

ORDER OF MAGNITUDE PHYSICS
A TEXTBOOK
WITH APPLICATIONS TO THE RETINAL ROD
AND TO THE DENSITY OF PRIME NUMBERS

Thesis by
Sanjoy Mahajan

In partial fulfillment of the requirements
for the degree of
Doctor of Philosophy



California Institute of Technology
Pasadena, California
1998
(Submitted 19 May 1998)

ACKNOWLEDGMENTS

I have had wonderful teachers. John Hopfield suggested that a textbook could be part of a thesis, and always shared the joy that he finds in physics. Sterl Phinney (my advisor) and Peter Goldreich taught me order-of-magnitude physics. Carver Mead shared his great scientific insight and his gentle personal wisdom. Steve Frautschi encouraged me to lead a section of Physics 1a, and generously gave me guidance on this unusual thesis project.

David Middlebrook taught me that to calculate is human, but to approximate is divine. Henry Lester, and the Computation and Neural Systems program, introduced me to biophysics. Lyn Dupré copyedited the parts of this dissertation with no mistakes, and, in the process, trained my ear. Erik Winfree showed me what clear thinking means. Joshua Zucker guided my experiments in teaching. Margaret Lambert taught me how to play Bach; and Don Caldwell welcomed me into the choir, although I had never sung before.

Comments from Jennifer Linden, Dap Hartmann, and John Hopfield have made the textbook useful to a wider audience. Comments from Joseph Keane, Dinakar Ramakrishnan, Warren Smith, and Erik Winfree have made the primality model more rigorous.

Donald Knuth wrote the \TeX typesetting system. John Hobby wrote MetaPost, with which I drew most of the figures. Thousands of volunteers wrote the Linux operating system.

The Hertz Foundation supported me in my last three years of graduate study. The Whitaker Foundation supported me in my first year. I also received support from Carver Mead, John Hopfield, and William Goddard, all of whom have been my advisors.

Many people here have shared their friendship. I particularly want to mention Sarah Bates, Carlos Brody, Craig Jin, Elisabeth Moyer, Sam Roweis, Rahul Sarpeshkar, Micah Siegel, Steven and Athina Quake, and Erik Winfree. To my other friends, here and from the years BC (before Caltech), thank you as well. Laura Rodriguez, den mother of the Hopfield group, kept us sane and happy.

My brothers, Sunit and Ashish, and my parents, Sushma and Subhash, have given me a home.

My thanks to all

ABSTRACT

I develop tools to amplify our mental senses: our intuition and reasoning abilities. The first five chapters—based on the Order of Magnitude Physics class taught at Caltech by Peter Goldreich and Sterl Phinney—form part of a textbook on dimensional analysis, approximation, and physical reasoning. The text is a resource of intuitions, problem-solving methods, and physical interpretations. By avoiding mathematical complexity, order-of-magnitude techniques increase our physical understanding, and allow us to study otherwise difficult or intractable problems. The textbook covers: (1) simple estimations, (2) dimensional analysis, (3) mechanical properties of materials, (4) thermal properties of materials, and (5) water waves.

As an extended example of order-of-magnitude methods, I construct an analytic model for the flash sensitivity of a retinal rod. This model extends the flash-response model of Lamb and Pugh with an approximate model for steady-state response as a function of background light I_b . The combined model predicts that the flash sensitivity is proportional to $I_b^{-1.3}$. This result roughly agrees with experimental data, which show that the flash sensitivity follows the Weber–Fechner behavior of I_b^{-1} over an intensity range of 100. Because the model is simple, it shows clearly how each biochemical pathway determines the rod’s response.

The second example is an approximate model of primality, the square-root model. Its goal is to explain features of the density of primes. In this model, which is related to the Hawkins’ random sieve, divisibility and primality are probabilistic. The model implies a recurrence for the probability that a number n is prime. The asymptotic solution to the recurrence is $(\log n)^{-1}$, in agreement with the prime-number theorem. The next term in the solution oscillates around $(\log n)^{-1}$ with a period that grows superexponentially. These oscillations are a model for oscillations in the density of actual primes first demonstrated by Littlewood, who showed that the number of primes $\leq n$ crosses its natural approximator, the logarithmic integral, infinitely often. No explicit crossing is known; the best theorem, due to de Rivecourt, says that the first crossing happens below 7×10^{370} . A consequence of the square-root model is the conjecture that the first crossing is near 10^{27} .

CONTENTS

1 Contributions	1
1.1 Textbook	1
1.2 Original research	2
1.3 Words to the reader	3
 <i>Part 1: Textbook</i>	
2 Wetting Your Feet	5
2.1 Warmup problems	5
2.2 Scaling analyses	16
2.3 What you have learned	25
3 Dimensional Analysis	27
3.1 Newton's law	27
3.2 Pendula	31
3.3 Drag in fluids	36
3.4 What you have learned	47
4 Materials I	49
4.1 Sizes	49
4.2 Energies	58
4.3 Elastic properties	60
4.4 Application to white dwarfs	68
4.5 What you have learned	71
5 Materials II	72
5.1 Thermal expansion	72
5.2 Phase changes	74
5.3 Specific heat	82
5.4 Thermal diffusivity of liquids and solids	86
5.5 Diffusivity and viscosity of gases	88
5.6 Thermal conductivity	89
5.7 What you have learned	92
6 Waves	93
6.1 Dispersion relations	93
6.2 Deep water	97
6.3 Shallow water	115
6.4 Combining deep- and shallow-water gravity waves	119
6.5 Combining deep- and shallow-water ripples	120
6.6 Combining all the analyses	121
6.7 What we did	121
 <i>Part 2: Applications</i>	
7 Retinal rod	122
7.1 Problem that the rod solves	122

7.2	Scope of this analysis	123
7.3	Flash response in darkness: The model of Lamb and Pugh	123
7.4	Steady-state response	126
7.5	Flash sensitivity	130
7.6	Conclusion	132
8	Density of primes	133
8.1	Approximate primality by stages	133
8.2	Convert the recurrence into a delay–differential equation	137
8.3	Solve the delay–differential equation	140
8.4	Analyze the oscillations	146
8.5	Speculate on the Riemann hypothesis	149
8.6	Go beyond the square root	149
8.7	Lower the temperature	150
8.8	Finish	151
	Bibliography	152

1 CONTRIBUTIONS

Human hands are weak, and human arms are short. Tools—such as the wedge, the crane, and the waterwheel—strengthen our grasp and lengthen our reach. To extend our mind’s reach, we need mental tools. This dissertation is a resource of such tools. It is in two parts: a textbook on order-of-magnitude physics, and original research in biophysics and number theory that illustrates order-of-magnitude methods.

In an order-of-magnitude analysis, we approximate because “an approximate answer is not only almost as useful as the exact answer, it’s more useful.” An approximate answer, because it is simple, shows us the important principles in a physical problem. Thirty thousand pages of detailed velocity and temperature tables might give us accurate knowledge of the winds, but we would rather know that a hurricane with 100 mph winds is approaching from sea. An approximate model, because it is compact, fits in our limited brains; once it is in there, we can refine it, and reason from it to new problems.

No exact, analytic solutions exist for many problems, such as hurricane motion or river turbulence. Even when solutions exist, they are often so complicated that they teach us little about the physical principles important in the problem. It is these principles that have general value, because we can use them to understand new pieces of the world. Often, these principles remain hidden behind a maze of mathematical manipulations; the principles determine the scaling laws—the functional form of the solution—and the complex mathematics determines a dimensionless constant that multiplies the functional form. An order-of-magnitude analysis reverses the traditional approach: You forget about the dimensionless constant, and hope that it is close to unity. You make estimates using procedures no more complicated than high-school algebra and geometry. If your scaling law is grossly wrong, you find out quickly, and can begin looking for new physics to create a new model; if it is accurate, you know what physical principles to include in the next, more refined analysis, and you have added a physical picture to your mental toolbox.

The textbook teaches such techniques. You learn how to identify important factors, how to combine the factors into an approximation, and how to appraise and refine an approximation. The theme woven throughout the chapters is successive approximation, because all order-of-magnitude techniques are derivatives of this one, and because an order-of-magnitude analysis is itself a first approximation to a more accurate analysis. The biophysics chapter (Chapter 7) shows how to apply these methods to a problem more in-depth than a typical textbook problem. The number-theory chapter (Chapter 8) shows how to use order-of-magnitude methods to create a series of successively more accurate, yet still tractable, models of primality; and how to use more standard methods to solve these models.

1.1 Textbook

The textbook is based on the Physics 103 class taught at the California Institute of Technology by Peter Goldreich and Sterl Phinney. My contribution has been to make explicit

the methods of approximation, and to show how to use physical reasoning to complement the methods of dimensional analysis.

Many books—for example those by Bridgman [4], Pankhurst [45], and Ipsen [26]—describe the method of dimensional analysis: how to combine physical quantities into groups that have no units. This textbook begins earlier and ends later than most texts. Chapter 2 introduces methods to estimate everyday numbers, such as city budgets or annual diaper production. If dimensional analysis is the first method of approximation, these methods are the zeroth. Chapter 3 introduces dimensional analysis. The discussion goes beyond the discussion in most textbooks: You learn how to use physical reasoning to complete dimensional-analysis arguments. With a repertoire that includes simple estimation methods, dimensional analysis, and physical reasoning, we then study the mechanical and thermal properties of materials (Chapters 4 and 5); the calculation of boiling points (Section 5.2.1) is an extended example of successive approximation, a method that the text introduces in successively more complicated problems. The final textbook chapter studies the properties of water waves. It analyzes one function, the dispersion relation, in all its limits. The chapter combines dimensional analysis and successive approximation with a physical model of wave propagation: the slab model.

1.2 Original research

I apply this style of reasoning to a problem in biophysics, the flash sensitivity of a retinal rod; and to a problem in number theory, the distribution of primes.

1.2.1 Retinal Rod

Imagine that we want to analyze the flash sensitivity of a retinal rod, to determine how the sensitivity varies as a function of background light I_b . (The flash sensitivity is the maximum change in membrane current divided by the flash strength.) To make an accurate model, we would model the biochemistry with a bundle of nonlinear differential equations that contain a slew of rate and time constants. We would throw up our hands and integrate the equations numerically; if we set up the equations correctly, our simulation would match reality. We could then get a valuable check on the completeness of our experimental knowledge. But what insight would this model give us about how the rod works? We would not know what function the various biochemical pathways perform. We could explain the response to light by saying only, “It works as the computer says.”

Instead, Chapter 7 develops an analytic model of the flash sensitivity. This model compactly represents important features of rod biochemistry; it shows how negative feedback and cooperative binding compute a flash sensitivity that approximates the Weber–Fechner sensitivity of I_b^{-1} . The model predicts that the flash sensitivity is $I_b^{-\theta}$, where $\theta^{-1} = m + 1 + 1/n$. The m results from cooperativity of calcium—the feedback signal—on the enzyme guanylate cyclase; and the 1 results from cooperativity of calcium on the phosphodiesterase hydrolytic rate. The two feedback paths are driven by the forward path—cooperativity of cyclic GMP on channel gating—which produces the $1/n$.

1.2.2 Primes

Although their definition is simple, prime numbers exhibit exotic properties, particularly in their distribution. For example, primes thin out: The density of prime numbers near n

is roughly $(\log n)^{-1}$. The journey from the definition of primality to the deduction of this density is arduous. Gauss began the journey; Riemann nearly completed it; and Hadamard and de la Vallée Poussin, in the proofs of the prime-number theorem, finished it. The journey becomes more arduous if we want to prove more stringent asymptotic statements about the density of primes.

The purpose of order-of-magnitude methods is to identify important features of a system; to incorporate them into a tractable model; and, from that model, to deduce properties of the original system. Why not treat prime numbers with these methods? To analyze the distribution of primes, Chapter 8 makes a series of successively more accurate probabilistic models of divisibility. One model is similar to the Hawkins' random sieve [21, 22], and one, the square-root model, turns out to be an extension of it. For the density of primes, the square-root model produces a recurrence; after we transform the recurrence properly, it becomes a delay-differential equation. The deduction and solution of this equation occupy much of the chapter. These apparently non-order-of-magnitude parts of the analysis show how to use order-of-magnitude reasoning to begin an analysis and to generate models, and how to complete the analysis using more standard methods.

The model predicts that the density of primes oscillates around $(\log n)^{-1}$ with a period that grows superexponentially. I believe that these oscillations model oscillations in the density of actual primes: Littlewood [39] showed that $\pi(n)$ (the number of primes $\leq n$) crosses its natural approximator, the logarithmic integral $\text{Li}(n)$, infinitely often. Littlewood's proof is nonconstructive. We still do not know an explicit n for which $\pi(n) > \text{Li}(n)$; all explicit computations of $\pi(n)$ have produced $\pi(n) < \text{Li}(n)$. Te Riele [51] showed that the first crossing occurs below 7×10^{370} . The square-root model leads to the conjecture that the first crossing is near 10^{27} . In perhaps one decade, computers will be fast enough to determine whether this prediction is valid.

1.3 Words to the reader

Although I can verify complicated derivations, my limited brain cannot understand them sufficiently well to use in other problems. So I search for pictures and for methods simple enough to treat as a mental unit and compact enough to transport to other problems. This dissertation—which combines simple models, teaching, and research—has matched my interests. I hope that you enjoy the content and the informal style, which reflects the pleasure that I have found in this work.

I want the textbook to converge to one that is letter perfect, and is fun and instructive to read; your comments are welcome. You can send paper mail to

Sanjoy Mahajan
8824 S. Poplar St
Tempe, AZ 85284
USA

or email to one of (in likeliest-to-unlikeliest order)

`mahajan@alumni.cs.cmu.edu`
`sanjoy@hope.caltech.edu`
`sanjoy@neuron.princeton.edu`
`sanjoy@alumni.caltech.edu`

until those addresses stop working. For as long as I can, I will place on the World Wide Web revised, electronic versions of these textbook chapters, as well as of newly written ones. With luck, you can find them at one of the following locations:

`http://www.neuron.princeton.edu`
`http://hope.caltech.edu`
`http://www.alumni.caltech.edu`
`http://www.pcmp.caltech.edu`

Eventually, the textbook will be officially published; it will then be more widely available (though maybe no longer electronically, alas).

2 WETTING YOUR FEET

Most technical education emphasizes exact answers. If you are a physicist, you solve for the energy levels of the hydrogen atom to six decimal places. If you are a chemist, you measure reaction rates and concentrations to two or three decimal places. In this book, you learn complementary skills. You learn that an approximate answer is not merely good enough; it's often more useful than an exact answer. When you approach an unfamiliar problem, you want to learn first the main ideas and the important principles, because these ideas and principles structure your understanding of the problem. It is easier to refine this understanding than to create the refined analysis in one step.

The adjective in the title, **order of magnitude**, reflects our emphasis on approximation. An order of magnitude is a factor of 10. To be “within an order of magnitude,” or to estimate a quantity “to order of magnitude,” means that your estimate is roughly within a factor of 10 on either side. This chapter introduces the art of determining such approximations.

Writer's block is broken by writing; estimator's block is broken by estimating. So we begin our study of approximation using everyday examples, such as estimating budgets or annual production of diapers. These warmups flex your estimation muscles, which may have lain dormant through many years of traditional education. After the warmup, we introduce a more subtle method: scaling relations.

2.1 Warmup problems

Everyday estimations provide practice for our later problems, and also provide a method to sanity check information that you see. Suppose that a newspaper article says that the annual cost of health care in the United States will soon surpass \$1 trillion. Whenever you read any such claim, you should automatically think: Does this number seem reasonable? Is it far too small, or far too large? You need methods for such estimations, methods that we develop in several examples. We dedicate the first example to physicists who need employment outside of physics.

2.1.1 *Armored cars*

How much money is there in a fully loaded Brinks armored car?

The amount of money depends on the size of the car, the denomination of the bills, the volume of each bill, the amount of air between the bills, and many other factors. The question, at first glance, seems vague. One important skill that you will learn from this text, by practice and example, is what assumptions to make. Because we do not need an exact answer, any reasonable set of assumptions will do. Getting started is more important than dotting every *i*; make an assumption—*any* assumption—and begin. You can correct the gross lies after you have got a feeling for the problem, and have learned which assumptions are most critical. If you keep silent, rather than tell a gross lie, you never discover anything.

Let's begin with our equality conventions, in ascending order of precision. We use \propto for proportionalities, where the units on the left and right sides of the \propto do not match; for example, Newton's second law could read $F \propto m$. We use \sim for dimensionally correct relations (the units do match), which are often accurate to, say, a factor of 5 in either direction. An example is

$$\text{kinetic energy} \sim Mv^2. \quad (2.1)$$

Like the \propto sign, the \sim sign indicates that we've left out a constant; with \sim , the constant is dimensionless. We use \approx to emphasize that the relation is accurate to, say, 20 or 30 percent. Sometimes, \sim relations are also that accurate; the context will make the distinction.

Now we return to the armored car. How much money does it contain? Before you try a systematic method, take a guess. Make it an educated guess if you have some knowledge (perhaps you work for an insurance company, and you happened to write the insurance policy that the armored-car company bought); make it an uneducated guess if you have no knowledge. Then, after you get a more reliable estimate, compare it to your guess: The wonderful learning machine that is your brain magically improves your guesses for the next problem. You train your intuition, and, as we see at the end of this example, you aid your memory. As a pure guess, let's say that the armored car contains \$1 million.

Now we introduce a systematic method. A general method in many estimations is to break the problem into pieces that we can handle: We **divide and conquer**. The amount of money is large by everyday standards; the largeness suggests that we break the problem into smaller chunks, which we can estimate more reliably. If we know the volume V of the car, and the volume v of a US bill, then we can count the bills inside the car by dividing the two volumes, $N \sim V/v$. After we count the bills, we can worry about the denominations (divide and conquer again). [We do not want to say that $N \approx V/v$. Our volume estimates may be in error easily by 30 or 40 percent, or only a fraction of the storage space may be occupied by bills. We do not want to commit ourselves.¹]

We have divided the problem into two simpler subproblems: determining the volume of the car, and determining the volume of a bill. What is the volume of an armored car? The storage space in an armored car has a funny shape, with ledges, corners, nooks, and crannies; no simple formula would tell us the volume, even if we knew the 50-odd measurements. This situation is just the sort for which order-of-magnitude physics is designed; the problem is messy and underspecified. So we **lie skillfully**: We pretend that the storage space is a simple shape with a volume that we can find. In this case, we pretend that it is a rectangular prism (Figure 2.1).

To estimate the volume of the prism, we divide and conquer. We divide estimating the volume into estimating the three dimensions of the prism. The compound structure of the formula

$$V \sim \text{length} \times \text{width} \times \text{height} \quad (2.2)$$

1. "Once at a Fourth-of-July celebration, a reporter wondered and later asked why Mr. Murphy (he was always Mr. Murphy even to his closest associates) did not join in the singing of the National Anthem. 'Perhaps he didn't want to commit himself,' the boss's aide explained." From the Introduction by Arthur Mann, to William L. Riordan, *Plunkitt of Tammany Hall* (New York: E. P. Dutton, 1963), page ix.

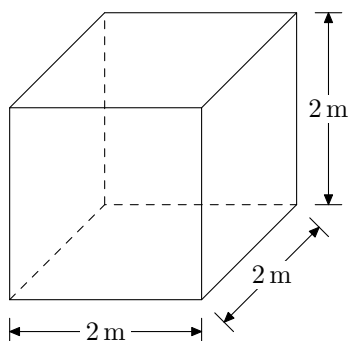


Figure 2.1. Interior of a Brinks armored car. The actual shape is irregular, but to order of magnitude, the interior is a cube. A person can probably lie down or stand up with room to spare, so we estimate the volume as $V \sim 2\text{ m} \times 2\text{ m} \times 2\text{ m} \sim 10\text{ m}^3$.

suggests that we divide and conquer. Probably an average-sized person can lie down inside with room to spare, so each dimension is roughly 2 m, and the interior volume is

$$V \sim 2\text{ m} \times 2\text{ m} \times 2\text{ m} \sim 10\text{ m}^3 = 10^7\text{ cm}^3. \quad (2.3)$$

In this text, $2 \times 2 \times 2$ is almost always 10. We are already working with crude approximations, which we signal by using \sim in $N \sim V/v$, so we do not waste effort in keeping track of a factor of 1.25 (from using 10 instead of 8). We converted the m^3 to cm^3 in anticipation of the dollar-bill-volume calculation: We want to use units that match the volume of a dollar bill, which is certainly much smaller than 1 m^3 .

Now we estimate the volume of a dollar bill (the volumes of US denominations are roughly the same). You can lay a ruler next to a dollar bill, or you can just guess that a bill measures 2 or 3 inches by 6 inches, or $6\text{ cm} \times 15\text{ cm}$. To develop your feel for sizes, guess first; then, if you feel uneasy, check your answer with a ruler. As your feel for sizes develops, you will need to bring out the ruler less frequently. How thick is the dollar bill? Now we apply another order-of-magnitude technique: **guerrilla warfare**. We take any piece of information that we can get.² What's a dollar bill? We lie skillfully and say that a dollar bill is just ordinary paper. How thick is paper? Next to the computer used to compose this textbook is a laser printer; next to the printer is a ream of laser printer paper. The ream (500 sheets) is roughly 5 cm thick, so a sheet of quality paper has thickness 10^{-2} cm . Now we have the pieces to compute the volume of the bill:

$$v \sim 6\text{ cm} \times 15\text{ cm} \times 10^{-2}\text{ cm} \sim 1\text{ cm}^3. \quad (2.4)$$

The original point of computing the volume of the armored car and the volume of the bill was to find how many bills fit into the car: $N \sim V/v \sim 10^7\text{ cm}^3/1\text{ cm}^3 = 10^7$. If the money is in \$20 bills, then the car would contain \$200 million.

2. "I seen my opportunities and I took 'em."—George Washington Plunkitt, of Tammany Hall, quoted by Riordan [52, page 3].

The bills could also be \$1 or \$1000 bills, or any of the intermediate sizes. We chose the intermediate size \$20, because it lies nearly halfway between \$1 and \$1000. You naturally object that \$500, not \$20, lies halfway between \$1 and \$1000. We answer that objection shortly. First, we pause to discuss a general method of estimating: **talking to your gut**. You often have to estimate quantities about which you have only meager knowledge. You can then draw from your vast store of implicit knowledge about the world—knowledge that you possess but cannot easily write down. You extract this knowledge by conversing with your gut; you ask that internal sensor concrete questions, and listen to the feelings that it returns. You already carry on such conversations for other aspects of life. In your native language, you have an implicit knowledge of the grammar; an incorrect sentence sounds funny to you, even if you do not know the rule being broken. Here, we have to estimate the denomination of bill carried by the armored car (assuming that it carries mostly one denomination). We ask ourselves, “How does an armored car filled with one-dollar bills sound?” Our gut, which knows the grammar of the world, responds, “It sounds a bit ridiculous. One-dollar bills are not worth so much effort; plus, every automated teller machine dispenses \$20 bills, so a \$20 bill is a more likely denomination.” We then ask ourselves, “How about a truck filled with thousand-dollar bills?” and our gut responds, “no, sounds way too big—never even seen a thousand-dollar bill, probably collectors’ items, not for general circulation.” After this edifying dialogue, we decide to guess a value intermediate between \$1 and \$1000.

We interpret “between” using a logarithmic scale, so we choose a value near the geometric mean, $\sqrt{1 \times 1000} \sim 30$. Interpolating on a logarithmic scale is more appropriate and accurate than is interpolating on a linear scale, because we are going to use the number in a chain of multiplications and divisions. Let’s check whether 30 is reasonable, by asking our gut about nearby estimates. It is noncommittal when asked about \$10 or \$100 bills; both sound reasonable. So our estimate of 30 is probably reasonable. Because there are no \$30 bills, we use a nearby actual denomination, \$20.

Assuming \$20 bills, we estimate that the car contains \$200 million, an amount much greater than our initial guess of \$1 million. Such a large discrepancy makes us suspicious of either the guess or this new estimate. We therefore **cross-check** our answer, by estimating the monetary value in another way. By finding another method of solution, we learn more about the domain. If our new estimate agrees with the previous one, then we gain confidence that the first estimate was correct; if the new estimate does not agree, it may help us to find the error in the first estimate.

We estimated the carrying capacity using the available space. How else could we estimate it? The armored car, besides having limited space, cannot carry infinite mass. So we estimate the mass of the bills, instead of their volume. What is the mass of a bill? If we knew the density of a bill, we could determine the mass using the volume computed in (2.4). To find the density, we use the guerrilla method. Money is paper. What is paper? It’s wood or fabric, except for many complex processing stages whose analysis is beyond the scope of this book. Here, we just used another order-of-magnitude technique, **punt**: When a process, such as papermaking, looks formidable, forget about it, and hope that you’ll be okay anyway. Ignorance is bliss. It’s more important to get an estimate; you can correct the egregiously inaccurate assumptions later. How dense is wood? Once again, use the guerrilla method: Wood barely floats, so its density is roughly that of water, $\rho \sim 1 \text{ g cm}^{-3}$. A bill,

which has volume $v \sim 1 \text{ cm}^3$, has mass $m \sim 1 \text{ g}$. And 10^7 cm^3 of bills would have a mass of $10^7 \text{ g} = 10 \text{ tons}$.³

This cargo is large. [Metric tons are 10^6 g ; English tons (may that measure soon perish) are roughly $0.9 \cdot 10^6 \text{ g}$, which, for our purposes, is also 10^6 g .] What makes 10 tons large? Not the number 10 being large. To see why not, consider these extreme arguments:

- In megatons, the cargo is 10^{-5} megatons, which is a tiny cargo because 10^{-5} is a tiny number.
- In grams, the cargo is 10^7 g , which is a gigantic cargo because 10^7 is a gigantic number.

You might object that these arguments are cheats, because neither grams nor megatons is a reasonable unit in which to measure truck cargo, whereas tons is a reasonable unit. This objection is correct; when you specify a reasonable unit, you implicitly choose a standard of comparison. The moral is this: A quantity with units—such as tons—cannot be large intrinsically. It must be large compared to a quantity with the same units. This argument foreshadows the topic of dimensional analysis, which is the subject of Chapter 3.

So we must compare 10 tons to another mass. We could compare it to the mass of a bacterium, and we would learn that 10 tons is relatively large; but to learn about the cargo capacity of Brinks armored cars, we should compare 10 tons to a mass related to transport. We therefore compare it to the mass limits at railroad crossings and on many bridges, which are typically 2 or 3 tons. Compared to this mass, 10 tons is large. Such an armored car could not drive many places. Perhaps 1 ton of cargo is a more reasonable estimate for the mass, corresponding to 10^6 bills. We can cross-check this cargo estimate using the size of the armored car's engine (which presumably is related to the cargo mass); the engine is roughly the same size as the engine of a medium-sized pickup truck, which can carry 1 or 2 tons of cargo (roughly 20 or 30 book boxes—see Example 4.1). If the money is in \$20 bills, then the car contains \$20 million. Our original, pure-guess estimate of \$1 million is still much smaller than this estimate by roughly an order of magnitude, but we have more confidence in this new estimate, which lies roughly halfway between \$1 million and \$200 million (we find the midpoint on a logarithmic scale). The Reuters newswire of 18 September 1997 has a report on the largest armored car heist in US history; the thieves took \$18 million; so our estimate is accurate for a well-stocked car. (Typical heists net between \$1 million and \$3 million.)

We answered this first question in detail to illustrate a number of order-of-magnitude techniques. We saw the value of lying skillfully—approximating dollar-bill paper as ordinary paper, and ordinary paper as wood. We saw the value of waging guerrilla warfare—using knowledge that wood barely floats to estimate the density of wood. We saw the value of cross-checking—estimating the mass and volume of the cargo—to make sure that we have not committed a gross blunder. And we saw the value of divide and conquer—breaking volume estimations into products of length, width, and thickness. Breaking problems into factors, besides making the estimation possible, has another advantage: It often reduces the error in the estimate. There probably is a general rule about guessing, that the *logarithm* is in error by a reasonably fixed fraction. If we guess a number of the order of 1 billion in one

3. It is unfortunate that *mass* is not a transitive verb in the way that *weigh* is. Otherwise, we could write that the truck masses 10 tons. If you have more courage than we have, use this construction anyway, and start a useful trend.

step, we might be in error by, say, a factor of 10. If we factor the 1 billion into four pieces, the estimate of each piece will be in error by a factor of $\gamma = 10^{1/4}$. We then can hope that the errors are uncorrelated, so that they combine as steps in a random walk. Then, the error in the product is $\gamma^{\sqrt{4}} = 10^{1/2}$, which is smaller than the one-shot error of 10. So breaking an estimate into pieces reduces the error, according to this order-of-magnitude analysis of error.

2.1.2 Cost of lighting Pasadena, California

What is the annual cost of lighting the streets of Pasadena, California?

Astronomers would like this cost to be huge, so that they could argue that street lights should be turned off at night, the better to gaze at heavenly bodies. As in Section 2.1.1, we guess a cost right away, to train our intuition. So let's guess that lighting costs \$1 million annually. This number is unreliable; by talking to our gut, we find that \$100,000 sounds okay too, as does \$10 million (although \$100 million sounds too high).

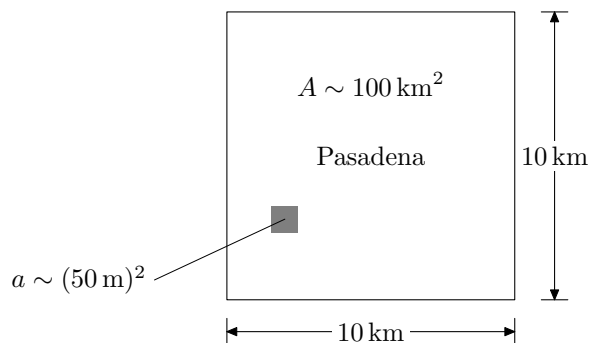


Figure 2.2. Map of Pasadena, California drawn to order of magnitude. The small shaded box is the area governed by one lamp; the box is not drawn to scale, because if it were, it would be only a few pixels wide. How many such boxes can we fit into the big square? It takes 10 min to leave Pasadena by car, so Pasadena has area $A \sim (10 \text{ km})^2 = 10^8 \text{ m}^2$. While driving, we pass a lamp every 3 s, so we estimate that there's a lamp every 50 m; each lamp covers an area $a \sim (50 \text{ m})^2$.

The cost is a large number, out of the ordinary range of costs, so it is difficult to estimate in one step (we just tried to guess it, and we're not sure within a factor of 10 what value is correct). So we divide and conquer. First, we estimate the number of lamps; then, we estimate how much it costs to light each lamp.

To estimate the number of lamps (another large, hard-to-guess number), we again divide and conquer: We estimate the area of Pasadena, and divide it by the area that each lamp governs, as shown in Figure 2.2. There is one more factor to consider: the fraction of the land that is lighted (we call this fraction f). In the desert, f is perhaps 0.01; in a typical city, such as Pasadena, f is closer to 1.0. We first assume that $f = 1.0$, to get an initial estimate; then we estimate f , and correct the cost accordingly.

We now estimate the area of Pasadena. What is its shape? We could look at a map, but, as lazy armchair theorists, we lie; we assume that Pasadena is a square. It takes, say,

10 minutes to leave Pasadena by car, perhaps traveling at 1 km/min; Pasadena is roughly 10 km in length. Therefore, Pasadena has area $A \sim 10 \text{ km} \times 10 \text{ km} = 100 \text{ km}^2 = 10^8 \text{ m}^2$. (The true area is 23 mi^2 , or 60 km^2 .) How much area does each lamp govern? In a car—say, at 1 km/min or $\sim 20 \text{ m s}^{-1}$ —it takes 2 or 3 s to go from lamppost to lamppost, corresponding to a spacing of $\sim 50 \text{ m}$. Therefore, $a \sim (50 \text{ m})^2 \sim 2.5 \cdot 10^3 \text{ m}^2$, and the number of lights is $N \sim A/a \sim 10^8 \text{ m}^2 / 2.5 \cdot 10^3 \text{ m}^2 \sim 4 \cdot 10^4$.

How much does each lamp cost to operate? We estimate the cost by estimating the energy that they consume in a year and the price per unit of energy (divide and conquer). Energy is power \times time. We can estimate power reasonably accurately, because we are familiar with lamps around the home. To estimate a quantity, try to compare it to a related, familiar one. Street lamps shine brighter than a household 100 W bulb, but they are probably more efficient as well, so we guess that each lamp draws $p \sim 300 \text{ W}$. All N lamps consume $P \sim Np \sim 4 \cdot 10^4 \times 300 \text{ W} \sim 1.2 \cdot 10^4 \text{ kW}$. Let's say that the lights are on at night—8 hours per day—or 3000 hours/year. Then, they consume $4 \cdot 10^7 \text{ kW-hour}$. An electric bill will tell you that electricity costs \$0.08 per kW-hour (if you live in Pasadena), so the annual cost for all the lamps is \$3 million.

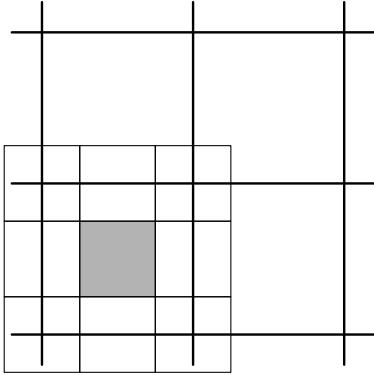


Figure 2.3. Fraction of Pasadena that is lighted. The streets (thick lines) are spaced $d \sim 100 \text{ m}$ apart. Each lamp, spaced 50 m apart, lights a $50 \text{ m} \times 50 \text{ m}$ area (the eight small, unshaded squares). The only area not lighted is in the center of the block (shaded square); it is one-fourth of the area of the block. So, if every street has lights, $f = 0.75$.

Now let's improve this result by estimating the fraction f . What features of Pasadena determine the value of f ? To answer this question, consider two extreme cases: the desert and New York city. In the desert, f is small, because the streets are widely separated, and many streets have no lights. In New York city, f is high, because the streets are densely packed, and most streets are filled with street lights. So the relevant factors are the spacing between streets (which we call d), and the fraction of streets that are lighted (which we call f_1). As all pedestrians in New York city know, 10 north-south blocks or 20 east-west blocks make 1 mile (or 1600 m); so $d \sim 100 \text{ m}$. In street layout, Pasadena is closer to New York city than to the desert. So we use $d \sim 100 \text{ m}$ for Pasadena as well. If every street were lighted, what fraction of Pasadena would be lighted? Figure 2.3 shows the computation; the

result is $f \sim 0.75$. In New York city, $f_L \sim 1$; in Pasadena, $f_L \sim 0.3$ is more appropriate. So $f \sim 0.75 \times 0.3 \sim 0.25$. Our estimate for the annual cost is then \$1 million. Our initial guess is unexpectedly accurate.

As you practice such estimations, you will be able to write them down compactly, converting units stepwise until you get to your goal (here, \$/year). The cost is

$$\begin{aligned} \text{cost} &\sim \underbrace{100 \text{ km}^2}_A \times \frac{10^6 \text{ m}^2}{1 \text{ km}^2} \times \underbrace{\frac{1 \text{ lamp}}{2.5 \cdot 10^3 \text{ m}^2}}_a \times \underbrace{\frac{8 \text{ hrs}}{1 \text{ day}}}_{\text{night}} \\ &\quad \times \frac{365 \text{ days}}{1 \text{ year}} \times \underbrace{\frac{\$0.08}{1 \text{ kW-hour}}}_{\text{price}} \times 0.3 \text{ kW} \times 0.25 \end{aligned} \quad (2.5)$$

$\sim \$1 \text{ million.}$

It is instructive to do the arithmetic without using a calculator. Just as driving to the neighbors' house atrophies your muscles, using calculators for simple arithmetic dulls your mind. You do not develop an innate sense of how large quantities should be, or of when you have made a mistake; you learn only how to punch keys. If you need an answer with 6-digit precision, use a calculator; that's the task for which they are suited. In order-of-magnitude estimates, 1- or 2-digit precision is sufficient; you can easily perform these low-precision calculations mentally.

Will Pasadena astronomers rejoice because this cost is large? A cost has units (here, dollars), so we must compare it to another, relevant cost. In this case, that cost is the budget of Pasadena. If lighting is a significant fraction of the budget, then can we say that the lighting cost is large.

2.1.3 Pasadena's budget

What fraction of Pasadena's budget is allotted to street lighting?

We just estimated the cost of lighting; now we need to estimate Pasadena's budget. First, however, we make the initial guess. It would be ridiculous if such a trivial service as street lighting consumed as much as 10 percent of the city's budget. The city still has road construction, police, city hall, and schools to support. 1 percent is a more reasonable guess. The budget should be roughly \$100 million.

Now that we've guessed the budget, how can we estimate it? The budget is the amount spent. This money must come from somewhere (or, at least, most of it must): Even the US government is moderately subject to the rule that income \approx spending. So we can estimate spending by estimating income. Most US cities and towns bring in income from property taxes. We estimate the city's income by estimating the property tax per person, and multiplying the tax by the city's population.

Each person pays property taxes either directly (if she owns land) or indirectly (if she rents from someone who does own land). A typical monthly rent per person (for a two-person apartment) is \$500 in Pasadena (the apartments-for-rent section of a local newspaper will tell you the rent in your area), or \$6000 per year. (Places with fine weather and less smog, such as the San Francisco area, have higher monthly rents, roughly \$1500 per person.) According to occasional articles that appear in newspapers when rent skyrockets and interest

in the subject increases, roughly 20 percent of rent goes toward landlords' property taxes. We therefore estimate that \$1000 is the annual property tax per person.

Pasadena has roughly $2 \cdot 10^5$ people, as stated on the road signs that grace the entries to Pasadena. So the annual tax collected is \$200 million. If we add federal subsidies to the budget, the total budget is probably double that, or \$400 million. A rule of thumb in these financial calculations is to double any estimate that you make, to correct for costs or revenues that you forgot to include. This rule of thumb is not infallible. We can check its validity in this case by estimating the federal contribution. The federal budget is roughly \$2 trillion, or \$6000 for every person in the United States (any recent almanac tells us the federal budget and the US population). One-half of the \$6000 funds defense spending and interest on the national debt; it would be surprising if fully one-half of the remaining \$3000 went to the cities. Perhaps \$1000 per person goes to cities, which is roughly the amount that the city collects from property taxes. Our doubling rule is accurate in this case.

For practice, we cross-check the local-tax estimate of \$200 million, by estimating the total land value in Pasadena, and guessing the tax rate. The area of Pasadena is $100 \text{ km}^2 \sim 36 \text{ mi}^2$, and $1 \text{ mi}^2 = 640 \text{ acres}$. You can look up this acre-square-mile conversion, or remember that, under the Homestead Act, the US government handed out land in 160-acre parcels—known as *quarter lots* because they were 0.25 mi^2 . Land prices are exorbitant in southern California (sun, sand, surf, and mountains, all within a few hours drive); the cost is roughly \$1 million per acre (as you can determine by looking at the homes-for-sale section of the newspaper). We guess that property tax is 1 percent of property value. You can determine a more accurate value by asking anyone who owns a home, or by asking City Hall. The total tax is

$$\begin{aligned}
 W &\sim \underbrace{36 \text{ mi}^2}_{\text{area}} \times \frac{640 \text{ acres}}{1 \text{ mi}^2} \times \underbrace{\frac{\$1 \text{ million}}{1 \text{ acre}}}_{\text{land price}} \times \underbrace{0.01}_{\text{tax}} \\
 &\sim \$200 \text{ million.}
 \end{aligned} \tag{2.6}$$

This revenue is identical to our previous estimate of local revenue; the equality increases our confidence in the estimates. As a check on our estimate, we looked up the budget of Pasadena. In 1990, it was \$350 million; this value is much closer to our estimate of \$400 million than we have a right to expect!

The cost of lighting, calculated in Section 2.1.2, consumes only 0.2 percent of the city's budget. Astronomers should not wait for Pasadena to turn out the lights.

2.1.4 Diaper production

How many disposable diapers are manufactured in the United States every year?

We begin with a guess. The number has to be at least in the millions—say, 10 million—because of the huge outcry when environmentalists suggested banning disposable diapers to conserve landfill space and to reduce disposed plastic. To estimate such a large number, we divide and conquer. We estimate the number of diaper users—babies, assuming that all babies use diapers, and that no one else does—and the number of diapers that each baby uses in 1 year. These assumptions are not particularly accurate, but they provide a start for our estimation. How many babies are there? We hereby define a baby as a child under

2 years of age. What fraction of the population are babies? To estimate this fraction, we begin by assuming that everyone lives exactly 70 years—roughly the life expectancy in the United States—and then abruptly dies. (The life expectancy is more like 77 years, but an error of 10 percent is not significant given the inaccuracies in the remaining estimates.)

How could we have figured out the average age, if we did not already know it? In the United States, government retirement (Social Security) benefits begin at age 65 years, the canonical retirement age. If the life expectancy were less than 65 years—say, 55 years—then so many people would complain about being short-changed by Social Security that the system would probably be changed. If the life expectancy were much longer than 65 years—say, if it were 90 years—then Social Security would cost much more: It would have to pay retirement benefits for $90 - 65 = 25$ years instead of for $75 - 65 = 10$ years, a factor of 2.5 increase. It would have gone bankrupt long ago. So, the life expectancy must be around 70 or 80 years; if it becomes significantly longer, expect to see the retirement age increased accordingly. For definiteness, we choose one value: 70 years. Even if 80 years is a more accurate estimate, we would be making an error of only 15 percent, which is probably smaller than the error that we made in guessing the cutoff age for diaper use. It would hardly improve the accuracy of the final estimate to agonize over this 15 percent.

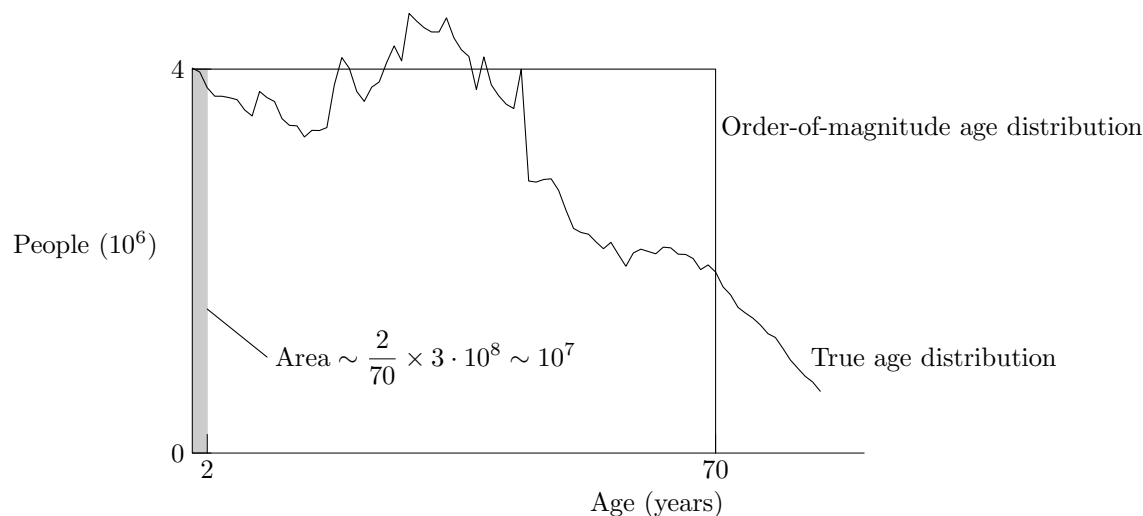


Figure 2.4. Number of people versus age (in the United States). The true age distribution is irregular and messy; without looking it up, we cannot know the area between ages 0.0 years and 2.0 years (to estimate the number of babies). The rectangular graph—which has the same area and similar width—immediately makes clear what the fraction under 2 years is: It is roughly $2/70 \sim 0.03$. The population of the United States is roughly $3 \cdot 10^8$, so the number of babies is $\sim 0.03 \times 3 \cdot 10^8 \sim 10^7$.

To compute how people are between the ages of 0 and 2.0 years, consider an analogous problem. In a 4-year university (which graduates everyone in 4 years and accepts no transfer students) with 1000 students, how many students graduate in each year's class? The

answer is 250, because $1000/4 = 250$. We can translate this argument into the following mathematics. Let τ be lifetime of a person. We assume that the population is steady: The birth and death rates are equal. Let the rates be \dot{N} . Then the total population is $N = \dot{N}\tau$, and the population between ages τ_1 and τ_2 is

$$N \frac{\tau_2 - \tau_1}{\tau} = \dot{N}(\tau_2 - \tau_1). \quad (2.7)$$

So, if everyone lives for 70 years exactly, then the fraction of the population whose age is between 0 and 2 years is $2/70$ or ~ 0.03 (Figure 2.4). There are roughly $3 \cdot 10^8$ people in the United States, so

$$N_{\text{babies}} \sim 3 \cdot 10^8 \times 0.03 \sim 10^7 \text{ babies.} \quad (2.8)$$

We have just seen another example of skillful lying. The jagged curve in Figure 2.4 shows a cartoon version of the actual mortality curve for the United States. We simplified this curve into the boxcar shape (the rectangle), because we know how to deal with rectangles. Instead of integrating the complex, jagged curve, we integrate a simple, civilized curve: a rectangle of the same area and similar width. This procedure is **order-of-magnitude integration**. Similarly, when we studied the Brinks armored-car example (Section 2.1.1), we pretended that the cargo space was a cube; that procedure was **order-of-magnitude geometry**.

How many diapers does each baby use per year? This number is large—maybe 100, maybe 10,000—so a wild guess is not likely to be accurate. We divide and conquer, dividing 1 year into 365 days. Suppose that each baby uses 8 diapers per day; newborns use many more, and older toddlers use less; our estimate is a reasonable compromise. Then, the annual use per baby is ~ 3000 , and all 10^7 babies use $3 \cdot 10^{10}$ diapers. The actual number manufactured is $1.6 \cdot 10^{10}$ per year, so our initial guess is low, and our systematic estimate is high.

This example also illustrates how to deal with **flows**: People move from one age to the next, leaving the flow (dying) at different ages, on average at age 70 years. From that knowledge alone, it is difficult to estimate the number of children under age 2 years; only an actuarial table would give us precise information. Instead, we invent a table that makes the calculation simple: Everyone lives to the life expectancy, and then dies abruptly. The calculation is simple, and the approximation is at least as accurate as the approximation that every child uses diapers for exactly 2 years. In a product, the error is dominated by the most uncertain factor; you waste your time if you make the other factors more accurate than the most uncertain factor.

2.1.5 Meteorite impacts

How many large meteorites hit the earth each year?

This question is not yet clearly defined: What does *large* mean? When you explore a new field, you often have to estimate such ill-defined quantities. The real world is messy. You have to constrain the question before you can answer it. After you answer it, even with crude approximations, you will understand the domain more clearly, will know which constraints were useful, and will know how to improve them. If your candidate set of assumptions produce a wildly inaccurate estimate—say, one that is off by a factor of 100,000—then you can be sure that your assumptions contain a fundamental flaw. Solving such an inaccurate

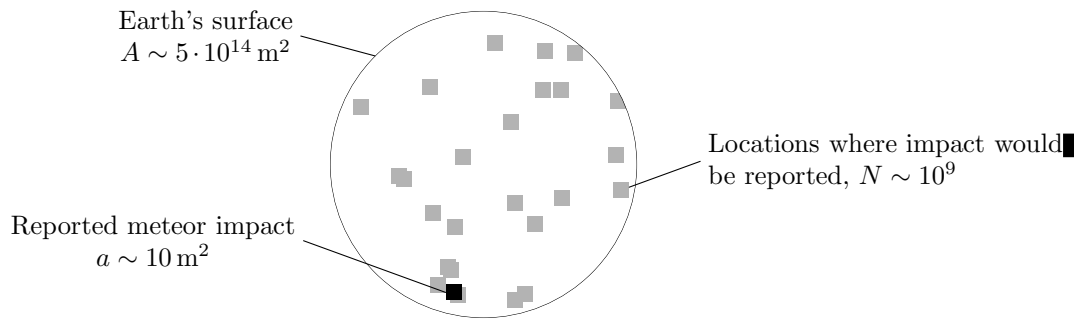


Figure 2.5. Large-meteorite impacts on the surface of the earth. Over the surface of the earth, represented as a circle, every year one meteorite impact (black square) causes sufficient damage to be reported by *Sky&Telescope*. The gray squares are areas where such a meteorite impact would have been reported—for example, a house or car in an industrial country; they have total area $Na \sim 10^{10} \text{ m}^2$. The gray squares cover only a small fraction of the earth's surface. The expected number of large impacts over the whole earth is $1 \times A/Na \sim 5 \cdot 10^4$, where $A \sim 5 \cdot 10^{14} \text{ m}^2$ is the surface area of the earth.

model exactly is a waste of your time. An order-of-magnitude analysis can prevent this waste, saving you time to create more realistic models. After you are satisfied with your assumptions, you can invest the effort to refine your model.

Sky&Telescope magazine reports approximately one meteorite impact per year. However, we cannot simply conclude that only one large meteorite falls each year, because *Sky&Telescope* presumably does not report meteorites that land in the ocean or in the middle of corn fields. We must adjust this figure upward, by a factor that accounts for the cross-section (effective area) that *Sky&Telescope* reports cover (Figure 2.5). Most of the reports cite impacts on large, expensive property such as cars or houses, and are from industrial countries, which have $N \sim 10^9$ people. How much target area does each person's car and living space occupy? Her car may occupy 4 m^2 , and her living space (portion of a house or apartment) may occupy 10 m^2 . [A country dweller living in a ranch house presents a larger target than 10 m^2 , perhaps 30 m^2 . A city dweller living in an apartment presents a smaller target than 10 m^2 , as you can understand from the following argument. Assume that a meteorite that lands in a city crashes through 10 stories. The target area is the area of the building roof, which is one-tenth the total apartment area in the building. In a city, perhaps 50 m^2 is a typical area for a two-person apartment, and 3 m^2 is a typical target area per person. Our estimate of 10 m^2 is a compromise between the rural value of 30 m^2 and the city value of 3 m^2 .]

Because each person presents a target area of $a \sim 10 \text{ m}^2$, the total area covered by the reports is $Na \sim 10^{10} \text{ m}^2$. The surface area of the earth is $A \sim 4\pi \times (6 \cdot 10^6 \text{ m})^2 \sim 5 \cdot 10^{14} \text{ m}^2$, so the reports of one impact per year cover a fraction $Na/A \sim 2 \cdot 10^{-5}$ of the earth's surface. We multiply our initial estimate of impacts by the reciprocal, A/Na , and estimate $5 \cdot 10^4$ large-meteorite impacts per year. In the solution, we defined large implicitly, by the criteria that *Sky&Telescope* use.

2.2 Scaling analyses

In most of the previous examples, we used opportunistic tricks to determine what numbers to multiply together. We now introduce a new method, **scaling**, for problems where simple multiplication is not sufficient. Instead of explaining what a scaling argument is, we first make one, and then explain what we did. The fastest way to learn a language is to hear and speak it. Physics is no exception; you hear it in the examples, and you speak it in the exercises.

2.2.1 Gravity on the moon

What is acceleration due to gravity on the surface of the moon?

First, we guess. Should it be 1 cm s^{-2} , or 10^6 cm s^{-2} , or 10^3 cm s^{-2} ? They all sound reasonable, so we make the guess of least resistance—that everywhere is like our local environment—and say that $g_{\text{moon}} \sim g_{\text{earth}}$, which is 1000 cm s^{-2} . Now we make a systematic estimate.

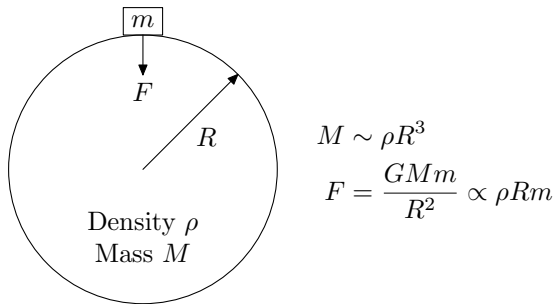


Figure 2.6. Order-of-magnitude astronomical body. The body—taken to be a sphere—has uniform density ρ and radius R . A block of mass m sits on the surface and feels a gravitational force $F = GMm/R^2$, where $M \sim \rho R^3$ is the mass of the astronomical body. The resulting acceleration is $g = F/m = G\rho R \propto \rho R$; if ρ is the same for all astronomical bodies in which we’re interested, then $g \propto R$.

This method that we use eventually shows you how to make estimates without knowing physical constants, such as the gravitational constant G . First, we give the wrong solution, so that we can contrast it with the right—and simpler—order-of-magnitude solution. The acceleration due to gravity at the surface of the moon is given by Newton’s law of gravitation (Figure 2.6):

$$g = \frac{F}{m} = \frac{GM}{R^2}. \quad (2.9)$$

In the wrong way, we look up—perhaps in the thorough and useful *CRC Handbook of Chemistry and Physics* [38]— M and R for the moon, and the fundamental constant G , and get

$$g_{\text{moon}} \sim \frac{6.7 \cdot 10^{-8} \text{ cm}^3 \text{ g}^{-1} \text{ s}^{-2} \times 7.3 \cdot 10^{25} \text{ g}}{(1.7 \cdot 10^8 \text{ cm})^2} \sim 160 \text{ cm s}^{-2}. \quad (2.10)$$

Here is another arithmetic calculation that you can do mentally, perhaps saying to yourself, “First, I count the powers of 10: There are 17 ($-8 + 25$) powers of 10 in the numerator,

and $16 (8+8)$ in the denominator, leaving 1 power of 10 after the division. Then, I account for the prefactors, ignoring the factors of 10. The numerator contains 6.7×7.3 , which is roughly $7 \times 7 = 49$. The denominator contains $1.7^2 \sim 3$. Therefore, the prefactors produce $49/3 \sim 16$. When we include one power of 10, we get 160.”

This brute-force method—looking up every quantity and then doing arithmetic—is easy to understand, and is a reasonable method for an initial solution. However, it is not instructive. For example, when you compare $g_{\text{moon}} \sim 160 \text{ cm s}^{-2}$ with g_{earth} , you may notice that g_{moon} is smaller than g_{earth} by a factor of only ~ 6 . With the huge numbers that we multiplied and divided in (2.10), g_{moon} could easily have been 0.01 cm s^{-2} or 10^6 cm s^{-2} . Why are g_{moon} and g_{earth} nearly the same, different by a mere factor of 6? The brute-force method shows only that huge numbers among G , M , and R^2 nearly canceled out to produce the moderate acceleration 160 cm s^{-2} .

So we try a more insightful method, which has the benefit that we do not have to know G ; we have to know only g_{earth} . This method is not as accurate as the brute-force method, but it will teach us more physics. It is an example of how approximate answers can be more useful than exact answers.

We find g_{moon} for the moon by scaling it against g_{earth} . [It is worth memorizing g_{earth} , because so many of our estimations depend on its value.] We begin with (2.9). Instead of M and R , we use density ρ and radius R as the independent variables; we lose no information, because we can compute density from mass and radius (assuming, as usual, that the astronomical body has the simplest shape: a sphere). We prefer density to mass, because density and radius are more orthogonal than mass and radius. In a thought experiment—and order-of-magnitude analyses are largely thought experiments—we might imagine a larger moon made out of the same type of rock. Enlarging the moon changes both M and R , but leaves ρ alone. To keep M fixed while changing R requires a larger feat of imagination (we shatter the moon and use scaffolding to hold the fragments at the right distance apart).

For a sphere of constant density, $M = (4\pi/3)\rho R^3$, so (2.9) becomes

$$g \propto \rho R. \quad (2.11)$$

This **scaling relation** tells us how g varies—scales—with density and radius. We retain only those variables and factors that change from the earth to the moon; the proportionality sign \propto allows us to eliminate constants such as G , and numerical factors such as $4\pi/3$.

If the earth and moon have the same radius and the same average density of rock, then we can further simplify (2.11) by eliminating ρ and R to get $g \propto 1$. These assumptions are not accurate, but they simplify the scaling relation; we correct them shortly. So, in this simple model, g_{moon} and g_{earth} are equal, which partially explains the modest factor of 6 that separates g_{moon} and g_{earth} . Now that we roughly understand the factor of 6, as a constant near unity, we strive for more accuracy, and remove the most inaccurate approximations. The first approximation to correct is the assumption that the earth and moon have the same radius. If R can be different on the earth and moon, then (2.11) becomes $g \propto R$, whereupon $g_{\text{earth}}/g_{\text{moon}} \sim R_{\text{earth}}/R_{\text{moon}}$.

What is R_{moon} ? Once again, we apply the guerrilla method. When the moon is full, a thumb held at arms length will just cover the moon perceived by a human eye. For a typical human-arm length of 100 cm, and a typical thumb width of 1 cm, the angle subtended is

$\theta \sim 0.01$ rad. The moon is $L \sim 4 \cdot 10^{10}$ cm from the earth, so its diameter is $\theta L \sim 0.01L$; therefore, $R_{\text{moon}} \sim 2 \cdot 10^8$ cm. By contrast, $R_{\text{earth}} \sim 6 \cdot 10^8$ cm, so $g_{\text{earth}}/g_{\text{moon}} \sim 3$. We have already explained a large part of the factor of 6. Before we explain the remainder, let's estimate L from familiar parameters of the moon's orbit. One of the goals of order-of-magnitude physics is to show you that you can make many estimates with the knowledge that you already have. Let's apply this philosophy to estimating the radius of the moon's orbit. One familiar parameter is the period: $T \sim 30$ days. The moon orbits in a circle because of the earth's gravitational field. What is the effect of earth's gravity at distance L (from the center of the earth)? At distance R_{earth} from the center of the earth, the acceleration due to gravity is g ; at L , it is $a = g(R_{\text{earth}}/L)^2$, because gravitational force (and, therefore, acceleration) are proportional to distance⁻². The acceleration required to move the moon in a circle is v^2/L . In terms of the period, which we know, this acceleration is $a = (2\pi L/T)^2/L$. So

$$\underbrace{g \left(\frac{R_{\text{earth}}}{L} \right)^2}_{a_{\text{gravity}}} = \underbrace{\left(\frac{2\pi L}{T} \right)^2 \frac{1}{L}}_{a_{\text{required}}} \quad (2.12)$$

The orbit radius is

$$\begin{aligned} L &= \left(\frac{g R_{\text{earth}}^2 T^2}{4\pi^2} \right)^{1/3} \\ &\sim \left(\frac{1000 \text{ cm s}^{-2} \times (6 \cdot 10^8 \text{ cm})^2 \times (3 \cdot 10^6 \text{ s})^2}{40} \right)^{1/3} \\ &\sim 5 \cdot 10^{10} \text{ cm}, \end{aligned} \quad (2.13)$$

which closely matches the actual value of $4 \cdot 10^{10}$ cm.

Now we return to explaining the factor of 6. We have already explained a factor of 3. (A factor of 3 is more than one-half of a factor of 6. Why?) The remaining error (a factor of 2) must come largely because we assumed that the earth and moon have the same density. Allowing the density to vary, we recover the original scaling relation (2.11). Then,

$$\frac{g_{\text{earth}}}{g_{\text{moon}}} \sim \frac{\rho_{\text{earth}}}{\rho_{\text{moon}}} \frac{R_{\text{earth}}}{R_{\text{moon}}} \quad (2.14)$$

Typically, $\rho_{\text{crust}} \sim \rho_{\text{moon}} \sim 3 \text{ g cm}^{-3}$, whereas $\rho_{\text{earth}} \sim 5 \text{ g cm}^{-3}$ (here, ρ_{crust} is the density of the earth's crust).

Although we did not show you how to deduce the density of moon rock from well-known numbers, we repay the debt by presenting a speculation that results from comparing the average densities of the earth and the moon. Moon rock is lighter than earth rock; rocks in the earth's crust are also lighter than the average earth rock (here "rock" is used to include all materials that make up the earth, including the core, which is nickel and iron); when the earth was young, the heavier, and therefore denser, elements sank to the center of the earth. In fact, moon rock has density close to that of the earth's crust—perhaps because the moon was carved out of the earth's crust. Even if this hypothesis is not true, it is plausible, and it suggests experiments that might disprove it. Its genesis shows an advantage of the

scaling method over the brute-force method: The scaling method forces us to compare the properties of one system with the properties of another. In making that comparison, we may find an interesting hypothesis.

Whatever the early history of the moon, the density ratio contributes a factor of $5/3 \sim 1.7$ to the ratio (2.14), and we get $g_{\text{earth}}/g_{\text{moon}} \sim 3 \times 1.7 \sim 5$. We have explained most of the factor of 6—as much of it as we can expect, given the crude method that we used to estimate the moon’s radius, and the one-digit accuracy that we used for the densities.

The brute-force method—looking up all the relevant numbers in a table—defeats the purpose of order-of-magnitude analysis. Instead of approximating, you use precise values and get a precise answer. You combine numerous physical effects into one equation, so you cannot easily discern which effects are important. The scaling method, where we first approximate the earth and moon as having the same density and radius, and then correct the most inaccurate assumptions, teaches us more. It explains why $g_{\text{moon}} \sim g_{\text{earth}}$: because the earth and moon are made of similar material and are roughly the same size. It explains why $g_{\text{moon}}/g_{\text{earth}} \simeq 1/6$: because moon rock is lighter than earth rock, and because the moon is smaller than the earth. We found a series of successive approximations:

$$\begin{aligned} g_{\text{moon}} &\sim g_{\text{earth}}, \\ g_{\text{moon}} &\sim \frac{R_{\text{moon}}}{R_{\text{earth}}} g_{\text{earth}}, \\ g_{\text{moon}} &\sim \frac{\rho_{\text{moon}}}{\rho_{\text{earth}}} \frac{R_{\text{moon}}}{R_{\text{earth}}} g_{\text{earth}}. \end{aligned} \tag{2.15}$$

Each approximation introduces only one physical effect, and is therefore easy to understand. Another benefit of the scaling method is that it can suggest new theories or hypotheses. When we considered the density of moon rock and earth rock, we were led to speculate on the moon’s origin from the earth’s crust. Order-of-magnitude reasoning highlights the important factors, so that our limited brains can digest them, draw conclusions from them, and possibly extend them.

2.2.2 Collisions

Imagine that you work for a government safety agency testing how safe various cars are in crashes. Your budget is slim, so you first crash small toy cars, not real cars, into brick walls. (Actually, you might crash cars in computer simulation only, but, as the order-of-magnitude analysis of computer programs is not the topic of this example, we ignore this possibility.) At what speed does such a crash produce mangled and twisted metal? Metal toy cars are still available (although hard to find), and we assume that you are using them.

For our initial guess, let’s estimate that the speed should be roughly 50 mph or 80 kph—roughly the same speed that would badly mangle a real car (mangle the panels and the engine compartment, not just the fenders). Why does a crash make metal bend? Because the kinetic energy from the crash distorts the metallic bonds. We determine the necessary crash speed using a scaling argument.

Figure 2.7 shows a car about to hit a brick wall. In an order-of-magnitude world, all cars, toy or real, have the same proportions, so the only variable that distinguishes them is their length, L . (Because we are assuming that all cars have the same proportions, we

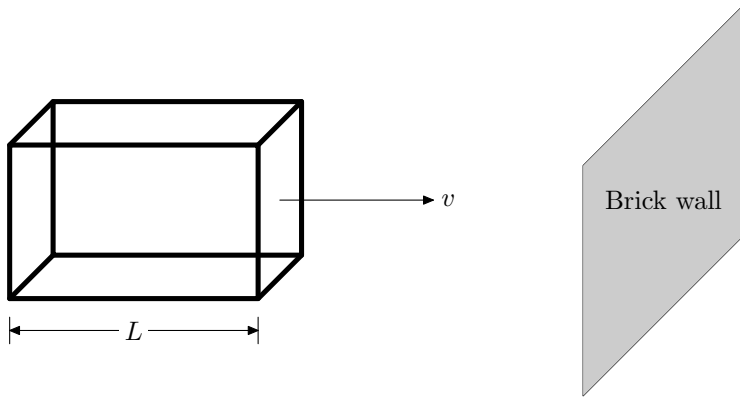


Figure 2.7. Order-of-magnitude car about to hit a brick wall. It hits with speed v , which provides kinetic energy $\sim Mv^2$, where M is the mass of the car. The energy required to distort a fixed fraction of the bonds is proportional to the number of bonds. If toy and real cars are made of the same metal, then the number of atoms, and the total bond-distortion energy, will be proportional to M , the mass of the car. The available kinetic energy also is proportional to M , so the necessary crash velocity is the same at all masses, and, therefore, at all sizes.

could use the width or height instead of the length.) The kinetic energy available is

$$E_{\text{kinetic}} \sim Mv^2. \quad (2.16)$$

The energy required to distort the bonds is

$$E_{\text{required}} \sim \underbrace{\frac{M}{m_{\text{atom}}}}_{\text{no. of atoms}} \times \epsilon_c \times f, \quad (2.17)$$

where ϵ_c is the binding, or cohesive, energy per atom; and f is a fractional fudge factor thrown in because the crash does not need to break every bond. We discuss and estimate cohesive energies in Section 4.2.2; for now, we need to know only that the cohesive energy is an estimate of how strong the bonds in the substance are. Let's assume that, to mangle metal, the collision must break a fixed fraction of the bonds, perhaps $f \sim 0.01$. Equating the available energy (2.16) and the required energy (2.17), we find that

$$Mv^2 \sim M \times \frac{\epsilon_c}{m_{\text{atom}}} \times f. \quad (2.18)$$

We assume (reasonably) that ϵ_c , f , and m_{atom} are the same for all cars, toy or real, so once we cancel M , we have $v \propto 1$. The required speed is the same at all sizes, as we had guessed.

Now that we have a zeroth-order understanding of the problem, we can improve our analysis, which assumed that all cars have the same shape. The metal in toy cars is proportionally thicker than the metal in real cars, just as roads on maps are proportionally wider than real roads. So a toy car has a larger mass, and is therefore stronger than the simple

scaling predicts. The metal in full-size cars mangles in a 80 kph crash; the metal in toy cars may survive an 80 kph crash, and may mangle only at a significantly higher speed, such as 200 kph.

Our solution shows the benefit of optimism. We do not know the fudge factor f , or the cohesive energy ϵ_c , but if we assume that they are the same for all cars, toy or real, then we can ignore them. The moral is this: **Use symbols** for quantities that you do not know; they might cancel at the end. Our example illustrated another technique: successive approximation. We made a reasonable analysis—implicitly assuming that all cars have the same shape—then improved it. The initial analysis was simple, and the correction was almost as simple. Doing the more accurate analysis in one step would have been more difficult.

2.2.3 Jump heights

We next apply scaling methods to understand how high an animal can jump, as a function of its size. We study a jump from standing (or from rest, for animals that do not stand); a running jump depends on different physics. This jump-height problem also looks under-specified. The height depends on how much muscle an animal has, how efficient the muscles are, what the animal's shape is, and much else. So we invoke another order-of-magnitude method: When the going gets tough, **lower your standards**. We cannot easily figure out the absolute height; we estimate instead how the height depends on size, leaving the constant of proportionality to be determined by experiment. In Section 2.2.3.1, we develop a simple model of jumping; in Section 2.2.3.2, we consider physical effects that we neglected in the crude approximations.

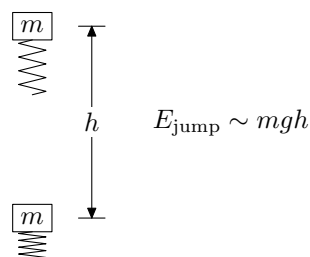


Figure 2.8. *Jumping animal.* An animal of mass m (the block) stores energy in its muscles (the compressed, massless spring). It uses the energy to jump a height h off the ground. The energy required is $E_{\text{jump}} \sim mgh$.

2.2.3.1 SIMPLE MODEL FOR JUMP HEIGHT. We want to determine only how jump height scales (varies) with body mass. Even this problem looks difficult; the height still depends on muscle efficiency, and so on. Let's see how far we get by just plowing along, and using symbols for the unknown quantities. Maybe all the unknowns cancel. We want an equation for the height h , such as $h \sim f(m)$, where m is the animal's mass. Jumping requires energy, which must be provided by muscles. [Muscles get their energy from sugar, which gets its energy from sunlight, but we are not concerned with the ultimate origins of energy here.] If we can determine the required energy, and compare it with the energy that all the muscles

in an animal can supply, then we have an equation for f . Figure 2.8 shows a cartoon version of the problem.

A jump of height h requires energy $E_{\text{jump}} \sim mgh$. So we can write

$$E_{\text{jump}} \propto mh. \quad (2.19)$$

The \propto sign means that we do not have to worry about making the units on both sides match. We exploited this freedom to get rid of the irrelevant constant g (which is the same for all animals on the earth, unless some animal has discovered antigravity). The energy that the animal can produce depends on many factors. We use symbols for each of these unknowns. First, the energy depends on how much muscle an animal has. So we approximate by assuming that a fraction, α , of an animal's mass is muscle, and that all muscle tissue can store the same energy density, \mathcal{E} (we are optimists). Then, the energy that can be stored in muscles is

$$E_{\text{stored}} \sim m\alpha\mathcal{E} \propto m. \quad (2.20)$$

Here we have derived a scaling relation, showing how energy stored varies with mass; we used the freedom provided by \propto to get rid of α and \mathcal{E} , presumed to be the same for all animals. Equating the required energy from (2.19) with the available energy from (2.20), we find that $mh \propto m$, or that $h \propto 1$; this proportionality says that h is independent of mass. This result seems surprising. Our intuition tells us that people should be able to jump higher than locusts. Table 2.1 shows measured jump heights for animals of various sizes and shapes; the data are also plotted in Figure 2.9. Surprising or not, our result is roughly correct.

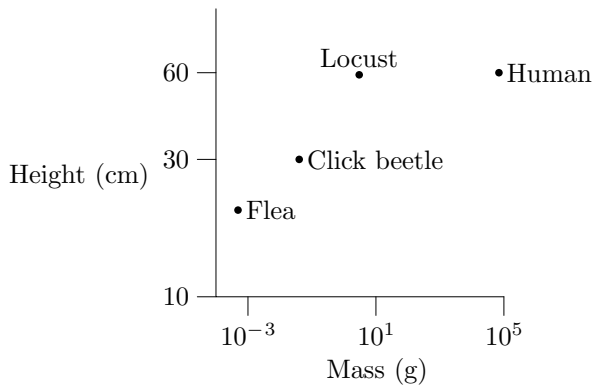


Figure 2.9. Jump height versus body mass. This graph plots the data in Table 2.1. Notice the small range of variation in height, compared to the range of variations in mass. The mass varies more than 8 orders of magnitude (a factor of 10^8), yet the jump height varies only by a factor of 3. The predicted scaling of constant h ($h \propto 1$) is surprisingly accurate. The largest error shows up at the light end; fleas and beetles do not jump as high as larger animals, due to air resistance.

<i>Animal</i>	<i>Mass (g)</i>	<i>Height (cm)</i>
Flea	$0.5 \cdot 10^{-3}$	20
Click beetle	0.04	30
Locust	3	59
Human	$7 \cdot 10^4$	60

Table 2.1. *Jump height as a function of mass. Source: Scaling: Why Animal Size is So Important [54, page 178].*

2.2.3.2 EXTENSIONS OF THE SIMPLE MODEL. Now that we have a crude understanding of the situation—that jump height is constant—we try to explain more subtle effects. For example, the scaling breaks down for tiny animals such as fleas; they do not jump as high as we expect. What could limit the jump heights for tiny animals? Smaller animals have a larger surface-to-volume ratio than do large animals, so any effect that depends on the surface area is more important for a small animal. One such effect is air resistance; the drag force F on an animal of size L is $F \propto L^2$, as we show in Section 3.3.3. The resulting deceleration is $F/m \propto L^{-1}$, so small animals (small L) get decelerated more than big animals. We would have to include the constants of proportionality to check whether the effect is sufficiently large to make a difference; for example, it could be a negligible effect for large animals, and 10 times as large for small animals, but still be negligible. If we made the estimate, we would find that the effect of air resistance is important, and can partially explain why fleas do not jump as high as humans. The constant jump height also fails for large animals such as elephants, who would break their bones when they landed if they jumped as high as humans.

You might object that treating muscle simply as an energy storage medium ignores a lot of physics. This criticism is valid, but if the basic model is correct, it's simpler to improve the model one step at a time, instead of including every effect in the early stages. As an example of successive refinement, let's consider the power requirements for jumping. How does the power required scale with animal size, and do limitations on power prevent animals from attaining their theoretical jump height?

Power is energy per time; in this case, it is energy required for the jump divided by time during which the energy is released. In (2.19) we found that $E \propto mh$; because h is constant, $E \propto m$. [Successive refinement, which we are doing here, depends on an at least rudimentary understanding of the problem. If we had not already solved the problem crudely, we would not know that $E \propto m$ or that $h \propto 1$.]

We now need to estimate the time required to release the energy, which is roughly the time during which the animal touches the ground while launching. Suppose that the animal blasts off with velocity v . The animal squats to zero height, the clock starts ticking, and the animal starts to push. At the end of the push, when the clock stops ticking, the animal is moving with speed v ; we assume that it moves with the same speed throughout its launch (the rectangle assumption). The clock, therefore, stops ticking at time $\tau \sim L/v$. The takeoff speed v is roughly the same for all animals, because $v \propto \sqrt{gh} \propto \sqrt{h}$, and h is roughly constant. So $\tau \propto L$.

How does the energy vary with L ? We make the simplest assumption—that all animals have the same density and the same cubical shape. Then, $E \propto m$, and $m \propto L^3$, so $E \propto L^3$.

From our estimates for the energy and the time, we estimate that the power required is $P \sim E/\tau \propto L^2$. Per unit volume, the power required is $\mathcal{P}_{\text{req}} \sim L^{-1}$. If there is a maximum power per unit volume, \mathcal{P}_{max} , that an animal can generate, then sufficiently tiny animals—for whom \mathcal{P}_{req} is large—might not be able to generate sufficient power. Click beetles overcome this problem by storing energy in their exoskeleton, and jumping only after they have stored sufficient energy: They increase the effective τ , and thus decrease \mathcal{P}_{req} .

The analysis of this extreme case—tiny animals—and the analysis of the power requirements show the value of making a simple analysis, and then refining it. To complete the more detailed analysis, we required results from the simple analysis. If we had tried to include all factors—such as air resistance, bone breakage, power consumption, and energy storage—from the beginning, we would have cooked up a conceptual goulash, and would have had trouble digesting the mess. The approximate model provides a structure on which we can build the more detailed analyses.

2.3 What you have learned

You now know a basic repertoire of order-of-magnitude techniques:

- *Divide and conquer*: Split a complicated problem into manageable chunks, especially when you must deal with tiny or huge numbers, or when a formula naturally factors into parts (such as $V \sim l \times w \times h$).
- *Guess*: Make a guess before solving a problem. The guess may suggest a method of attack. For example, if the guess results in a tiny or huge number, consider using divide and conquer. The guess may provide a rough estimate; then you can remember the final estimate as a correction to the guess. Furthermore, guessing—and checking and modifying your guess—improves your intuition and guesses for future problems.
- *Talk to your gut*: When you make a guess, ask your gut how it feels. Is it too high? Too low? If the guess is both, then it's probably reliable.
- *Lie skillfully*: Simplify a complicated situation by assuming what you need to know to solve it. For example, when you do not know what shape an object has, assume that it is a sphere or a cube.
- *Cross-check*: Solve a problem in more than one way, to check whether your answers correspond.
- *Use guerrilla warfare*: Dredge up related facts to help you make an estimate.
- *Punt*: If you're worried about a physical effect, do not worry about it in your first attempt at a solution. The productive strategy is to start estimating, to explore the problem, and then to handle the exceptions once you understand the domain.
- *Be an optimist*: This method is related to *punt*. If an assumption allows a solution, make it, and worry about the damage afterward.
- *Lower your standards*: If you cannot solve the entire problem as asked, solve those parts of it that you can, because the subproblem might still be interesting. Solving the subproblem also clarifies what you need to know to solve the original problem.
- *Use symbols*: Even if you do not know a certain value—for example, the energy density

stored in muscle—define a symbol for it. It may cancel later. If it does not, and the problem is still too complex, then lower your standards.

We apply these techniques, and introduce a few more, in the chapters to come. With a little knowledge and a repertoire of techniques, you can estimate many quantities.

3 DIMENSIONAL ANALYSIS

In this chapter, you learn a general method of coping with complicated equations without actually solving them: **dimensional analysis**. Because the method is fast—much faster than finding an honest solution would be—you can use it to discard unpromising approaches early, sparing you time to think of new approaches. Rather than bore you with a theoretical discussion of dimensional analysis, we illustrate the method with examples.

The idea of dimensional analysis is that units—be they meters, feet, seconds, pints, fortnights, or furlongs—are artificial. The universe cares not for our choice of units. Valid physical laws must have the same form in any system of units. Only dimensionless quantities—pure numbers—are the same in every unit system, so we write equations in a universe-friendly, dimensionless form. Often, there is only one such form. Then, without doing any work, we have solved the problem.

3.1 Newton’s law

For example, suppose that we live in Newton’s time. Lacking his insight, we do not realize that force (F), mass (m), and acceleration (a) are related by $F = ma$. We know, perhaps from experiments, only that F , m , and a are related. Our problem is to find that way. The most general relation between the three variables is

$$h(F, m, a) = 0, \quad (3.1)$$

where h is a function that we must find.

We use dimensional analysis to find h . Dimensional analysis tells us that we can rewrite the general form (3.1) using only quantities without units. With luck, there is only one such form: If we find it, we obtain a law of nature.

3.1.1 Finding dimensionless groups

To get rid of the units, we need first to know them. In cgs units, the units of force are g cm s^{-2} ; in a general mass–length–time system of units, the units are

$$[F] = [M][L][T]^{-2}. \quad (3.2)$$

We introduce the useful notation that $[M]$ stands for a unit of mass, $[L]$ for a unit of length, and $[T]$ for a unit of time. Also, $[x]$ stands for the units of the enclosed variable x ; context distinguishes the two usages of $[\cdot]$. Similarly, the units of acceleration are

$$[a] = [L][T]^{-2}, \quad (3.3)$$

and the units of mass are

$$[m] = [M]. \quad (3.4)$$

We can combine the three variables into a **dimensionless group**—a quantity with no units. An example of an illegal group is F/a , which has units of mass. Legal groups include

$$\frac{F}{ma}, \quad \left(\frac{F}{ma}\right)^2, \quad \sin\left(\frac{F}{ma}\right), \quad \text{and} \quad \frac{ma}{F}. \quad (3.5)$$

These legal groups—all legal groups—are functions of F/ma . They are also functions of ma/F or $2F/ma$. It does not matter which base form we choose, so we choose F/ma . This form does not contain an arbitrary constant (a weakness of $2F/ma$), which is superfluous in a dimensional analysis. This form also makes later manipulations most familiar. The general form of a dimensionless group is therefore

$$f\left(\frac{F}{ma}\right), \quad (3.6)$$

where f is a dimensionless function: It takes in dimensionless numbers and spits out a dimensionless number.

3.1.2 Finding the dimensionless relation

The most general relation using (3.6) is

$$f\left(\frac{F}{ma}\right) = 0. \quad (3.7)$$

This form is the dimensionless version of (3.1). To make sure that we have not gone astray, let's check that (3.7) does not exclude what we know to be the right answer: $F = ma$. If we choose $f(x) = x - 1$, then (3.7) becomes $F/ma = 1$, or $F = ma$.

We have made progress. The function h has three parameters; f has only one.¹ We can further simplify (3.7), because we need to find only the zeros of f ; we do not need to find f itself. We have sketched a dimensionless function in Figure 3.1. Its six zero-crossings determine the possible values of F/ma .

Which zero is correct? In some problems, more than one zero is possible; for example, an electron in a hydrogen atom can have an infinity of possible energies. In this problem, only one zero is possible—a fact that we could discover from experiment; dimensional analysis cannot decide this question for us. We call the unknown zero Π , in honor of the Buckingham Pi theorem (Theorem 3.1). The solution of (3.7) is $F/ma = \Pi$, or

$$F = \Pi ma. \quad (3.8)$$

This result is Newton's second law. We deduced it without knowing much physics. [This example is unfair to Newton, because it implies that he was dense for not instantly realizing that $F = ma$. However, we took for granted much knowledge that was unknown in Newton's

1. "A good table of functions of one variable may require a page; that of a function of two variables a volume; that of a function of three variables a bookcase; and that of a function of four variables a library."

—H. Jeffries [45, p. 82]

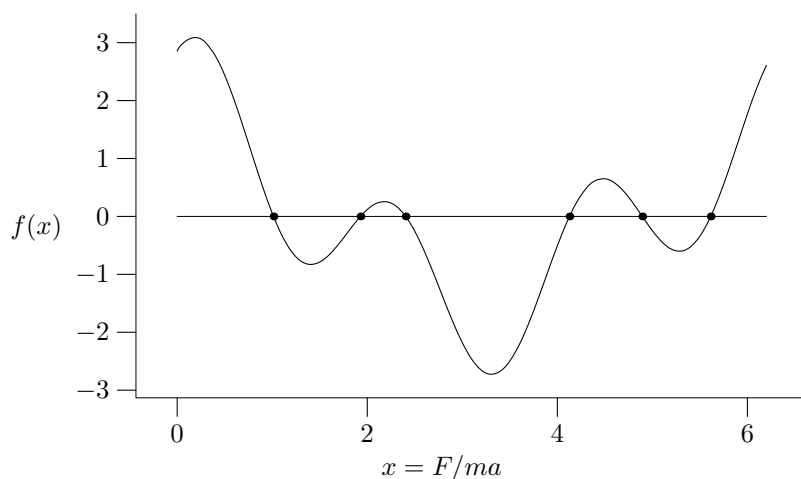


Figure 3.1. Zero-crossings of a random function f . They are located at 4.894, 5.61392, 1.01346, 1.9288, 2.4041, and 4.12762 and are marked with dots.

time. At that time, the concept of force—let alone its units—was not clearly understood, and it was the great triumph of Newton that he gave an implicit definition of force in $F = ma$.] Even with our extra knowledge, we have not matched Newton. Dimensional analysis does not tell us the value of Π . It cannot, because numbers, such as Π , have no units; they are therefore invisible to dimensional analysis, which cares about only the units of a quantity, not its magnitude. We could do an experiment to determine Π ; if we did, we would find that $\Pi = 1$. Much of this text assumes (or hopes) that Π is near 1. We often use dimensional analysis to solve problems for which we cannot easily determine Π ; at the end, we pretend that $\Pi = 1$. The closer Π is to unity, the more accurate is the resulting estimate.

3.1.3 Summarizing the argument

Let's extract the general style of argument that we used in this example; we use a similar pattern in many order-of-magnitude analyses. What did we have to know before we could begin? We had to know which variables were important: F , m , and a . In this example, the variables were given by fiat, as part of the thought experiment. In more realistic examples, we have to think carefully to choose the variables: Choosing them is the hardest part in dimensional analysis. Aristotle thought that force, mass, and velocity were the relevant variables; it took almost 2000 years before scientists realized that acceleration, rather than velocity, was the relevant variable.

If we leave out a necessary variable, we invite trouble. For example, suppose that we use only F and a . No combination of F and a is dimensionless; this failure is a clue that we've left out a necessary variable—in this case, the mass. We can also make more subtle mistakes. Suppose that we correctly realize that mass should be a variable, but we incorrectly think that the relevant mass is the mass of the earth m_{\oplus} , because of a vague hunch about gravity. We form the dimensionless group $F/m_{\oplus}a$, and find that $F = \Pi m_{\oplus}a$. Although dimensionally correct by construction, this equation is empirically bogus. Newton's law is accurate in far regions of the solar system, where the earth's gravity, and m_{\oplus} , are irrelevant.

You must think physically, appeal to experiment, and make lucky guesses if you are to include all relevant variables.

We do not want to include every possible variable, however, because irrelevant variables multiply the possibilities for dimensionless relations. For example, suppose that, to hedge our bets, we include both m and m_{\oplus} . Then, there are two dimensionless groups, F/ma and $F/m_{\oplus}a$. The most general relation using these groups is

$$f\left(\frac{F}{ma}, \frac{F}{m_{\oplus}a}\right) = 0, \quad (3.9)$$

where f is a dimensionless function of two variables. We now have to use physics knowledge to restrict the form of f . Here, we add the knowledge that the mass of the earth is irrelevant, because Newton's law works in the far solar system; then, we recover the simpler relation $f(F/ma) = 0$ that we derived before, but, this time, we had to work harder. Section 3.3.2 contains another example of the excess-variable problem; there, we use knowledge of fluid mechanics to simplify the resulting relations. We know of no recipe for choosing the right set of variables, except to practice.

[We sloppily used the same letter f to denote two different functions: In (3.7), it is a function of one variable, yet in (3.9), it is a function of two variables. The number of parameters disambiguates the two uses. For maximum clarity, we should call the two functions by different names, such as f_1 and f_2 . In dimensional-analysis problems, however, we often start with, say, a three-variable function; then use physics knowledge to reduce it to a two-variable function; then use more knowledge to reduce it to a one-variable function. The sloppy notation is therefore convenient. We avoid redundant subscripts, and we do not need to invent arbitrary names such as f , g , and h .]

Once we chose the variables, we found the only dimensionless group (apart from transformations) F/ma . The general relation, $f(F/ma) = 0$ simplified to $F = \Pi ma$. The unknown constant Π we can determine by experiment, or from other knowledge. Figure 3.2 illustrates the steps of the argument.

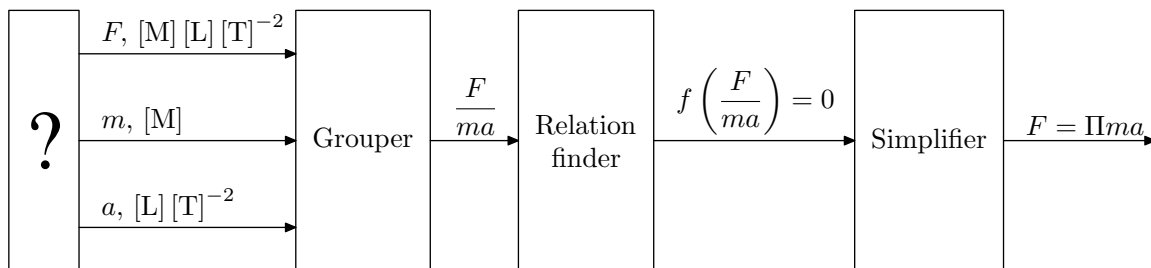


Figure 3.2. The Pi machine. After lucky guessing, we decide what variables to feed the grouper. We feed them in, along with their units. The grouper hunts for combinations that have no units (for dimensionless groups), and spits them out (here, there is only one group). That group enters the relation finder, which produces the most general dimensionless relation from its inputs. The simplifier simplifies this equation, presenting the answer $F = \Pi ma$.

3.1.4 Generalizing the argument: the Buckingham Pi theorem

The art of dimensional analysis lies in choosing the right set of relevant variables; a computer would find this stage difficult, if not impossible. The part that we could program—finding the dimensionless groups—we normally do by guessing. In this example, we started with three relevant variables, composed of two basic units, $[M]$, and $[L][T]^{-2}$. It looks like the variables contain units— $[M]$, $[L]$, and $[T]$ —but the single combination $[L][T]^{-2}$ suffices to replace $[L]$ and $[T]$. As an analogy, imagine three functions— a , b , and c —with the forms

$$\begin{aligned} a(x, y, z) &= xyz^{-2}, \\ b(x, y, z) &= x, \\ c(x, y, z) &= yz^{-2}. \end{aligned} \tag{3.10}$$

Here, a , b , and c play the roles of $[F]$, $[m]$, and $[a]$, respectively; and x , y , and z play the roles of $[M]$, $[L]$, and $[T]$, respectively. The variables y and z occur in only the combination yz^{-2} . So there are only two independent variables: x and yz^{-2} . In the $F = ma$ example, there are only two independent units: $[M]$ and $[L][T]^{-2}$. From $n = 3$ variables, composed of $r = 2$ units, we found $n - r = 1$ dimensionless group. This pattern is general, and is stated in Theorem 3.1.

Theorem 3.1 (Buckingham Pi theorem) A system described by n variables, built from r independent dimensions, is also described by $n - r$ independent dimensionless groups.

We also call the groups Π variables. Now we apply this method to examples with more than one dimensionless group, where the argument is more complex.

3.2 Pendula

The next example is familiar to many of you from freshman-physics problem sets. A pendulum bob hangs at the end of a massless rope, and, from a resting state, the bob starts oscillating (Figure 3.3). What is the oscillation period?

3.2.1 Finding dimensionless groups

Following the pattern of Section 3.1, we first choose the relevant variables. We make sure to include the variable for which we're trying to solve, the period τ ; a common mistake is to forget the goal variable, and thus to end up one variable short. The period may depend on the initial angle, so we include θ_0 . To decide which other variables to include, we imagine life as a pendulum bob. Why do we move at all? We move because of gravity, so g belongs on our list. A more massive bob feels a stronger gravitational force, so we include m in our list.²

From τ , g , and m , θ_0 —four variables with three units—we can form one dimensionless group, θ_0 . We cannot form one that includes τ (try to do so); so we really have only three

2. You may object that we have personified the pendulum bob, that we have endowed it with the capacity for feeling, at least, for feeling forces. Even worse, we have become a pendulum bob. We plead guilty, with an explanation. We personify the bob, because picturing an active bob enhances our intuition for how it behaves; to make a vivid picture of an active bob, we even pretend to be a bob. Similarly, looking for a lost marble, we often ask ourselves, “If I were a marble and somebody dropped me, where would I roll?” This style may not enhance your intuition; try it for month as an experiment.

variables, and zero dimensionless groups. We must have left out a variable. A parameter of the system that we have not mentioned is the length of the rope. Perhaps that variable will rescue us, so we include it in our list. Table 3.1 contains the variables that we have collected. (From previous experience with this problem, you know that m does not matter, and that l does; we will use those facts as a sanity check on our solution.)

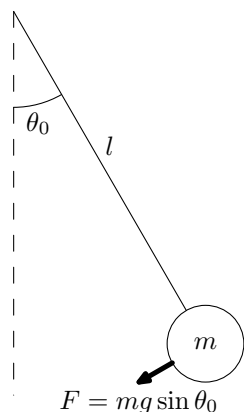


Figure 3.3. A pendulum bob of mass m hangs from a massless rope of length l . The bob is released from rest at an angle θ_0 .

<i>Var.</i>	<i>Units</i>	<i>Description</i>
θ_0	—	angle of release
m	[M]	mass of bob
τ	[T]	period
g	[L][T] ⁻²	acceleration due to gravity
l	[L]	length of rope

Table 3.1. Variables that may determine a pendulum’s oscillation period.

We now form dimensionless groups. Finding these groups requires that we have handy the units of the variables, which is why we listed the variables in a table. These $n = 5$ variables contain $r = 3$ basic dimensions: length, mass, and time. The Buckingham Pi theorem says that we can form $n - r = 2$ independent dimensionless groups from this set. Guesswork is the easiest way to find them. One original variable, θ_0 , is already dimensionless (why are angles dimensionless?), so we start with $\Pi_1 \equiv \theta_0$. Only one original variable, m , contains a unit of mass. Therefore, no dimensionless group can contain m . [To be dimensionless, the group has to contain another variable that cancels the unit of mass contributed by m ; however, m is the only mass in this problem, so m must cancel itself from any group that it tries to join.³] We form the second independent dimensionless variable using g , l , and τ .

3. “Any club that would admit me as a member, I wouldn’t want to join.”

—Groucho Marx

[If we did not require that the $n - r$ Π variables be independent, then we could choose, for example, the useless pair θ_0 and θ_0^2 .] The units of g show us how to form the group: Cancel the length in g with l^{-1} , and cancel the time with τ^2 , to make $\Pi_2 \equiv g\tau^2/l$. The two groups are then

$$\begin{aligned}\Pi_1 &= \theta_0, \\ \Pi_2 &= \frac{g\tau^2}{l}.\end{aligned}\tag{3.11}$$

3.2.2 Finding dimensionless relations

The most general relation using the groups in (3.11) is

$$f(\Pi_1, \Pi_2) = 0,\tag{3.12}$$

where f is a dimensionless function. What is f ? Figure 3.4 outlines the argument that we use find it.

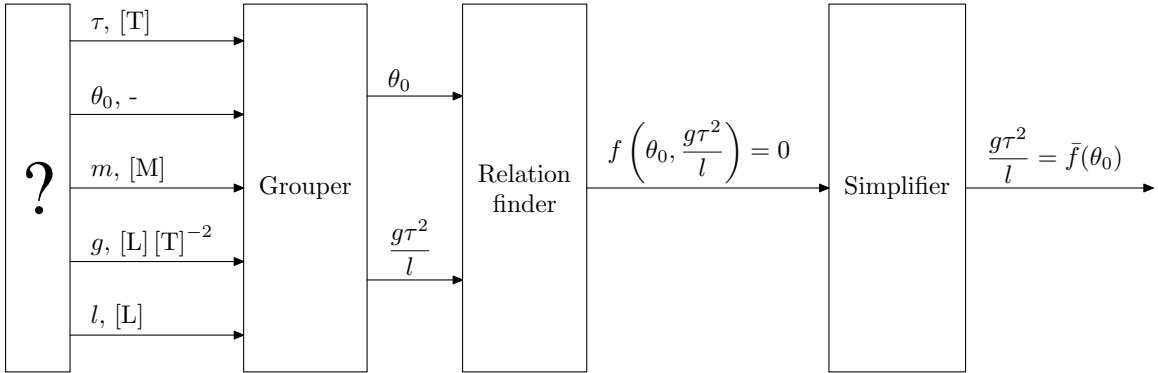


Figure 3.4. Dimensional analysis for the pendulum period. After thinking, we choose five relevant variables. They enter the grouper. It counts the inputs (five) and the number of independent units in the inputs (three), and spits out two dimensionless groups. The relation finder produces the most general relation using the two groups. The simplifier transforms the relation into a form more convenient for solving for τ .

In Section 3.1.2, we rewrote $f(\Pi_1) = 0$ as $\Pi_1 = \Pi$, where $\Pi_1 = F/ma$ was the dimensionless group, and Π was an unknown constant—the zero of f that is experimentally correct. With the more complicated relation (3.12), we can make the same transformation. The two-parameter function f is equivalent to a family of one-parameter functions, f_{Π_1} , indexed by Π_1 (Figure 3.5). The general form (3.12) is then

$$f_{\Pi_1}(\Pi_2) = 0.\tag{3.13}$$

In Section 3.1.2, where f was a one-parameter function, we transformed $f(\Pi_1) = 0$ to $\Pi_1 = \Pi$. Here, we transform $f_{\Pi_1}(\Pi_2) = 0$ to

$$\Pi_2 = \bar{f}(\Pi_1),\tag{3.14}$$

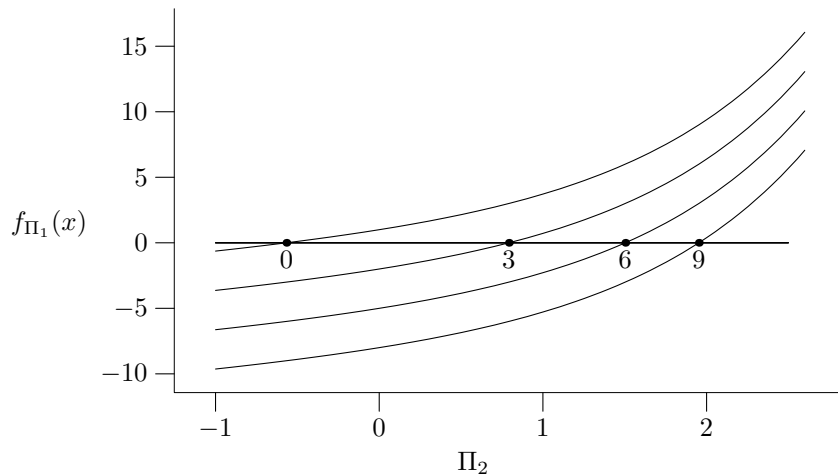


Figure 3.5. The two-parameter function $f(\Pi_1, \Pi_2) = e^{\Pi_2} + \Pi_2 - \Pi_1$, plotted for various Π_1 . The result is a family of functions, $f_{\Pi_1}(\Pi_2)$. Their zero-crossings are marked with dots, and labeled with the values of Π_1 .

where $\bar{f}(\Pi_1)$ is the zero of f_{Π_1} (Figure 3.6). (This transformation is valid when f_{Π_1} has only one physically valid zero, which we assume for simplicity.) We have extracted Π_2 from its hiding place in f . Alternatively, we could have written $f(\Pi_1, \Pi_2) = f_{\Pi_2}(\Pi_1)$, and have defined $\bar{f}(\Pi_2)$ as the zero of f_{Π_2} . Then, we could have extracted Π_1 :

$$\Pi_1 = \bar{f}(\Pi_2). \quad (3.15)$$

We choose which form to use by the location of the goal variable. If it belongs to Π_2 , we use (3.14). If it belongs to Π_1 , we use (3.15). If it belongs to both, then we have to think. If it belongs to neither, then we have too few starting variables and dimensionless groups.

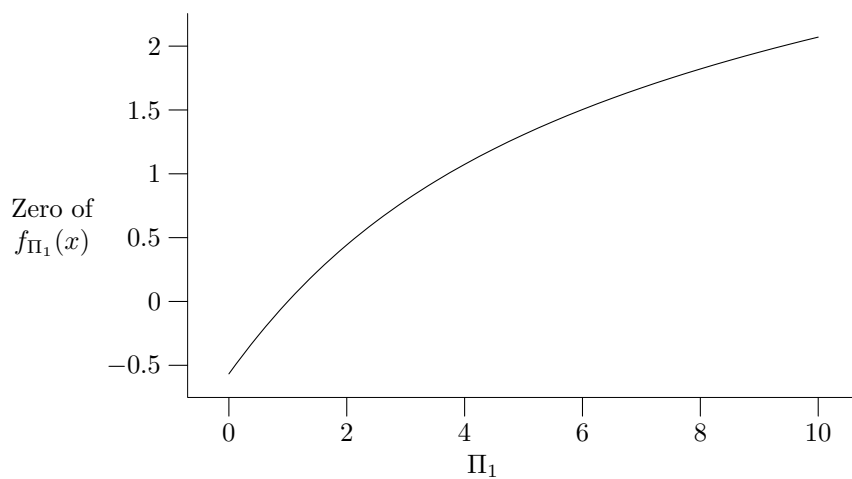


Figure 3.6. Zero of $f_{\Pi_1}(\Pi_2)$, as a function of Π_1 . The function $f(\Pi_1, \Pi_2)$ is from Figure 3.5.

We wish to solve for the period, which is contained in $\Pi_2 = g\tau^2/l$, so we use the form (3.14). In terms of the original variables, (3.14) is $g\tau^2/l = \bar{f}(\theta_0)$, or

$$\tau = \bar{f}(\theta_0) \sqrt{\frac{l}{g}}. \quad (3.16)$$

So τ depends on l but not on m (the promised sanity check). We do not know the function $\bar{f}(\theta_0)$. When we derived Newton's law in Section 3.1.2, we could not, using dimensional analysis, determine the constant (or zero-parameter function) Π . Here, we cannot determine the one-parameter function \bar{f} . We can determine it by experiment: We release a pendulum at various θ_0 , and measure $\tau(\theta_0)$. Then, \bar{f} is

$$\bar{f}(\theta_0) = \tau(\theta_0) \sqrt{\frac{g}{l}}. \quad (3.17)$$

We do not have to repeat the experiments for different l (say, for another pendulum) or g (say, on another planet), because \bar{f} is universal. All pendulums—long or short, on the earth or on Mars—obey the same \bar{f} .

In the small-amplitude limit ($\theta_0 \rightarrow 0$), we can simplify (3.16):

$$\tau = \bar{f}(0) \sqrt{\frac{l}{g}}, \quad (3.18)$$

where $\bar{f}(0)$ is a constant. In speaking of $\bar{f}(0)$, we are tacitly assuming that $\lim_{x \rightarrow 0} \bar{f}(x)$ exists. Dimensional analysis makes no promises that it does; we have to appeal to physics. Here is a scaling argument that makes the existence of the limit plausible. We estimate the oscillation time by estimating the acceleration and the oscillation distance. The pendulum feels a force $F \sim mg \sin \theta_0$ (see Figure 3.3), which makes it accelerate at $a \sim g \sin \theta_0$. For small θ_0 , the acceleration is $a \sim g\theta_0$, because $\sin \theta_0 \sim \theta_0$ for small θ_0 . In time τ , it travels a distance $a\tau^2 \sim g\theta_0\tau^2$. It needs to travel a distance $d \sim l\theta_0$ (neglecting constants) to complete a cycle, so

$$g\theta_0\tau^2 \sim l\theta_0. \quad (3.19)$$

The amplitude θ_0 cancels, which is the physical equivalent of the mathematical statement that $\bar{f}(0)$ exists. We even get an estimate for τ :

$$\tau \sim \sqrt{\frac{l}{g}}, \quad (3.20)$$

which is a physical derivation of (3.18). Alternatively, we can estimate a typical velocity for the bob, and from the velocity, estimate the period. The maximum potential energy is $\text{PE} \sim mgh$, where the change in height is $h = l(1 - \cos \theta_0) \sim l\theta_0^2$. The maximum kinetic energy is equal to the maximum potential energy. So the maximum velocity is given by $mv^2 \sim mgl\theta_0^2$. (This relation equates two energies, potential and kinetic. This method is an example of balancing; we discuss balancing more fully in Section 4.1.2.) The maximum

velocity is then $v \sim \theta_0 \sqrt{gl}$; we use this value as an estimate for a typical velocity. The time to complete a cycle is $\tau \sim d/v \sim \sqrt{l/g}$, as we found using the acceleration method.

We call $\bar{f}(0)$ the constant Π , and (3.18) becomes

$$\tau = \Pi \sqrt{\frac{l}{g}}. \quad (3.21)$$

We cannot evaluate Π with dimensional analysis. It turns out to be 2π , but we would either have to solve an equation to determine it—and solving equations is beyond the scope of the text—or have to determine it experimentally, by timing one pendulum swing. We release the pendulum from a small angle—say, $\theta_0 = 0.1$ rad—and measure the period, τ . Knowing τ , l , and g , we use (3.21) to determine Π .

3.3 Drag in fluids

Every freshman-physics course solves the pendulum differential equation to find (3.21). The freshman course usually continues beyond where dimensional analysis can compete: The solution to the differential equation contains the value of Π . The pendulum problem does not show the benefit of dimensional analysis. So we try our hand at fluid mechanics—a subject notorious for its mathematical and physical complexity; Chandrasekhar's books [6, 7] show how complex the mathematics can become. As usual, we study examples. The examples are a marble falling through corn syrup (Section 3.3.2), and a person falling from the sky after being thrown from an airplane (Section 3.3.3). We calculate the **terminal velocity**: The velocity that the object reaches after falling for sufficiently long. The examples illustrate two extremes of fluid flow: oozing and turbulent.

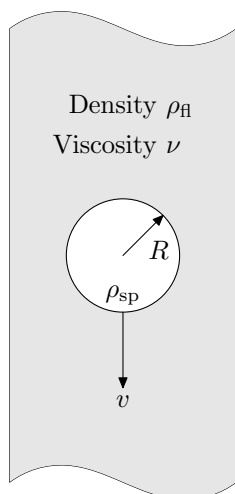


Figure 3.7. Sphere falling in a liquid. Its terminal velocity is v .

We begin by solving the partial-differential equations of fluid mechanics (for the in-

compressible flow of a Newtonian fluid):

$$\begin{aligned}\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} &= -\frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{v}, \\ \nabla \cdot \mathbf{v} &= 0.\end{aligned}\tag{3.22}$$

Here, \mathbf{v} is the fluid velocity field, ρ is the fluid density, ν is the kinematic viscosity, and p is the pressure. We have to solve six coupled, nonlinear partial-differential equations, and to find the velocity of the marble as $t \rightarrow \infty$. A miserable task.

Do not worry—we will not solve these equations. We put them there to scare you into using dimensional analysis, to show you that it is easier than solving equations.

So let's use dimensional analysis to find the terminal velocity, v . We could start with our usual method: Choose a group of variables that includes v , form dimensionless groups, and then solve for the velocity. This direct approach would give us too many dimensionless groups; we would have to add physics knowledge at the end, to simplify the resulting dimensionless relations. Instead, we use physics knowledge at the beginning, to simplify the derivation early on. You save effort if you simplify early. If you simplify later, you carry around baggage that you eventually discard. If you're going on holiday to the Caribbean, why pack snow shoes?

The adjective *terminal* in the phrase terminal velocity hints at the physics that determines the velocity. It indicates that the velocity is constant, which can happen only if there is no net force on the marble. What forces act on the marble? There is gravity, which speeds it up. There is drag, which slows it down. To find the terminal velocity, then, we find the drag and gravitational forces, and equate them. We have split the terminal-velocity problem into two simpler problems (divide and conquer). The gravitational force is mg (or close to that, as we will see), where m is the mass of the marble. What is the drag force?

3.3.1 Dimensional analysis for the drag force

We use dimensional analysis to find the drag force, F_d . The first step is to choose the relevant variables. We begin by including the variable for which we're solving, F_d . What characteristics of the sphere are relevant to the drag force? The drag force has no idea what is inside the sphere. Picture the fluid as a huge computer that implements the laws of fluid dynamics. How can it know how dense the sphere or person is? It cannot see inside the object. To the drag force, the parameters v and R are the only relevant attributes of a sphere moving through the atmosphere. What lies underneath the person's surface does not affect the fluid flow.⁴ So the computer can determine the flow (if it has tremendous processing power) knowing only v and R . Therefore, ρ_{sp} is irrelevant. What characteristics of the fluid are relevant? The fluid supercomputer needs to know the density of the fluid, to determine how fast the pieces of fluid move because of the object. So we include ρ_{fl} .

What about viscosity? Viscosity is a measure of the tendency of a fluid to smear velocity differences in the flow. You can observe an analogue of viscosity in traffic flow on a multilane highway. If one lane moves much faster than another, drivers switch from the slower to the faster lane, eventually slowing down the faster lane. Local decisions of the drivers reduce the velocity gradient. Similarly, in a fluid, molecular motion transports speed

4. Viscosity is only skin deep.

(really, momentum) from fast- to slow-flowing regions. This transport reduces the velocity difference between the regions. Thicker, oozier (more viscous) fluids probably cause more drag than thin fluids do. So we include viscosity in our list.

Fluid mechanics⁵ have defined two viscosities, dynamic viscosity μ , and kinematic viscosity ν . They are related by $\mu = \rho_{\text{fl}} \nu$. *Life in Moving Fluids* [59, pp. 23–25] discusses the two types of viscosity in detail. In Section 5.5, we estimate the viscosity of gases by examining the molecular origin of viscosity. For the analysis of drag force, we need to know only that viscous forces are proportional to viscosity. Which viscosity should we use? Dynamic viscosity hides ρ_{fl} inside the product $\nu \rho_{\text{fl}}$; equations that contain ρ_{sp} and μ then look less dimensionless than they are, because ρ_{sp} has no partner. So we use the kinematic viscosity, ν .

<i>Var.</i>	<i>Units</i>	<i>Description</i>
F_{d}	$[\text{M}][\text{L}][\text{T}]^{-2}$	drag force
R	$[\text{L}]$	marble radius
ν	$[\text{L}]^2[\text{T}]^{-1}$	kinematic viscosity
ρ_{fl}	$[\text{M}][\text{L}]^{-3}$	fluid density
v	$[\text{L}][\text{T}]^{-1}$	terminal velocity

Table 3.2. Variables that determine the drag force.

Table 3.2 lists the variables that we have collected. The drag force does not depend on ρ_{sp} or g ; it is simpler to estimate than is the terminal velocity (which does depend on g and ρ_{sp}). The five variables in the list are composed of three basic dimensions; we look for two dimensionless groups. We find one group by the method of divide and conquer. The list already includes a velocity, v . If we can concoct another quantity, V , also with units of velocity, then we can form the dimensionless group v/V . The viscosity ν almost works. It has an extra length, which R can eliminate: $V \equiv \nu/R$. The dimensionless group is

$$\Pi_1 \equiv \frac{v}{V} = \frac{vR}{\nu}, \quad (3.23)$$

which is the **Reynolds’ number**, Re . It is a dimensionless measure of the flow speed. [In Section 3.3.3, we give a more detailed physical interpretation of Re .] The velocity alone cannot distinguish fast from slow flow, because v is not dimensionless (see the discussion of large cargos in Section 2.1.1). If you hear that a quantity is small, or fast, or large, or whatever, your first reaction should be, “compared to what?” To distinguish fast from slow flow, we have to use a dimensionless quantity related to v —the Reynolds’ number. It compares v to V . Low values of Re indicate slow, viscous flow (cold honey oozing out of a jar); high values indicate turbulent flow (a jet flying at 600 mph). We once again refer you to

5. Sadly, we could not use the more mellifluous term *fluid mechanics* to signify a host of physicists agonizing over the equations of fluid mechanics; it would not distinguish the toilers from their toil.

the excellent *Life in Moving Fluids* [59], this time for more details on dimensionless ratios in fluid mechanics. We introduce another dimensionless ratio—the Froude number—when we discuss speedboating in Section 6.2.1.2.

What is the second dimensionless group? The drag force is absent from Π_1 , so it has to be part of Π_2 (else how can we solve for F_d ?). Instead of dreaming up the dimensionless group in one lucky guess, we construct two quantities with the same units, and divide them to get a dimensionless quantity. Notice that $F_d R$ is an energy. We construct another energy (call it E), then form the dimensionless ratio $F_d R/E$. Energy contains one power of mass; the only variable other than F_d that contains mass is ρ_{fl} . So E contains one power of ρ_{fl} . How can we make an energy from ρ_{fl} ? We start with $\rho_{\text{fl}} R^3$, which is a mass. Kinetic energy is $mv^2/2$, so it has units of $[M][V]^2$. [A formula is a reliable and simple method for determining the units of a quantity. Even if the formula is approximate— $mv^2/2$ is valid at speeds only much less than the speed of light—it still tells you the correct units.] The second energy is therefore

$$E \equiv \underbrace{\rho_{\text{fl}} R^3}_{\text{mass}} v^2. \quad (3.24)$$

The group Π_2 is the ratio of the two energies:

$$\Pi_2 \equiv \frac{F_d R}{E} = \frac{F_d}{\rho_{\text{fl}} R^2 v^2}. \quad (3.25)$$

The most general relation is $f(\Pi_1, \Pi_2) = 0$. We want to solve for F_d , which is contained in Π_2 , so we use the (3.14): $\Pi_2 = \bar{f}(\Pi_1)$. In terms of the original variables, the drag force is

$$F_d = \rho_{\text{fl}} R^2 v^2 \bar{f}\left(\frac{vR}{\nu}\right). \quad (3.26)$$

Now we are stuck; we can go no farther using only dimensional analysis. To learn the form of \bar{f} , we specialize our problem to extreme conditions: to turbulent, high-speed flow ($Re \gg 1$), or to viscous, low-speed flow ($Re \ll 1$).

3.3.2 Viscous limit

We first analyze the low-speed limit: a marble falling in corn syrup. You may wonder how often marbles fall in corn syrup, and why we bother with this example. The short answer to the first question is, “not often.” However, the same physics that determines the behavior of marbles in syrup also determines, for example, the behavior of fog droplets in air, of bacteria swimming in water, or of oil drops in the Millikan oil-drop experiment. We study the marble problem because it illustrates the physical principles, and because we can check our estimates with a home experiment.

In slow, viscous flows, the drag force comes from viscous forces, which are proportional to ν . Therefore, $F_d \propto \nu$. The viscosity appears exactly once: in the argument of \bar{f} , and in the denominator. So that ν flips into the numerator, we choose $\bar{f}(x) \sim 1/x$. [We used \sim to avoid writing the Π constant repeatedly. The \sim symbol means that the two sides have the same units (none in this case), but that they may differ by a dimensionless constant.] With this $f(x)$, the drag force (3.26) becomes

$$F_d = \rho_{\text{fl}} R^2 v^2 \Pi \frac{\nu}{vR} = \Pi \nu \rho_{\text{fl}} v R, \quad (3.27)$$

where Π is a dimensionless constant.

Sadly, we cannot compute the magic Π factor using dimensional analysis. We must call a fluid mechanic to do the messy work of calculating it; but we can hope that her burden is light, because we have worked out the solution except for one number. Stokes first derived the value of Π , and found that $\Pi = 6\pi$ (for a sphere). With this value, (3.27) becomes

$$F_d = 6\pi\rho_{\text{fl}}\nu v R. \quad (3.28)$$

Let's pause to sanity check this result. Larger or faster marbles should feel more drag; and a denser or more viscous fluid should produce more drag. The drag force (3.28) passes these tests.

We can make a more detailed check by explaining the result physically. The argument is similar in style to the argument that we made in Section 3.2.2 to justify $\tau \sim \sqrt{l/g}$ for a pendulum. In the viscous limit, drag is caused by viscosity only: The drag force is the viscous force. To estimate the viscous force, we have to introduce more knowledge about viscosity. We also work in the reference frame of the sphere. Viscosity is a consequence of the **no-slip boundary condition**: At the surface of the sphere, the fluid is stationary (relative to the sphere). Far away from the sphere, the fluid does not know about the sphere: The fluid moves as if the sphere were not there, with velocity v (relative to the sphere). Viscosity reduces this velocity difference by transporting momentum from the sphere to the distant fluid: It slows the sphere. The momentum transport produces a stress $\vec{\sigma}$ (roughly, a force per area) that is proportional to dynamic viscosity and to velocity gradient:

$$\vec{\sigma} \sim \rho_{\text{fl}}\nu \times \nabla \mathbf{u}, \quad (3.29)$$

where \mathbf{u} is the fluid velocity as a function of distance from the sphere's center. The velocity gradient $\nabla \mathbf{u}$ is a complicated object: a tensor. We can pretend that it is an ordinary number (a scalar), and approximate its magnitude using two data points: $u = 0$ near the sphere, and $u = v$ far away from the sphere. How far is far away? Equivalently, we can ask: Far relative to what length? The natural length to use as a standard of comparison is R , the radius of the sphere. So let's say that $u(R) = 0$ and $u(2R) = v$. We could also use $u(10R) = v$ or $u(8R) = v$; the constant that exactly characterizes "far away" is not important, because we are going to ignore it anyway. Because \mathbf{u} changes by v in a length R , its gradient is approximately v/R . This result is a consequence of a general order-of-magnitude rule of thumb:

$$\nabla f \sim \frac{\text{typical value of } f}{\text{length over which } f \text{ changes significantly}}. \quad (3.30)$$

This rule of thumb is the method of **order-of-magnitude differentiation**. [One way to remember it is to write

$$\frac{df}{dx} \sim \frac{f}{x}, \quad (3.31)$$

once we cancel the ds .] The viscous stress is therefore $\sigma \sim \rho_{\text{fl}}\nu v/R$. The viscous force is the stress times the surface area over which it acts:

$$F_{\text{viscous}} \sim \underbrace{\rho_{\text{fl}}\nu v/R}_{\sigma} \times \underbrace{R^2}_{\text{area}} \sim \rho_{\text{fl}}\nu v R. \quad (3.32)$$

This result agrees with the estimate result (3.27), which we got from dimensional analysis. We neglected constants throughout this physical argument, which is therefore incapable of justifying the factor of 6π in (3.28).

This magic factor of 6π in (3.28) comes from doing **honest** physics, here, from solving the Navier–Stokes equations (3.22). In this book, we wish to teach you how to avoid suffering, so we do not solve such equations. Occasionally, we quote the factor that honest physics produces, to show you how accurate (or sloppy) the order-of-magnitude approximations are. The honest-physics factor is usually of order unity. Such a number suits our neural hardware: It is easy to remember and to use. If we know the order-of-magnitude derivation and remember this one number, we can reconstruct the exact result without solving difficult equations.

We now return to the original problem: finding the terminal velocity, which we estimate by balancing drag against gravity. The gravitational force is $F_g = mg$, where m is the mass of the marble. Instead of using m , we use ρ_{sp} —the density of the marble. We can recover the mass from ρ_{sp} and R if we need it. We prefer to use density, because the drag force contains a density, ρ_{fl} . We therefore write the gravitational force as

$$F_g \sim \rho_{\text{sp}} g R^3, \quad (3.33)$$

where we have neglected the $4\pi/3$ factor. We equate the drag force from (3.27) with the gravitational force:

$$\nu \rho_{\text{fl}} v R \sim \rho_{\text{sp}} g R^3. \quad (3.34)$$

The terminal velocity is

$$v = \frac{g R^2}{\nu} \frac{\rho_{\text{sp}}}{\rho_{\text{fl}}}. \quad (3.35)$$

We can increase our confidence in this expression by checking whether the correct variables are upstairs (a pictorial expression meaning in the numerator) and downstairs (in the denominator). Denser marbles should fall faster than less dense marbles, so ρ_{sp} should live upstairs. Gravity accelerates marbles, so g should also live upstairs. Viscosity slows marbles, so ν should live downstairs. The terminal velocity (3.35) passes these tests. We therefore have more confidence in our result (although the tests did not check the exponents, or the location of R).

Our result is correct, except that we neglected an important physical effect: buoyancy. Equivalently, we forgot one dimensionless group: $\Pi_3 \equiv \rho_{\text{fl}}/\rho_{\text{sp}}$. Buoyancy reduces the effective weight of the marble. Equivalently, it reduces the value of g . If $\rho_{\text{sp}} = \rho_{\text{fl}}$, then the effective g is 0; if $\rho_{\text{fl}} = 0$, then gravity retains its full effect. Between these two limits, the effective g varies linearly with ρ_{fl} . Why must the variation be linear? Gravity makes pressure in a fluid increase with depth; this increase causes buoyancy (Figure 3.8). So gravity is responsible for buoyancy. Gravitational force—and therefore buoyant force—is proportional to ρ_{fl} . So the effective g should vary linearly with ρ_{fl} . Therefore, we make the transformation

$$g \rightarrow g \left(1 - \frac{\rho_{\text{fl}}}{\rho_{\text{sp}}} \right). \quad (3.36)$$

The gravitational force (3.33) becomes

$$F_g \sim (\rho_{\text{sp}} - \rho_{\text{fl}}) g R^3, \quad (3.37)$$

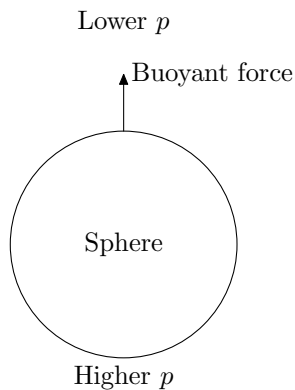


Figure 3.8. Gravity causes buoyancy. Because of gravity, the pressure at the bottom of the sphere (submerged in a fluid that is not shown) is greater than the pressure at the top. The result is an upward force: buoyancy.

and the terminal velocity becomes

$$v \sim \frac{gR^2}{\nu} \left(\frac{\rho_{\text{sp}}}{\rho_{\text{fl}}} - 1 \right). \quad (3.38)$$

If we had carried through the constants of proportionality, starting with the magic 6π from the drag force (3.28) and including the $4\pi/3$ that belongs in the gravitational force (3.33), then we would have found that the missing constant of proportionality in the terminal velocity (3.38) is $2/9$.

Using (3.38), we tried a home experiment to determine the viscosity of corn syrup. We dropped a glass marble with ($R = 0.8\text{ cm}$) in a tube of corn syrup, and measured $v \sim 2\text{ cm s}^{-1}$. With the magic $2/9$ factor, the viscosity is

$$\nu = \frac{2}{9} \frac{gR^2}{v} \left(\frac{\rho_{\text{sp}}}{\rho_{\text{fl}}} - 1 \right). \quad (3.39)$$

The density of glass is $\rho_{\text{sp}} \sim 2\text{ g cm}^{-3}$, and that of corn syrup is $\rho_{\text{sp}} \sim 1\text{ g cm}^{-3}$. With all the numbers, the viscosity estimate is

$$\nu \sim 0.2 \times \frac{1000\text{ cm s}^{-2} \times (0.8\text{ cm})^2}{2\text{ cm s}^{-2}} \times 1 \sim 60\text{ cm}^2\text{ s}^{-1}. \quad (3.40)$$

Many tables list μ , for which we predict $60\text{ g cm}^{-1}\text{ s}^{-1}$ (because $\rho_{\text{sp}} \sim 1\text{ g cm}^{-3}$).

A search on the World Wide Web for a site listing the viscosity of corn syrup yielded $60\text{ g cm}^{-1}\text{ s}^{-1}$ for deionized corn syrup, and $25\text{ g cm}^{-1}\text{ s}^{-1}$ for regular corn syrup. Presumably, the corn syrup that we purchased from the grocery store was merely regular corn syrup. Our estimate is therefore in error by roughly a factor of 2. The modest size of the factor indicates that we have included much of the relevant physics in our model. In this case, we have captured much of the physics (although we did insert the $2/9$ by fiat); the major uncertainty is in the viscosity of our corn syrup. (Another source of error is the effect

of the tube wall; this order-of-magnitude analysis assumed that the tube is infinitely wide, so that the marble feels no effects from the wall.)

In deriving the drag force (3.28), we assumed that the drag force was proportional to ν —which it is indeed in slow flows, or in highly viscous liquids. Is our flow sufficiently slow? The Reynolds' number—defined in (3.23)—is the correct measure of flow speed. Let's make sure that Re is small:

$$Re \sim \frac{vR}{\nu} \sim \frac{2 \text{ cm s}^{-1} \times 0.8 \text{ cm}}{60 \text{ cm}^2 \text{ s}^{-1}} \sim 0.03, \quad (3.41)$$

which is $\ll 1$. The flow oozes; so our assumption is self-consistent.

3.3.3 Turbulent limit

We now compute drag in the other flow extreme: high-speed, or turbulent, flow. We estimate the terminal velocity for a person thrown from a high-flying airplane. How high is the airplane flying? High enough so that the person reaches terminal velocity. You may reasonably question whether people frequently get thrown from airplanes. Fortunately, they do not. However, our results are generic. For example, even when a child rises from a chair, the airflow around her is high-speed flow. Say that the child is 20 cm wide, and that she rises with velocity 50 cm s⁻¹. Then,

$$Re \sim \frac{vR}{\nu_{\text{air}}} \sim \frac{50 \text{ cm s}^{-1} \times 20 \text{ cm}}{0.2 \text{ cm}^2 \text{ s}^{-1}} \sim 5000, \quad (3.42)$$

which is $\gg 1$. The flow is turbulent. Larger objects, such as adults and cars, create turbulence even when they travel only 10 cm s⁻¹. Most fluid flow—large raindrops falling in air, ships traveling in water, cyclists racing in air—is turbulent.

We now specialize the results of Section 3.3.1 to high-speed flow. To begin our analysis, we lie: We assume that a person is a sphere. It is a convenient approximation. Even if people are cylindrical in everyday life, a person thrown out of an airplane might, from a sensible fear reflex, tuck into fetal position, and vindicate our approximation. At high speeds (more precisely, at high Re), the flow is turbulent. Viscosity—which affects only slow flows, but does not influence the shearing and whirling of turbulent flows—becomes irrelevant. Let's see how much we can understand about turbulent drag, knowing only that drag is nearly independent of viscosity. The viscosity appears in only Π_1 . If viscosity disappears, so does Π_1 . This argument is slightly glib. More precisely, we remove ν from the list in Table 3.2, and search for dimensionless groups. From four variables, we can find one dimensionless group: the Π_2 from the marble example.

Why is drag independent of Re at high speeds? Equivalently, why can we remove ν from the list of variables and still get the correct form for the drag force? The answer is not obvious. Our construction of Re in (3.23)—as a ratio of v and V —provides a partial answer. A natural length in this problem is R , so from each velocity, we can form a quantity with units of time:

$$\begin{aligned} \tau_v &\equiv \frac{R}{v}, \\ \tau_V &\equiv \frac{R}{V} \sim \frac{R^2}{\nu}. \end{aligned} \quad (3.43)$$

Note that $Re \equiv \tau_V/\tau_v$. The quantity τ_v is the time that fluid takes to travel around the sphere (apart from constants). As we discuss in Section 5.5, kinematic viscosity is the

diffusion coefficient for momentum, and the time for momentum to diffuse a distance x is x^2/ν . So τ_V is the time that momentum takes to diffuse around the sphere (apart from constants). If $\tau_V \ll \tau_v$ —in which case, $Re \ll 1$ —then momentum diffuses before fluid travels around the sphere. Momentum diffusion equalizes velocities, if it has time. In this low- Re limit, it does. It prevents flow at the front from being radically different from the flow at the back. There is no turbulence. In the other limit— $\tau_V \gg \tau_v$, or $Re \gg 1$ —momentum diffusion gets soundly beaten by fluid flow, so the fluid is free to do whatever it wants: It is turbulent. The degree to which fluid flow beats momentum diffusion does not affect the flow; the flow is already unconstrained by viscosity. (This statement is a lie, but it's reasonably accurate.)

The dimensionless group that remains is

$$\Pi_2 = \frac{F_d}{\rho_{\text{fl}} R^2 v^2}. \quad (3.44)$$

The most general relation is $f(\Pi_2) = 0$, so $\Pi_2 = \Pi$, and

$$F_d \sim \rho_{\text{fl}} R^2 v^2. \quad (3.45)$$

We computed the drag force for a sphere; what about other shapes? So that the drag force generalizes to more complex shapes, we express it using the cross-sectional area of the object—here, $A = \pi R^2$, so $F_d \sim \rho_{\text{fl}} A v^2$. The constant of proportionality is called $c_d/2$, where c_d is the **drag coefficient**. The drag coefficient is a dimensionless measure of the drag force, and it depends on the shape of the object (on how streamlined the object is). Table 3.3 lists c_d for various shapes. (The drag coefficient is proportional to our function $\bar{f}(Re)$.) With the drag coefficient, the drag force is

$$F_d \sim \frac{1}{2} c_d \rho_{\text{fl}} v^2 A. \quad (3.46)$$

We For most purposes, we can assume that $c_d \sim 1$. The drag coefficient varies slightly with Re , so we lied a little when we said that, at high Re , drag is independent of viscosity (which appears in the Reynolds' number). The lie is sufficiently accurate for most order-of-magnitude purposes.

<i>Object</i>	c_d
Sphere	0.5
Cylinder	1.0
Flat plate	2.0
Car	0.4

Table 3.3. Approximate drag coefficients (c_d) of various objects at high speeds ($Re \gg 1$). The cylinder moves perpendicular to its axis; the flat plate moves perpendicular to its surface.

Now that we have an expression for the drag force, we estimate the gravitational force; computing it requires far less work. The gravitational force—including the buoyant force—does not care whether the flow is turbulent. So the result for viscous flow (3.37) applies here. For your convenience, we restate that result:

$$F_g \sim (\rho_{\text{sp}} - \rho_{\text{fl}})gR^3. \quad (3.47)$$

To find the terminal velocity, we equate F_g with the drag force (3.45):

$$\rho_{\text{fl}}R^2v^2 \sim g(\rho_{\text{sp}} - \rho_{\text{fl}})R^3. \quad (3.48)$$

The terminal velocity is

$$v \sim \sqrt{gR \frac{\rho_{\text{sp}} - \rho_{\text{fl}}}{\rho_{\text{fl}}}}. \quad (3.49)$$

We pause to sanity check this result. Are the right variables upstairs and downstairs? We consider each variable in turn.

- ρ_{fl} : The terminal velocity is smaller in a denser fluid (try running in a swimming pool), so ρ_{fl} should be in the denominator.
- g : Imagine a person falling on a planet that has a gravitational force stronger than that of the earth.⁶ The drag force does not depend on g , so gravity wins the tug of war, and v increases: g should be upstairs.
- ρ_{sp} : Imagine a person who has twice normal density (but has normal size). The gravitational force doubles. The drag force stays the same, as we saw using the fluid-is-a-computer argument in Section 3.3.1. So ρ_{sp} should be upstairs in (3.49).
- R : To determine where the radius lives, we require a more subtle argument, because changing R changes the gravitational force and the drag force. The gravitational force increases as R^3 ; the drag force increases as R^2 . At larger R , gravity wins; greater gravity means greater v , so R should be live upstairs.
- ν : At high speeds, viscosity does not affect drag, so ν should appear neither upstairs nor downstairs.

The velocity (3.49) passes all tests.

At last, we can compute the terminal velocity for the (involuntary) skydiver. We assume that he tucks into a fetal position, making him roughly spherical, with, say, $R \sim 40$ cm. A person is mostly water, so $\rho_{\text{sp}} \sim 1$ g cm⁻³. The density of air is $\rho_{\text{fl}} \sim 10^{-3}$ g cm⁻³, so $\rho_{\text{fl}} \ll \rho_{\text{sp}}$; buoyancy is not an important effect. In (3.49), we can neglect ρ_{fl} in $\rho_{\text{sp}} - \rho_{\text{fl}}$. We put these numbers into the terminal velocity (3.49):

$$v \sim \left(1000 \text{ cm s}^{-2} \times 40 \text{ cm} \times \frac{1 \text{ g cm}^{-3}}{10^{-3} \text{ g cm}^{-3}} \right)^{1/2} \sim 6 \cdot 10^3 \text{ cm s}^{-1}, \quad (3.50)$$

or 120 mph (200 kph).

6. Gravity partially determines atmospheric pressure and density. Holding the atmospheric density constant while increasing gravity might be impossible in real life, but we can do it easily in a thought experiment.

This calculation assumed that $Re \gg 1$. We now check that assumption. We do not need to calculate Re from scratch, because, as we worked out in (3.42), a child ($R \sim 20\text{ cm}$) rising from her chair ($v \sim 50\text{ cm s}^{-1}$) creates a flow with $Re \sim 5000$; this flow is turbulent. The flow created by the skydiver is faster by a factor of 100, and the skydiver is larger by a factor of 2, so the flow is certainly turbulent. Scaling the Reynolds' number from (3.42), we get

$$\begin{aligned} Re &\sim Re_{\text{child}} \times \frac{v_{\text{skydiver}}}{v_{\text{rising}}} \times \frac{R_{\text{skydiver}}}{R_{\text{child}}} \\ &\sim 5000 \times 100 \times 2 = 10^6, \end{aligned} \quad (3.51)$$

which is $\gg 1$. The flow is extremely turbulent.

Now that we have found the terminal velocity, let's extract the pattern of the solution. The order that we followed was assume, derive, calculate, and check (Figure 3.9). This order is more powerful than the usual order of derive and calculate. Without knowing whether the flow is fast or slow, we cannot derive a closed-form expression for F_d (such a derivation is probably beyond present-day understanding of turbulence). Blocked by this mathematical Everest, we would remain trapped in the derive box; we would never determine F_d ; and we would never realize that Re is large (the assume box), which is the assumption that makes it possible to estimate F_d . We used the same, correct order when we solved for the terminal velocity in corn syrup; Figure 3.10 shows how similar the logic is.

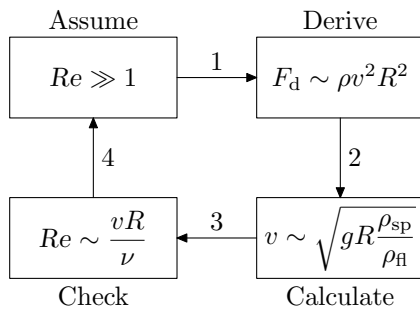


Figure 3.9. Correct solution order for terminal velocity in turbulent flow. We simplified the solution procedure by starting in the assume box. Step 1: On that assumption, we estimated the drag force (the derive box). Step 2: From the drag force, we estimated the terminal velocity (the calculate box). Step 3: From the terminal velocity, we estimated the Reynolds' number (the check box). Step 4: To close the loop, we verified the starting condition. Note: For compactness, the terminal-velocity formula ignores the normally small effect of buoyancy.

3.3.4 Combining solutions from the two limits

We computed the drag force in two limits. Let's compare the results, to see whether we can combine them. We can write the solution for viscous flow (3.28) in terms of the Reynolds' number and cross-sectional area:

$$F_d = \frac{1}{\pi} \rho_{fl} A v^2 \bar{f}(Re). \quad (3.52)$$

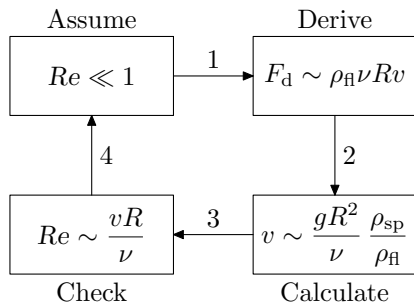


Figure 3.10. Correct solution order for terminal velocity in viscous flow. Note the similarities with Figure 3.9.

We showed that $\bar{f}(x) = \Pi/x$, where the magic constant $\Pi = 6\pi$ comes from honest physics, which we did not do.⁷ So

$$F_d = \frac{6}{Re} \rho_f A v^2. \quad (3.53)$$

Compare this viscous drag force to the turbulent drag force (3.46). If we substitute $c_d = 12/Re$, then (3.46) becomes (3.53). So we can use the drag formula (3.46) for both high- and low-speed flow, provided that we let c_d vary with Re . For high Re , we quoted $c_d \sim 1/2$. If the low- Re approximation were valid at all Reynolds' numbers, then c_d would cross $1/2$ near $Re \sim 24$, at which point we would expect the high- Re approximation to take over. The crossing point is a reasonable estimate for the transition between low- and high-speed flow. Experimental data place the crossover nearer to $Re \sim 5$, at which point $c_d \sim 2$. If a dimensionless variable, such as Re , is close to unity, calculations become difficult. Approximations that depend on a quantity being either huge or tiny are no longer valid. When all quantities are roughly the same size, you cannot get rid of many terms in your equations. To get results in these situations, you have to do honest work: You must solve the nonlinear equations (3.22) numerically or analytically.

3.4 What you have learned

- Every valid physical equation can be written in a form without units. To find such forms, you follow these steps:
 1. Write down—by magic, intuition, or luck—the physically relevant variables. For illustration, let's say that there are n of them.
 2. Determine the units of each variable. Count how many independent dimensions these variables comprise. Call this number r . Usually, length, mass, and time are all that you need, so $r = 3$.
 3. By playing around, or by guessing with inspiration, find $n - r$ independent dimensionless combinations of the variables. These combinations are the dimensionless groups, or Pi variables—named for the Buckingham Pi theorem.

7. We agree with the Fowlers [17] and Dupré [12]: *that* is restrictive, *which* is not. If we had said ...*from honest physics that we did not do*, we would imply that, somewhere in this book, we did do honest physics. We plead not guilty.

4. Write down the most general relation that uses these groups. Then try to eliminate dimensionless groups, or to restrict the form of the relation, using physical arguments.
 5. Using physical arguments, simplify the dimensionless relation by eliminating dimensionless groups, or by otherwise constraining the form of the relation.
- In dimensionless form, even fluid mechanics is simple.
 - The Reynolds' number is a dimensionless measure of flow speed; it distinguishes viscous from turbulent flow, and compares the momentum transport time to the fluid transport time.
 - When you solve a difficult problem—such as computing the drag force—simplify by assuming an extreme case: Assume that one or more of the dimensionless variables are nearly zero or nearly ∞ .
 - Using your solution, check your assumption!

In the next two chapters, we use these dishonest methods, along with physical reasoning, to estimate the properties of materials.

4 MATERIALS I

Why are there no 100 km mountains on earth (Section 4.3.2)? Why must a Thanksgiving turkey cook all morning (Example 5.2)? How large are white dwarfs (Section 4.4)? The answers to these questions depend on the mechanical and thermal properties of materials. We estimate these properties by using the techniques of the previous chapters, and introduce a new technique, balancing. In this chapter, we study mechanical properties of materials: their density, strength, and surface tension. In Chapter 5, we study thermal properties of materials.

Our estimates of material properties depend on the **atomic theory**. In Feynman's words:

If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, what statement would contain the most information in the fewest words? I believe it is the *atomic hypothesis* (or the *atomic fact*, or whatever you wish to call it) that *all things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another*. In that one sentence, you will see, there is an *enormous* amount of information about the world. . . [15, p. 1-2]

The atomic theory was first stated by Democritus. Bertrand Russell, himself a scientist and philosopher, describes early Greek science and philosophy with wit, sympathy, and insight in his *History of Western Philosophy* [53]. Democritus could not say much about the properties of atoms. With our modern knowledge of classical and quantum mechanics, we can go farther. We begin by determining how large atoms are. After we understand what forces determine the sizes of atoms, we extend our understanding to molecules: We divide and conquer.

4.1 Sizes

We begin our study of atoms with a simple picture—always draw a simple picture—of the simplest atom, hydrogen (Figure 4.1). What is the radius of the orbit?

4.1.1 Dimensional analysis

We estimate the radius first using dimensional analysis. We make a list of relevant variables by considering what physics determines the properties of hydrogen. The electrostatic force holds the electron in orbit, so e , the charge on the electron, is certainly on the list. We use e^2 instead of e , because, in cgs units, e^2 has familiar dimensions: $[E][L]$. If we had used e , then we would wonder how to combine esu with more familiar dimensions. [In mks units, we would have to introduce the permittivity of free space, ϵ_0 , along with the charge. Symbolic results in electrostatics often easier to understand in cgs units than in mks units. Numerical estimates are often simpler in mks units, because tables quote values in ohms, or Coulombs, rather than in statohms or esu. If you want both to think and to

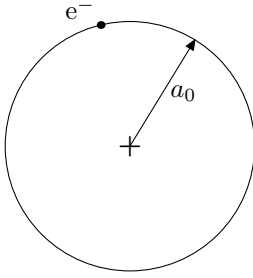


Figure 4.1. Hydrogen. A proton (the “+”) holds an electron in jail. The jail radius is a_0 , the **Bohr radius**.

calculate, become comfortable with both systems.] The mass of the electron m_e determines the acceleration that the electrostatic force produces, so we include m_e . The mass of the proton, m_p , might also be relevant. From these four variables— a_0 , e^2 , m_e , and m_p —we can form one dimensionless group,

$$\Pi_1 \equiv \frac{m_e}{m_p}. \quad (4.1)$$

That group does not contain a_0 ; we are stuck. Our failure to find a suitable dimensionless group tells us that we’ve left out physics that determines a_0 .

Var.	Units	Description
a_0	[L]	radius
e^2	[E][L]	charge ²
\hbar	[E][T]	quantum mechanics
m_p	[M]	proton mass
m_e	[M]	electron mass

Table 4.1. Variables that determine the size of hydrogen. The notation [E] is a shorthand for $[M][L]^2[T]^{-2}$.

We get a clue to this missing physics, because the picture of hydrogen in Figure 4.1 cannot be classical. A classical electron, moving in a circle, would radiate, lose energy, and spiral into its doom as the proton swallowed it; classical atoms cannot exist. Fortunately, quantum mechanics comes to the rescue. Unfortunately, quantum mechanics is a large, fearsome subject. Thanks to dimensional analysis, we do not have to understand quantum mechanics to compute results based on it. For the purposes of dimensional analysis, the whole content of quantum mechanics is contained in another constant of nature: \hbar , whose value is $1.05 \cdot 10^{-27}$ erg s. By adding \hbar to our four variables, we get the list in Table 4.1.

One more variable produces one more dimensionless group. We are sunk if this group does not contain a_0 , so let’s assume that it does. How can we make a length from the other four variables, to cancel the length in a_0 ? We divide and conquer. Because $[e^2]$ is [E][L], we make an energy from e^2 , \hbar , and m_e (we could also have used m_p , but using both m_p and

m_e allows us no additional units), and use the constructed energy to cancel the energy in $[E][L]$. Notice that e^2/\hbar is a velocity, so $m_e(e^2/\hbar)^2$ is an energy, with which we can cancel energy from e^2 . We do so, and construct the length

$$\frac{e^2}{m_e(e^2/\hbar)^2}, \quad (4.2)$$

which simplifies to $\hbar^2/m_e e^2$. The new dimensionless group is a_0 divided by this length:

$$\Pi_2 = \frac{a_0}{\hbar^2/m_e e^2}. \quad (4.3)$$

The other group is $\Pi_1 = m_e/m_p$, so we write $\Pi_2 = f(\Pi_1)$, where f is a dimensionless function, or, equivalently,

$$a_0 = \frac{\hbar^2}{m_e e^2} f\left(\frac{m_e}{m_p}\right). \quad (4.4)$$

The proton is much more massive than the electron, so we exaggerate and pretend that $m_e/m_p = 0$; we **simplify by exaggerating**. If $f(x)$ has a limit as $x \rightarrow 0$, then (4.4) becomes

$$a_0 = \Pi \frac{\hbar^2}{m_e e^2}, \quad (4.5)$$

where the constant Π is $f(0)$. This important result deserves a cross-check, which we perform in Section 4.1.2.

Before we do the cross-check, we discuss the validity of assuming that $f(0)$ exists. If we had constructed Π_2 from m_p instead of from m_e , then $f(x)$ would be $\sim 1/x$, which does not have a finite limit. How can you decide which mass avoids this problem, short of trying both, or getting lucky? There is no foolproof procedure, but we can formulate rules of thumb.

First, choose dimensionless groups with minimal overlap. In the definition of Π_2 , we could have used $m_p + m_e$ instead of m_e . This choice would make m_p and m_e appear in two dimensionless groups; the choice that we made—to use m_e in Π_2 , and m_e and m_p in Π_1 —has less overlap among the original quantities. When we consider the limit of f , we usually hope to eliminate one dimensionless group. This elimination is more likely if there are fewer overlaps.

Second, decide which variable is likely to vanish, and segregate it in one dimensionless group. In this example, the proton and electron mass are so disparate (their ratio is ~ 2000) that we immediately imagine a world where the disparity is more pronounced. There are many such imaginary worlds; $m_p \sim 1000$ kg and $m_e \sim 1$ g, or perhaps $m_p \sim 10^{-50}$ g and $m_e \sim 10^{-60}$ g. These examples are unreasonable. If possible, we vary only one quantity, so as to tell the fewest lies about the actual world. In our imaginary world, should we reduce m_e or increase m_p ? To decide, we use physical reasoning. The momenta of the proton and the electron are equal (and opposite) with, say, magnitude P . A typical kinetic energy is $E \sim P^2/m$. The mass is downstairs, so the heavier particle (the proton) contributes much less energy than the lighter particle contributes. We therefore imagine that m_p gets larger; that choice eliminates only a minor contributor to the total energy. This argument shows

that hydrogen has a physical limit when $m_p \rightarrow \infty$ and m_e stays fixed. Therefore, $f(m_e/m_p)$ has a limit at 0. What would happen if we had not segregated m_p ? Suppose, for example, that we chose $\Pi_1 = a_0 m_p e^2 / \hbar^2$ and $\Pi_2 = m_e / m_p$. As with our previous pair of dimensionless groups, we would have $\Pi_1 = g(\Pi_2)$, where g is a dimensionless function, or

$$a_0 = \frac{\hbar^2}{m_p e^2} g(m_e/m_p). \quad (4.6)$$

Now g behaves badly. As $m_p \rightarrow \infty$, the factor in front goes to 0; so g must go to infinity. To make progress in spite of g 's intransigence, we import the knowledge that a_0 is independent of m_p as $m_p \rightarrow \infty$. Therefore, we must have $g(x) \sim 1/x$, so that $g(x)$ can cancel the m_p that sits in front. If we segregate m_p to begin with—as we did in (4.1) and (4.3)—then we can simply discard Π_1 without going through this more involved argument.

4.1.2 Physical meaning

For the second computation of a_0 , we use a method that substitutes physical reasoning for dimensional analysis. We get a physical interpretation of the result (4.5). We want to find the orbit radius with the lowest energy (the ground state) of hydrogen. The energy is a sum of kinetic and potential energy: kinetic from motion of the electron, and potential from electrostatic attraction.

What is the origin of the kinetic energy? The electron does not orbit in any classical sense. If it did, it would radiate its energy and spiral into the nucleus. Instead, according to quantum mechanics, the proton confines the electron to some region of size r .¹ Confinement gives the electron energy according to the **uncertainty principle**, which says that $\Delta x \Delta p \sim \hbar$, where Δx is the position uncertainty, and Δp is the momentum uncertainty of the electron. In hydrogen, $\Delta x \sim r$, as shown in Figure 4.2, so $\Delta p \sim \hbar/r$. The kinetic energy of the electron is

$$\text{KE} \sim \frac{(\Delta p)^2}{m_e} \sim \frac{\hbar^2}{m_e r^2}. \quad (4.7)$$

This energy is the **confinement energy**, or the **uncertainty energy**. We use this idea many times in this book; our next application for it is in Section 4.4.

The potential energy is the classical expression

$$\text{PE} \sim -\frac{e^2}{r}. \quad (4.8)$$

Here the \sim sign indicates the electron is not precisely at a radius r ; rather, its typical, or characteristic, distance is r .

The total energy is the combination

$$E = \text{PE} + \text{KE} \sim -\frac{e^2}{r} + \frac{\hbar^2}{m_e r^2}. \quad (4.9)$$

1. In an order-of-magnitude estimate, we ignore details, such as the precise path that an object follows, and replace them by a typical size and, perhaps, speed. Quantum mechanics justifies this slothfulness as a principle of physics.

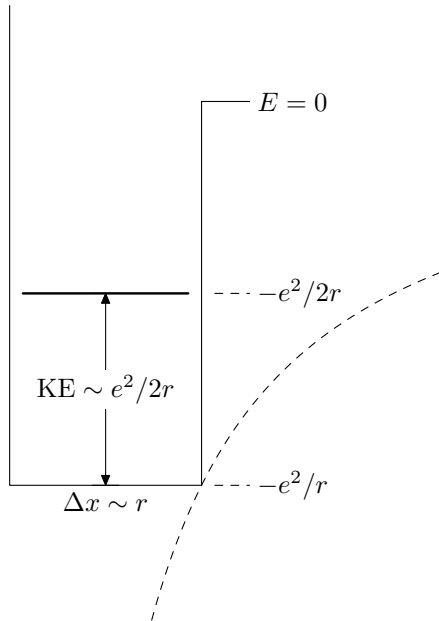


Figure 4.2. Order-of-magnitude hydrogen. We replace the Coulomb potential (dashed curve) by a box potential (solid line) with the same characteristic width and height. This rectangle trick is familiar from Section 2.1.4. The total energy of the electron (thick line) is the sum of the potential energy, $-e^2/r$, and the kinetic energy, $e^2/2r$.

The two energies compete. At small r , kinetic energy wins, because of the $1/r^2$; at large r , potential energy wins, because it goes to zero less rapidly. Is there a minimum combined energy at some intermediate value of r ? There has to be. At small r , $\partial E/\partial r$ is negative; at large r , it is positive; at an intermediate r , the $\partial E/\partial r$ must be zero. The potential has a minimum at this r . The location is easy to estimate if we write (4.9) in dimensionless form; such a rewriting is not necessary in this example, but it is helpful in more complicated examples. In deriving (4.3), we constructed a length, $l = \hbar^2/m_e e^2$. So let's define scaled length as $\bar{r} \equiv r/l$. Furthermore, e^2/l is an energy, so let's define scaled energy as $\bar{E} \equiv E/(e^2/l)$. Then (4.9) becomes

$$\bar{E} \sim -\frac{1}{\bar{r}} + \frac{1}{\bar{r}^2}. \quad (4.10)$$

We can find the minimum-energy \bar{r} by calculus, which gives $\bar{r}_{\min} = 2$. Or, we can use **order-of-magnitude minimization**: When two terms compete, the minimum occurs when the terms are roughly equal (Figure 4.3). We therefore equate \bar{r}^{-1} and \bar{r}^{-2} ; we get $\bar{r}_{\min} \sim 1$. In normal units, the minimum-energy radius is $r_{\min} = l\bar{r}_{\min} = \hbar^2/m_e e^2$, which is exactly the Bohr radius; in this case, the sloppiness in definition cancels the error in order-of-magnitude minimization.

To justify this method, consider a reasonable general form for E :

$$E(r) = \frac{A}{r^n} - \frac{B}{r^m}. \quad (4.11)$$

This form captures the important feature of (4.9): The two parts compete. To find the minimum, we solve $E'(r_{\min}) = 0$, or

$$-n \frac{A}{r_{\min}^{n+1}} + m \frac{B}{r_{\min}^{m+1}} = 0. \quad (4.12)$$

The solution is

$$\frac{A}{r_{\min}^n} = \frac{n}{m} \frac{B}{r_{\min}^m}. \quad (4.13)$$

The order-of-magnitude method minimizes (4.11) by equating A/r^n and B/r^m , so it lacks the n/m factor in (4.13). The ratio of the two estimates for r_{\min} is

$$\frac{\text{order-of-magnitude estimate}}{\text{calculus estimate}} \sim \left(\frac{n}{m}\right)^{1/(m-n)}, \quad (4.14)$$

which is < 1 if $n > m$. So the order-of-magnitude method underestimates minima and overestimates maxima (for maxima, the same argument carries through, except that m and n change places).

The potential between nonpolar atoms and molecules is well approximated with $m = 6$ and $n = 12$:

$$U(r) \sim \frac{A}{r^{12}} - \frac{B}{r^6}. \quad (4.15)$$

This potential is known as the Lennard–Jones potential. It describes the interaction of two nonpolar atoms (or molecules). Such atoms (for example, helium or xenon) have no permanent dipole moment. However, the charge distribution around the nucleus (and therefore the dipole moment) fluctuates. This fluctuation favors opposite fluctuations in the other atom. The induced dipole moments have opposite directions; therefore, the atoms attract each other, no matter what the sign of the fluctuation. Let's use this physical picture to explain the r^{-6} dependence in the attractive part of the potential. We begin by estimating the force on a dipole of moment μ sitting in an electric field E . (In this paragraph, E is an electric field, not an energy. Sorry!) A dipole is an ideal version of a positive and a negative charge ($+q$ and $-q$) separated by a small distance d ; the dipole moment is the product $\mu = qd$. The force on a dipole is therefore a sum of the force on the positive charge and on the negative charge. So

$$F \sim q(E(r+d) - E(r)). \quad (4.16)$$

Here, $E(r)$ is the field at a distance r from the other atom. Because d is small, this force is approximately $qdE'(r)$, or $\mu E'(r)$, where $E'(r) = \partial E/\partial r$. We assume that $E \propto r^b$, for some constant b ; then $E'(r) \sim E(r)/r$ and $F \sim \mu E/r$. What is the dipole moment? The dipole is induced by the field of the other atom; that field polarizes this atom. So $\mu = aE$, where a is the polarizability of the atom (a characteristic of the atom that is independent of field, as long as the field is not too strong). Then $F \sim aE^2/r$. What is the field E ? The field from a dipole is proportional to r^{-3} , so $F \propto r^{-7}$. Therefore, $U \propto r^{-6}$, which explains the attractive part of the Lennard–Jones potential. The r^{-12} term represents short-range repulsion; it is a fit to experimental data, and cannot be estimated with these simple arguments.

For the Lennard–Jones potential, the ratio n/m is 2, so the order-of-magnitude estimate of r_{\min} is in error by a factor of $2^{1/6} \sim 1.1$. Here is another example: Minimize the function

$f(r) = e^r + 1/r^2$. The minimum is at $r_0 \approx 0.925$ (the solution of $e^r = 2/r^3$), whereas the order-of-magnitude method predicts that $r_0 \sim 0.7$ (the solution of $e^r = 1/r^2$). So even when the two contributors to f are not power laws, the order-of-magnitude method gives a reasonable answer. [To understand the range of validity of this method, construct functions for which it is grossly inaccurate.]

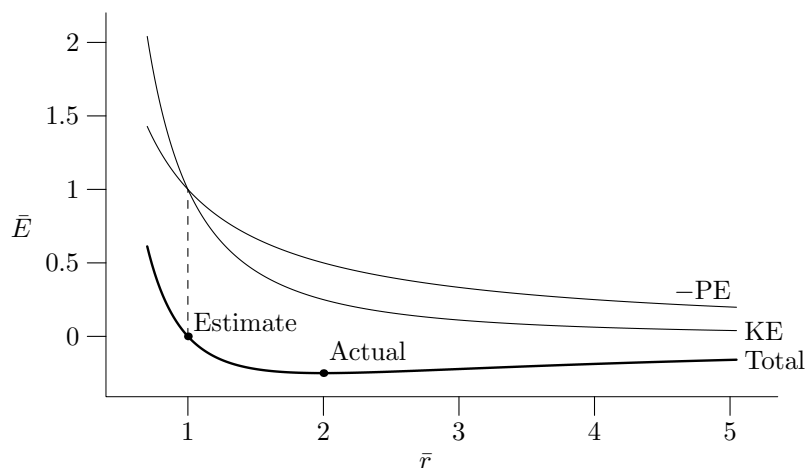


Figure 4.3. Order-of-magnitude calculus: minimizing scaled energy \bar{E} versus scaled bond length \bar{r} . The scaled energy is the sum of potential and kinetic energy. In absolute value, potential and kinetic energy are equal at $\bar{r}_{\min} = 1$, so the order-of-magnitude minimum is at $\bar{r}_{\min} = 1$. Honest calculus produces $\bar{r}_{\min} = 2$.

Now we return to the original problem: determining the Bohr radius, for which order-of-magnitude minimization predicts the correct value. Even if the minimization were not so charmed, there would be no point in doing the a calculus minimization; the calculus method is too accurate. The kinetic-energy estimate uses a crude form of the uncertainty principle, $\Delta p \Delta x \sim \hbar$, whereas the true statement is that $\Delta p \Delta x \geq \hbar/2$. It uses a crude definition of Δx , that $\Delta x \sim r$. It uses the approximate formula $\text{KE} \sim (\Delta p)^2/m$. This formula not only contains m instead of $2m$ in the denominator, but also assumes that we can convert Δp into an energy as though it were a true momentum, rather than merely a crude estimate of the root-mean-square momentum. After making these approximations, it's pointless to minimize using the elephant gun of differential calculus. The order-of-magnitude method is as accurate as the approximations that we used to derive the energy.

Our name for this method of equating competing terms is **balancing**; we balanced the kinetic energy against the potential energy and found that $a_0 \sim \hbar^2/m_e e^2$. Nature could have been unkind: The potential and kinetic energies could have been different by a factor of 10 or 100. But she is kind: The two energies are roughly equal, up to a constant that is nearly 1. “Nearly 1” is also called **of order unity**. This rough equality occurs in many examples that we examine; we often get a reasonable answer if we pretend that two energies (or, in general, two quantities with the same units) are equal. When the quantities are potential and kinetic energy, the virial theorem, which we discuss in Section 4.2.1, protects us against

large errors.

4.1.3 Numerical evaluation

Now that we have derived

$$a_0 \sim \frac{\hbar^2}{m_e e^2} \quad (4.17)$$

in two ways, we evaluate a_0 : We put in numbers. You might think that we need to know or look up \hbar , e , and m_e . Not so. We can avoid looking up \hbar , m_e , and e , by using three useful combinations of fundamental constants. The first two are round values that use eV as the unit of energy:

$$\hbar c \simeq 2000 \text{ eV } \text{\AA}, \quad (4.18)$$

$$m_e c^2 \simeq 500 \text{ keV}. \quad (4.19)$$

The third combination is

$$\alpha \equiv \frac{e^2}{\hbar c} \simeq \frac{1}{137} \sim 0.01. \quad (4.20)$$

This combination is the **fine-structure constant**, a dimensionless measure of the strength of the electrostatic force. Dimensionless numbers, such as α , are worth memorizing, because their value is the same in any unit system.

The Bohr-radius estimate (4.17) contains e^2 downstairs, so we multiply by $\hbar c/\hbar c$ to manufacture α^{-1} :

$$a_0 \sim \frac{\hbar^2}{m_e \hbar c} \frac{\hbar c}{e^2}. \quad (4.21)$$

This expression contains m_e , so we multiply by c^2/c^2 . After canceling one $\hbar c$ in the numerator and in the denominator, we find that

$$a_0 \sim \frac{\hbar c}{m_e c^2} \frac{\hbar c}{e^2}. \quad (4.22)$$

We could have gone directly from (4.17) to (4.22) by multiplying by c^2/c^2 . We wrote out the individual steps to distinguish two tricks that often go together. The first trick is to insert $\hbar c/\hbar c$. The second trick is to insert c^2/c^2 to make $m_e c^2$, which has the round value 500 keV. Often, if you add enough powers of c to convert every \hbar into $\hbar c$, you also convert every m_e into $m_e c^2$. In such cases, the two tricks combine into one, which here is multiplication by c^2/c^2 .

Now we can put easy-to-remember numbers into (4.22) and determine the Bohr radius:

$$a_0 \sim \frac{(2000 \text{ eV } \text{\AA})}{5 \cdot 10^5 \text{ keV}} \times 100 \sim 0.5 \text{ \AA}. \quad (4.23)$$

Knowing the size of hydrogen, we can understand the size of more complex atoms. Hydrogen is the simplest atom; it has one electron, and therefore one energy shell. The second row of the periodic table contains elements with two shells; the third row contains elements with three shells. The most abundant elements on earth (oxygen, carbon, silicon) come from the second and third rows. As a rule of thumb, the diameter of an atom with n shells is $n \text{ \AA}$, for $n \leq 3$; for $n > 3$, the diameter is still 3 \AA , because the extra nuclear charge in those atoms drags the electrons closer, and makes up for the increased number of shells.

The following argument expands on this statement about extra nuclear charge. Consider an atom with atomic number Z ; it has Z protons and Z electrons. The outermost electron moves in a potential created by the Z protons and the $Z - 1$ other electrons. We cannot easily calculate the charge distribution, so we need to simplify. Imagine that the other electrons orbit inside the outermost electron. This assumption is a crude approximation; it does not account for important facts from quantum mechanics, such as the Pauli exclusion principle,² nor does it accurately represent atoms in which two or more electrons are in the outermost shell. However, it is a simple assumption, and has truth in it, so let's use it. What charge distribution does the outermost electron see? It sees a nucleus with charge e (effectively, a single proton): The Z protons and the $Z - 1$ electrons almost cancel. An outermost electron orbits a single proton—this configuration is the description of hydrogen. So the environment of the outermost charge is independent of Z , and every large- Z atom is the size of hydrogen. Part of this conclusion is reasonably accurate: that every large- Z atom is a similar size. Part of the conclusion is not correct: that the size is the size of hydrogen. It is incorrect because of the extreme approximation in assuming that every other electron orbits inside the outermost electron, and because it neglects the Pauli exclusion principle. We retain the reasonably correct part, and use $a \sim 3 \text{ \AA}$ for a typical atomic diameter.

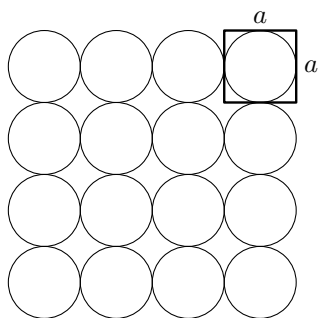


Figure 4.4. Packing of atoms in a solid or liquid. For clarity, we draw only a two-dimensional slice of the substance. In this crude picture, each atom occupies a cube of side length $a \sim 3 \text{ \AA}$, or a volume of $\sim 30 \text{ \AA}^3$ or $3 \cdot 10^{-23} \text{ cm}^3$.

4.1.4 Densities

From atomic sizes, we can estimate densities. An atom is a positive nucleus surrounded by a cloud of negative charge. A solid or liquid contains atoms jammed together. The electron cloud limits how closely the atoms can approach each other, because the repulsive force between the electron clouds is large when the clouds overlap. At large distances (relative to the Bohr radius), two atoms hardly interact. Between these extremes, there is a minimum-energy distance: a , the diameter of the electron cloud. So $a \sim 3 \text{ \AA}$ is also a typical interatomic spacing in a solid or liquid.

2. See any textbook on quantum mechanics for an explanation of the Pauli principle; Gasiorowicz's text [18] is particularly compact and clear.

Let A be the atomic mass of the atom; A is roughly the number of protons and neutrons in the nucleus. Although it is called a mass, A is dimensionless. Each atom occupies a cube of side length $a \sim 3 \text{ \AA}$ (Figure 4.4), and has mass Am_p . The density of the substance is

$$\rho = \frac{\text{mass}}{\text{volume}} \sim \frac{Am_p}{(3 \text{ \AA})^3}, \quad (4.24)$$

or, with $m_p \sim 2000m_e \sim 2 \cdot 10^{-24} \text{ g}$,

$$\rho \sim \frac{A \times 2 \cdot 10^{-24} \text{ g}}{3 \cdot 10^{-23} \text{ cm}^3} \sim \frac{A}{15} \text{ g cm}^{-3}. \quad (4.25)$$

In Table 4.2, we compare this estimate against reality. This estimate explains why most densities lie between 1 and 10 g cm^{-3} : Most everyday elements have atomic masses between 15 and 150.³ We even threw into the pack a joker—water, which is not an element—and our estimate was accurate to 20 percent. In Example 4.1, we answer a question that may be painfully familiar if you have moved house or apartment.

<i>Element</i>	$\rho (\text{g cm}^{-3})$	
	<i>Estimated</i>	<i>Actual</i>
Li	0.47	0.54
H ₂ O	1.2	1.0
Si	1.87	2.4
Fe	3.73	7.9
Hg	13.4	13.5
Au	13.1	19.3
U	15.9	18.7

Table 4.2. Actual and estimated densities of solids and liquids. The estimate is from (4.25).

Example 4.1 How heavy is a small box filled with books?

Books are mostly paper; as we reasoned in Section 2.1.1, paper has the same density as water, so $\rho_{\text{book}} \sim 1 \text{ g cm}^{-3}$. In the United States, the canonical book box is the small United Parcel Service box. Its volume is $60 \text{ cm} \times 30 \text{ cm} \times 30 \text{ cm} \sim 5 \cdot 10^4 \text{ cm}^3$, so its mass is $m \sim 50 \text{ kg}$ —approximately the mass of a person. It is no wonder that these boxes are so heavy, and no wonder that they are no larger.

3. This density estimate also shows why, for materials physics, cgs units are more convenient than mks units. A typical cgs density is 3, which is a modest, easy-to-work with number. A typical mks density is 3000, which is unwieldy.

4.2 Energies

Part of the order-of-magnitude picture of materials is the spacing between atoms (bond size); another part is the force or interaction energy between atoms (bond energy). Following the pattern of Section 4.1, we first estimate the **binding energy** of hydrogen, and then extend our understanding to more common materials.

4.2.1 Binding energy of hydrogen

In hydrogen, the binding energy is the energy required to drag the electron infinitely far from the proton. In more complex atoms, it is the energy to remove all the electrons. We can estimate the binding energy by first estimating the potential energy in hydrogen. The potential energy of an electron and a proton separated by the Bohr radius is

$$\text{PE} \sim -\frac{e^2}{a_0} \sim -\frac{m_e e^4}{\hbar^2}. \quad (4.26)$$

The binding energy is $-E_{\text{total}}$, where E_{total} is the total energy—it includes kinetic as well as potential energy. What is the kinetic energy? The **virial theorem** says that, in a $1/r^n$ potential ($n = 2$ for hydrogen),

$$\text{PE} = -n\text{KE}. \quad (4.27)$$

Therefore, $E_{\text{total}} = \text{PE} + \text{KE} = \text{PE}/2$, and the binding energy is

$$E_0 = -E_{\text{total}} = \frac{1}{2} \frac{m_e e^4}{\hbar^2}. \quad (4.28)$$

To evaluate this energy, we create $(\hbar c)^2$ and $m_e c^2$ by multiplying E_0 by c^2/c^2 :

$$\begin{aligned} E_0 &\sim \frac{1}{2} m_e c^2 \frac{e^4}{(\hbar c)^2} = \frac{1}{2} m_e c^2 \alpha^2 \\ &\simeq \frac{1}{2} \times 5.1 \cdot 10^5 \text{ eV} \times \frac{1}{137^2} = 13.6 \text{ eV}. \end{aligned} \quad (4.29)$$

By luck, the errors in our approximations have canceled: The honest-physics factor is 1. Equation (4.29) is the correct ground-state energy of hydrogen (neglecting relativistic effects, such as spin). We can rewrite the binding energy as $m_e(\alpha c)^2/2$, which is the kinetic energy of an electron with $v = \alpha c$: the fine-structure constant is the velocity of an electron in **atomic units**, where charge is measured in units of e , length in units of a_0 , and velocity in units of c .

For future reference, we quote useful energy conversions:

$$\begin{aligned} 1 \text{ eV} &\simeq 1.6 \cdot 10^{-12} \text{ erg} \\ &\simeq 1.6 \cdot 10^{-19} \text{ J}, \end{aligned} \quad (4.30)$$

$$1 \text{ cal} \simeq 4 \text{ J}, \quad (4.31)$$

and

$$\begin{aligned} 1 \text{ eV/molecule} &\simeq 25 \text{ kcal mole}^{-1} \\ &\simeq 100 \text{ kJ mole}^{-1}. \end{aligned} \quad (4.32)$$

4.2.2 Covalent- and ionic-bond energies

Covalent and ionic bonds are formed by attractions between electrons and protons; the hydrogen atom is a crude model of this interaction. The main defect of this model is that the electron–proton distance in a hydrogen atom is much smaller than it is in most materials. In most materials, the distance is roughly $a \sim 3 \text{ \AA}$, rather than $a_0 \sim 0.5 \text{ \AA}$. For covalent and ionic bonds—their binding energy is from the electrostatic attraction of monopoles—the binding energy is smaller than E_0 by a factor of 6: $E_{\text{bond}} \sim 2 \text{ eV}$. The factor of 6 occurs because $a \sim 6a_0$, and electrostatic energy scales as $E \propto 1/r$. [Scaling E is more direct than is evaluating e^2/a from scratch. It does not clutter the derivation, or our thinking, with irrelevant information, such as the value of \hbar or α .] Table 4.3 lists bond energies. For bonds that include carbon, oxygen, or hydrogen—which have $a \sim 1.5 \text{ \AA}$ —we expect $E_{\text{bond}} \sim 4 \text{ eV}$. This expectation is confirmed by the tabulated values. Van der Waals bonds are much weaker than covalent or ionic bonds, as you would expect from the high reciprocal powers of r in the Lennard–Jones potential (4.15).

<i>Bond</i>	<i>Energy (eV)</i>
C–C	6.3
H–H	4.5
Na–Cl	4.3
H–O	4.4
Fe–O	4.0
C–H	3.5
Si–Si	3.3
Hg–S	2.2
Cd–S	2.1
Zn–S	2.1

Table 4.3. *Approximate covalent and ionic bond energies. Source: CRC Handbook of Chemistry and Physics [38, pages 9-123–9-128].*

The **cohesive energy** is the energy required to break the weakest type of bond in an object. In water, it is the energy required to remove a water molecule from the liquid into the vapor. In hydrogen, it is the energy required to break the electron–proton bond (given by (4.29)). We denote the cohesive energy by ϵ_c . The typical magnitude of ϵ_c indicates why the electron–Volt is a convenient unit in materials physics: Cohesive energies measured in eV are of order unity (for strongly bonded substances).

4.3 Elastic properties

The size of white dwarfs or of raindrops, the speed of sound—these quantities we can estimate using our knowledge of atomic sizes and energies, and the methods that we have developed. We begin with the speed of sound. It illustrates a method that we use frequently: approximation by springs.

4.3.1 Speed of sound

In a gas at pressure P , the speed of sound is roughly

$$c_s \sim \sqrt{P/\rho}, \quad (4.33)$$

where P is the pressure and ρ is the density. We pretend that the gas is an ideal gas, even though this pretense introduces a theoretical difficulty: Sound cannot travel in an ideal gas, because sound waves require molecular collisions to propagate, and in an ideal gas, molecules do not collide. A more detailed analysis shows that sound waves with frequency ν can propagate provided that each molecule has a collision frequency $f \gg \nu$. In air, which for most purposes can be treated as an ideal gas, the collision frequency is roughly 10^9 Hz. This frequency is high enough to include most sound waves that we are interested in. The slight deviation from ideal-gas behavior is sufficient to make sound possible, so we will not worry about this theoretical incompatibility between sound propagation and the ideal-gas approximation.

The sound-speed formula (4.33) is the only dimensionally correct combination of c_s , P , and ρ . For a solid or liquid, we expect a similar relation, but we need to replace pressure with an analogous quantity. The units of pressure are force per area, or energy per volume: Pressure is energy density. We already have a volume: the interatomic volume a^3 . We have an energy: the cohesive energy ϵ_c . An analogue of pressure could be

$$\mathcal{M} \sim \frac{\text{energy}}{\text{volume}} \sim \frac{\epsilon_c}{a^3}. \quad (4.34)$$

The speed of sound is then

$$c_s \sim \sqrt{\frac{\mathcal{M}}{\rho}} \sim \sqrt{\frac{\epsilon_c}{\rho a^3}} \sim \sqrt{\frac{\epsilon_c}{m}} \quad (4.35)$$

where $\rho a^3 \sim m$ is the mass of an atom.

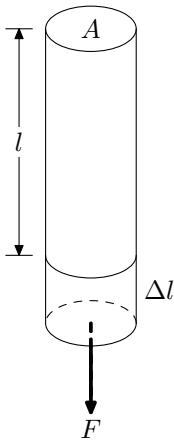


Figure 4.5. Wire stretched by a force. The force F stretches the wire by Δl from its relaxed length l ; it has cross-sectional area A .

The quantity \mathcal{M} is the **elastic modulus**. It is analogous to the spring constant, but it is more useful. Imagine a wire (Figure 4.5) stretched by a force. Its spring constant determines how far it stretches: $F = k_w \Delta l$. A thicker wire (of the same material)—with, say, area $4A$ —has a correspondingly larger spring constant: $k_1 = 4k_w$. A longer wire—with, say, length $2l$ (and the same area)—has a correspondingly smaller spring constant: $k_w/2$. The spring constant is not a property of the wire alone; it depends on the wire's length and area as well. The elastic modulus fixes this problem; it depends on only the substance.

How can we define such a quantity macroscopically, in terms of lengths, areas, and forces? (Equation (4.34) is a microscopic definition.) One dependence is on area: k/A is independent of area. The other is on length: $k_w l$ is independent of length. The combined quantity $k_w l/A$ is independent of area and length. In terms of this quantity, $F = k_w \Delta l$ becomes

$$F = \left(\frac{k_w l}{A} \right) \frac{A}{l} \Delta l, \quad (4.36)$$

or

$$\underbrace{\frac{F}{A}}_{\sigma} = \underbrace{\frac{k_w l}{A}}_{\mathcal{M}} \underbrace{\frac{\Delta l}{l}}_{\epsilon}, \quad (4.37)$$

where σ is stress (or pressure), \mathcal{M} is elastic modulus, and ϵ is fractional change in length, or **strain**. The strain is the dimensionless measure of extension. We show shortly that

$$\mathcal{M} = k_w l/A \quad (4.38)$$

is the macroscopic equivalent of (4.34).

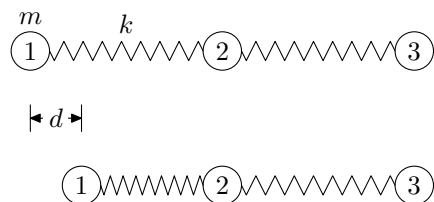


Figure 4.6. Small piece of a solid. The atoms, with mass m , are connected by bonds approximated as ideal springs of spring constant k . At $t = 0$, atom 1 gets a kick, and moves a distance d (second row). When does atom 2 follow in atom 1's footsteps?

We just imagined a wire as a spring. We can apply to the spring picture to the atomic scale as well. On the atomic scale, the bonds are springs. Imagine a line of atoms, each connected to its neighbors by complex forces: electrostatics and kinetic energy in delicate balance. We replace the complexity of a bond with the simplicity of a spring (Figure 4.6). We can use this picture to estimate the sound speed, and then compare the microscopic estimate with the dimensional guess (4.35). A sound wave transmits atomic motion from one location to another. To idealize the complex motion of the atoms, imagine that atom 1 gets a swift kick (an impulse) that moves it a distance d to the right. When does atom 2

hear the news, and move a distance d ? Before the kick, the atoms rested peacefully in their equilibrium positions. Right after atom 1 moves, the force on atom 2 is $F \sim kd$, where k is the spring constant, so atom 1's acceleration is $a = F/m \sim kd/m$. After time t , atom 2 has traveled a distance $at^2 \sim kt^2d/m$. This distance reaches d at $\tau \sim \sqrt{m/k}$. The sound speed is then $c_s \sim a/\tau \sim \sqrt{ka^2/m}$. To estimate the spring constant, k , we use $\epsilon_c \sim ka^2$. This approximation is equivalent to assuming that a bond breaks when its length is doubled (Figure 4.7). With this approximation, $c_s \sim \sqrt{\epsilon_c/m}$, as we derived in (4.35) by analogy with gases.

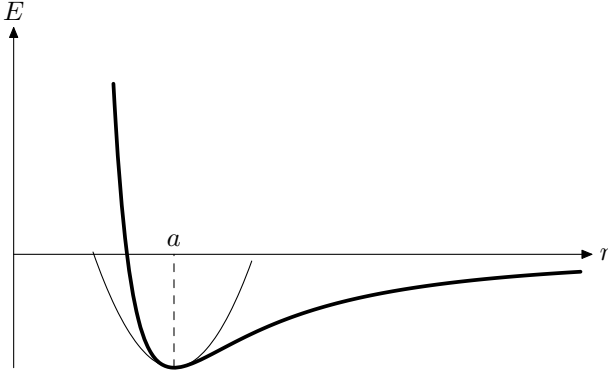


Figure 4.7. Bond energy versus separation. As two atoms approach, electron repulsion contributes a large positive energy. As they separate, electrostatic attraction weakens (becomes less negative). In between is a happy compromise, the energy minimum, a . Around that minimum, we fit a parabola: We assume that the bond is spring (we discuss the near-universal value of modeling physical processes with springs in Section 6.1).

We can count bond springs to justify the macroscopic expression $\mathcal{M} = k_w l/A$ from (4.38). The wire of Figure 4.5 is a bundle of filaments, where each filament is a line of atoms. Each filament has $N_l \sim l/a$ springs; the spring constant of a filament is therefore reduced by l/a compared to the spring constant of a bond. The wire contains $N_f \sim A/a^2$ filaments; this factor increases the spring constant of the wire compared to the spring constant of a filament. The spring constant of the wire is then $k_w = kN_f/N_l \sim kA/la$. With this expression, the macroscopic definition of elastic modulus (4.38) becomes $\mathcal{M} \sim k/a$. We need to estimate k , for which we use $ka^2 \sim \epsilon_c$. Then we have $\mathcal{M} \sim \epsilon_c/a^3$, which is the microscopic estimate (4.34).

To estimate a typical sound speed, we first evaluate a typical elastic modulus using (4.34):

$$\begin{aligned} \mathcal{M} &\sim \frac{\epsilon_c}{a^3} \sim \frac{2 \text{ eV}}{(3 \text{ \AA})^3} \sim \frac{2 \times 1.6 \cdot 10^{-12} \text{ erg}}{3 \cdot 10^{-23} \text{ cm}^3} \\ &\sim 10^{11} \text{ erg cm}^{-3}. \end{aligned} \tag{4.39}$$

This estimate for \mathcal{M} is reasonable for many materials (Table 4.4). A “typical” solid has

<i>Substance</i>	\mathcal{M} ($10^{11} \text{ erg cm}^{-3}$)
Steel	20
Cu	11
Ice (-5°C)	0.9
Al	7.9
Pb	1.8
C (diamond)	44
C (graphite)	
\perp to planes	3.5
\parallel to planes	100
Ash (white)	12
Glass	6 ± 0.5

Table 4.4. Elastic moduli. There are numerous elastic moduli. We list for each substance the Young's modulus, which is a combination of the shear modulus and the bulk modulus. Note how strong \mathcal{M} for graphite depends on direction, because result of graphite's layered structure. If \mathcal{M}_\perp were as high as \mathcal{M}_\parallel , then lead pencils would be diamond styluses. Source: Smithsonian Physical Tables [16].

atomic mass, say, 40; from (4.25), it has density $\rho \sim 3 \text{ g cm}^{-3}$, so

$$c_s \sim \sqrt{\frac{\mathcal{M}}{\rho}} \sim \sqrt{\frac{10^{11} \text{ erg cm}^{-3}}{3 \text{ g cm}^{-3}}} \sim 1.7 \text{ km s}^{-1}. \quad (4.40)$$

This estimate is reasonably accurate (Table 4.5), indicating that we have included the important physics.

4.3.2 Yield strength

How strong are materials? To break a perfect material (for example, diamond or carbon filament with no flaws), we would have to apply a stress $\sim \mathcal{M}$. Most materials break long before the stress reaches \mathcal{M} , because flaws in their crystal structure concentrate stress, so locally the stress may reach \mathcal{M} even if the global stress is much less than \mathcal{M} . A typical breaking stress (or yield stress) is between $0.001\mathcal{M}$ and $0.01\mathcal{M}$. The dimensionless factor is the yield or breaking strain, ϵ_y , which we list in Table 4.6. A typical yield stress is 10^9 erg cm^{-3} . We now apply these estimates to mountain climbing.

4.3.2.1 MOUNTAIN HEIGHTS ON THE EARTH. How much energy does it take to climb the tallest mountain on a planet? We first determine the height of such a mountain (Figure 4.8). The strength of the rock limits its height. The mass of the top conical block is ρh^3 , where ρ is the density of rock; it produces a force $F \sim \rho g h^3$. The resulting stress is

$$\sigma \sim \frac{F}{A} \sim \frac{\rho g h^3}{h^2} = \rho g h. \quad (4.41)$$

<i>Substance</i>	<i>v</i> (km s ⁻¹)
Glass	5.5
Steel	5.0
Brick	3.7
Pine	3.3
Water	1.5
Hg	1.5
Cu	3.6
Cork	0.5
Granite	4.0

Table 4.5. Speed of sound at room temperature. Our generic estimate is that $v \sim 1.7 \text{ km s}^{-1}$. It is a useful rule of thumb; but the rule of thumb underestimates the speed in metals (and stiff substances such as glass), and overestimates the speed in soft substances such as cork. Source: Smithsonian Physical Tables [16, p. 306].

<i>Substance</i>	ϵ_y
Steel	0.005
Cu	0.002
Al	0.001
Rock	0.001

Table 4.6. Approximate breaking strains, ϵ_y .

The strain is

$$\epsilon = \frac{\sigma}{\mathcal{M}} \sim \frac{\rho g h}{\mathcal{M}}. \quad (4.42)$$

For rock, $\mathcal{M} \sim 10^{12} \text{ erg cm}^{-3}$. When $\epsilon \sim \epsilon_y$, the rock yields, and the mountain shrinks until ϵ falls below ϵ_y . Therefore, the maximum height for a mountain on the earth is

$$h_{\max} \sim \frac{\mathcal{M} \epsilon_y}{\rho g}. \quad (4.43)$$

For a typical rock, $\epsilon_y \sim 10^{-3}$ and $\rho \sim 3 \text{ g cm}^{-3}$, so

$$h_{\max} \sim \frac{10^{12} \text{ erg cm}^{-3} \times 10^{-3}}{3 \text{ g cm}^{-3} \times 1000 \text{ cm s}^{-2}} \sim 3 \text{ km}. \quad (4.44)$$

In this estimate, we neglected many numerical factors; not surprisingly, some mountains, such as Mount Everest (where $h \sim 10 \text{ km}$), are larger than our estimated h_{\max} . (Perhaps Everest is made of extra-hard rock, such as granite.)

4.3.2.2 MOUNTAIN HEIGHTS ON OTHER PLANETS. Is it more difficult to climb the highest mountains on Mars, or Mercury, or the asteroid Ceres, compared to the highest mountains

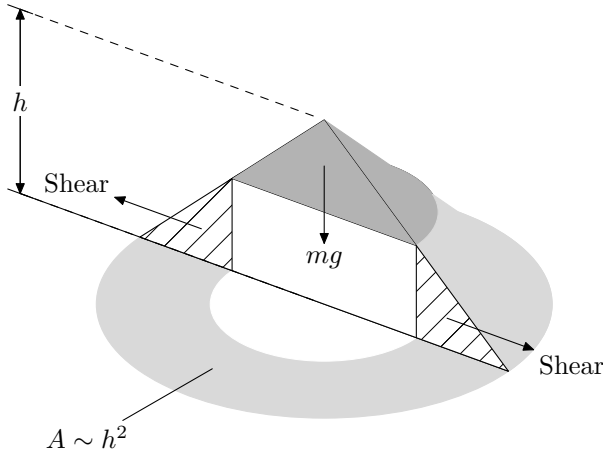


Figure 4.8. *Mountain.* This approximate mountain is a cone with 45° slopes. The shaded top block with mass m squeezes the hatched chunk (whose footprint is the lightly shaded ring). When the mountain gets too high, the squeeze becomes greater than the mountain rock can withstand.

on the earth? We assume that all astronomical objects are made of the same type of rock. The energy required to climb to the top is $E \sim mgh_{\max} \sim m\mathcal{M}\epsilon_y/\rho$, where m is the mass of the climber. For a typical climber, $m \sim 50$ kg, so

$$\begin{aligned}
 E &\sim \frac{5 \cdot 10^4 \text{ g} \times 10^{12} \text{ erg cm}^{-3} \times 10^{-3}}{3 \text{ g cm}^{-3}} \\
 &\sim 1.7 \cdot 10^{13} \text{ erg} \\
 &\sim 400 \text{ kcal},
 \end{aligned} \tag{4.45}$$

or one chocolate bar (at perfect efficiency). This energy is independent of g . People do not climb with perfect efficiency; no one could climb Everest on 1 or even 10 chocolate bars.

How high are these mountains? The energy to climb the mountains does not depend on g , but the maximum height does. As we found in (2.2.1), $g \propto R$. So $h_{\max} \propto R^{-1}$, and the fractional size of mountains, h_{\max}/R is $\propto R^{-2}$. For the earth, $h_{\max} \sim 3$ km and $R_{\oplus} \sim 6000$ km, so $h/R_{\oplus} \sim 10^{-3}$. For an asteroid of radius $R \sim R_{\oplus}/30$ (roughly 200 km), the fractional height is roughly 1. This estimate is reasonably accurate. Asteroids with $R \sim 200$ km have mountains roughly their own size; these mountains make the asteroids look oblate. In Section 4.4, we estimate mountain heights on white dwarfs.

4.3.3 Surface tension

Water droplets are spherical, because spheres have the smallest surface area for a given volume. This mathematical explanation hides a physical question: Why do water droplets minimize their surface area? The reason is that water has **surface tension**. Surface tension is the energy required to create one unit of new surface (surface tension has units of energy per area). On our picture of a solid, every atom in the interior is bonded to six neighbors. Each atom at the surface has only five neighbors. In a surface of area A , there are $N = A/a^2$ atoms. To make such a surface, we must break N bonds, one for each atom. The cohesive

energy is the energy to break all six bonds, so $E_{\text{bond}} \sim \epsilon_c/3$. To avoid counting each bond twice, we divide by 3 instead of by 6. So the surface energy is $E \sim AE_{\text{bond}}/a^2 \sim A\epsilon_c/3a^2$. The surface tension is the ratio E/A :

$$\gamma \equiv \frac{E}{A} \sim \frac{\epsilon_c}{3a^2}. \quad (4.46)$$

We can estimate cohesive energies from heats of vaporization, as we discuss in Section 5.2. For water, $\epsilon_c \sim 0.5 \text{ eV}$. Using our standard spacing $a \sim 3 \text{ \AA}$, we get

$$\begin{aligned} \gamma_{\text{water}} &\sim \frac{0.5 \text{ eV} \times 1.6 \cdot 10^{-12} \text{ erg eV}^{-1}}{3 \times (3 \cdot 10^{-8} \text{ cm})^2} \\ &\sim 300 \text{ erg cm}^{-2}. \end{aligned} \quad (4.47)$$

Although our basic model is sound, this estimate is large by a factor of 4, partly because of the crudity of our six-neighbor picture of solids.

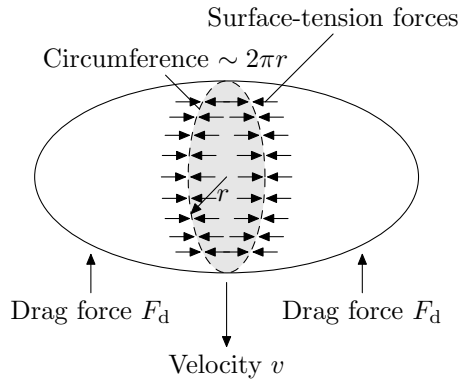


Figure 4.9. Falling raindrop of density ρ_w and radius r . The drop has reached terminal velocity v . Surface tension (the facing arrows) binds the two halves of the drop along the circumference. The shaded area shows a cross section of the drop.

Surface tension determines the size of raindrops. Falling raindrops feel a drag force, which tries to split the drop into smaller droplets. Many small raindrops cost extra surface energy, which fights the splitting. There is a happy, intermediate size, which we estimate by balancing the surface tension (as a force, F_γ) and the drag force, F_{drag} (Figure 4.9).

We first estimate the surface-tension force. Surface tension is energy per area, which is also force per length. To obtain a force, we multiply γ by the only length involved: the circumference, which is roughly $2\pi r$. So $F_\gamma \sim 2\pi r\gamma$. The drop moves at constant velocity (terminal velocity), so we equate the drag force and the weight: $F_{\text{drag}} = \rho r^3 g$, where ρr^3 is the mass of the drop. Equating the drag and surface-tension forces, we get $\rho r^3 g \sim 2\pi r\gamma$, or

$$r_{\text{max}} \simeq \sqrt{\frac{\gamma}{\rho_w g}}. \quad (4.48)$$

For water, $\gamma \sim 70 \text{ dyne cm}^{-1}$, so

$$r_{\max} \sim \sqrt{\frac{70 \text{ dyne cm}^{-1}}{1 \text{ g cm}^{-3} \times 10^3 \text{ cm s}^{-2}}} \sim 0.25 \text{ cm.} \quad (4.49)$$

The terminal velocity is given by the high-speed result (3.49):

$$v \sim \sqrt{gr_{\max} \frac{\rho_{\text{water}}}{\rho_{\text{air}}}}. \quad (4.50)$$

Because $\rho_{\text{water}}/\rho_{\text{air}} \sim 1000$, we get $v \sim 10 \text{ m/s}$, which is the speed of slow driving, or fast sprinting. This velocity seems about right; if you drive slowly in a rainstorm, the drops strike at roughly a 45° angle.

4.4 Application to white dwarfs

White dwarfs are small, dense stars in which the pressure is high enough to completely ionize the atoms. They are dead stars; there is no nuclear fusion in their interiors. We apply our estimation methods to this bizarre state of matter, to show you that the principles have wide use.

4.4.1 Size of a white dwarf

How large is a white dwarf? In a white dwarf, the atoms live in a sea of electrons. What prevents gravity from squeezing the white dwarf to zero size? You could ask a similar question about hydrogen: What prevents the electrostatic force from squeezing the radius to zero? In hydrogen, and in white dwarfs, confinement energy prevents collapse. For a white dwarf, we can find the minimum-energy radius by balancing gravitational and confinement energy. [The concept of confinement energy is a simple version of a more complicated argument that uses the Pauli exclusion principle; see Section 5.6.2.]

We first work out the confinement energy. Imagine a star of radius R and mass M , composed of atoms with atomic number Z and atomic mass A . We define the dimensionless ratio $\beta \equiv Z/A$, because this factor shows up often in the derivation. The star contains M/Am_p atoms, and $N_e = (M/m_p)(Z/A) = \beta M/m_p$ electrons. The number density of electrons is n_e , so each electron is confined to a region of size $\Delta x \sim n_e^{-1/3}$. (We shamelessly neglect numerical factors such as $4\pi/3$.) The electron's momentum is $\Delta p \sim \hbar n_e^{1/3}$, and its confinement energy is $(\Delta p)^2/m_e \sim \hbar^2 n_e^{2/3}/m_e$. This result is valid for nonrelativistic electrons. The total confinement (or kinetic) energy is this value times N_e :

$$\text{KE} \sim \frac{\hbar^2 n_e^{2/3}}{m_e} N_e \sim \frac{\hbar^2 N_e^{5/3}}{m_e R^2}, \quad (4.51)$$

because $n_e^{2/3} = N_e^{2/3}/R^2$. In terms of M and m_p , this energy is

$$\text{KE} \sim \frac{\hbar^2 M^{5/3}}{m_e m_p^{5/3} R^2} \beta^{5/3}. \quad (4.52)$$

We next work out the gravitational (potential) energy, which is the sum of all pairwise interactions between all particles in the star (protons, electrons, and neutrons). What a

mess! We first try a dimensional argument. We construct an energy using the variables m_e , m_p , M , and R . Any expression of the form

$$\text{PE} \sim \frac{G \times \text{mass} \times \text{mass}}{R}, \quad (4.53)$$

where G is the gravitational constant, is an energy. Unfortunately, there are three masses from which to choose, and two slots in which to put them. The possibilities are numerous. Fortunately, we can restrict the freedom. Gravity does not care about the composition of the star; muons, quarks, electrons, and protons all obey universal gravitation. So m_e and m_p are irrelevant, and

$$\text{PE} \sim \frac{GM^2}{R}. \quad (4.54)$$

A physical argument also gives us this result. Two random points in the white dwarf have a separation of $\sim R$. Imagine one-half of the star clumped at one point, and one-half clumped at the other point. The gravitational potential energy between the two clumps is $\sim -G(M/2)^2/R$. Once we neglect the numerical factors, we get $\text{PE} \sim -GM^2/R$, in agreement with (4.54), but with the correct sign.

Now we equate PE and KE (in absolute value) to solve for R :

$$\frac{GM^2}{R} \sim \frac{\hbar^2 \beta^{5/3} M^{5/3}}{m_e m_p^{5/3} R^2}, \quad (4.55)$$

so

$$R \sim M^{-1/3} \frac{\hbar^2 \beta^{5/3}}{G m_e m_p^{5/3}}. \quad (4.56)$$

This expression contains two factors of \hbar , which suggests that we multiply it by c^2/c^2 to manufacture two factors of $\hbar c$. After regrouping, we find that the radius is

$$R \sim \frac{\hbar c}{m_e c^2} \frac{\hbar c}{G m_p^2} \left(\frac{m_p}{M} \right)^{1/3} \beta^{5/3}. \quad (4.57)$$

This expression contains factors familiar from our computation of the Bohr radius. The factor $\hbar c/m_e c^2$ is also in (4.22). This parallel suggests that we compare the second factor, $\hbar c/G m_p^2$, to the second in (4.22), which is $\hbar c/e^2$ (the reciprocal of the fine-structure constant). We therefore call $G m_p^2/\hbar c$ the gravitational fine-structure constant. Its value is $6 \cdot 10^{-39}$: Gravity is much weaker than electrostatics. The remaining factors in (4.57), which are dimensionless, have no analogue in the hydrogen calculation.

If our sun were a white dwarf, how large would it be? We use (4.57) and scale the result against M_\odot . For most atoms, $\beta = Z/A \sim 0.5$. The radius is

$$R \sim \frac{\hbar c}{m_e c^2} \frac{\hbar c}{G m_p^2} \left(\frac{m_p}{M_\odot} \right)^{1/3} \beta^{5/3} \left(\frac{M}{M_\odot} \right)^{-1/3}. \quad (4.58)$$

The numerical factors combine to give

$$R \sim 2 \cdot 10^3 \text{ km} \left(\frac{M}{M_\odot} \right)^{-1/3}. \quad (4.59)$$

When $M = M_\odot$, R is roughly 0.3 earth radii. A more exact calculation—honest physics—produces a factor of 4.5 in front: The actual radius is roughly one earth radii. White dwarfs are dense.

Our computation of the mass–radius relation followed the same procedure as did the computation of the Bohr radius. Both systems have a potential energy source: gravity in the star, electrostatics in hydrogen. Without opposition, the potential energy would be a minimum at zero size ($R = 0$ or $a_0 = 0$). In both systems, confinement energy fights collapse.

Note that the radius of a white dwarf is $\propto M^{-1/3}$, so the volume is *inversely* proportional to the mass. For fixed Z and A , the electron density is therefore $\propto M^2$. At a certain mass, n_e becomes large enough that the electrons become relativistic (this derivation assumed that the electrons are nonrelativistic). We can estimate this critical mass equating the kinetic energy per electron to $m_e c^2$, and solving for the white-dwarf mass. The kinetic energy per electron is KE/N_e , where KE is given in (4.51):

$$E_e \sim \frac{\hbar^2 N_e^{2/3}}{m_e R^2}. \quad (4.60)$$

We equate this energy to $m_e c^2$:

$$\frac{\hbar^2 N_e^{2/3}}{m_e R^2} \sim m_e c^2. \quad (4.61)$$

We now use the expression for R in (4.57), and the expression $N_e = \beta M/m_p$ for the number of electrons. With these substitutions, (4.61) becomes

$$\frac{\hbar^2 \beta^{2/3} M^{4/3}}{m_p^{2/3} m_e} \sim m_e c^2 \underbrace{\frac{M^{-2/3} \hbar^2 \beta^{10/3}}{G^2 m_e^2 m_p^{10/3}}}_{R^2}. \quad (4.62)$$

After canceling common factors and rearranging, we find that the critical mass is

$$M_{\text{critical}} \sim \beta^2 \left(\frac{\hbar c}{G m_p^2} \right)^{3/2} m_p. \quad (4.63)$$

With an honest-physics factor of 3.15 in front, this mass is the Chandrasekhar mass, which he derived in 1931; it is the largest mass that a white dwarf can have, and its value is roughly $1.4M_\odot$. We did not try to calculate this mass—we tried to calculate only the transition mass between nonrelativistic and relativistic regimes—but we found it anyway. The reason that the transition mass and the maximum mass are related is that a relativistic white dwarf is less stiff than a nonrelativistic one (it does not withstand compression as well). Imagine a nonrelativistic white dwarf to which we slowly add mass. As the mass increases, the radius decreases ($R \propto M^{-1}$), and so does the electron velocity. As the electron velocity approaches c , however, the less-stiff relativistic behavior begins to dominate, and the radius decreases faster than M^{-1} . The electron velocity then increases even more than we would expect from the nonrelativistic relation, and the less-stiff relativistic behavior dominates even more. You can see positive feedback in this description. At a large enough mass, this feedback makes the radius go to zero. The mass at which this collapse occurs is the Chandrasekhar mass. (The radius does not actually go to zero—the white dwarf becomes instead a much denser object, a neutron star.) Because the same physics determines the Chandrasekhar mass and the transition mass, our estimate is also an estimate of the Chandrasekhar mass.

4.4.2 Mountains on white dwarfs

How high are the tallest mountains on white dwarfs? The energy density is the elastic modulus: $\mathcal{M} \sim \text{PE}/R^3 \sim GM^2/R^4$. The mass density is $\rho \sim M/R^3$. And the acceleration due to gravity is $g \sim GM/R^2$. The height estimate (4.43) becomes

$$h_{\text{max}} \sim \frac{\epsilon_y \times GM^2/R^4}{(M/R^3) \times (GM/R^2)} \sim \epsilon_y R. \quad (4.64)$$

The morass of proton and electron masses, β , and so on, have canceled. All that remains is the yield strain and the white-dwarf radius. Perhaps white-dwarf matter has $\epsilon_y \sim 0.01$ (because the high pressures squash imperfections in the lattice). We use (4.59) for R . The mountain height is then

$$h_{\text{max}} \sim 20 \text{ km} \left(\frac{M}{M_\odot} \right)^{-1/3}. \quad (4.65)$$

If $M = M_\odot$, the mountain is ~ 20 km tall. It would be a many-chocolate-bar climb.

4.5 What you have learned

- *Quantum mechanics:* We can understand many of the consequences of quantum mechanics by simply adding \hbar to the list of relevant variables for dimensional analysis. Quantum mechanics, through \hbar , introduces a new momentum scale, the uncertainty momentum Δp , and indirectly, the uncertainty energy. This energy is also called the confinement energy.
- *Balancing:* Many physical systems contain one process that competes with another. For example, one energy (perhaps gravity) competes with another energy (perhaps the uncertainty energy). The energies usually are equal near the minimum-energy state. Look for these competitions!
- *Atomic sizes:* Atoms have a diameter of a few Å.
- *Binding energies:* Typical covalent and ionic bond energies are a few eV, as a consequence of electrostatic attraction.

5 MATERIALS II

How fast does a turkey cook? How quickly do the ground and the air cool on a cloudless night? How high a fever do you run after 30 minutes in a hot tub? At what temperature does water boil? We can answer such questions by estimating the thermal properties of materials. When we estimate boiling temperatures, we introduce an extended example of successive approximation—the main technique of order-of-magnitude physics. As a warmup, we begin with thermal expansion.

5.1 Thermal expansion

Most materials expand when they get hotter, and contract when they get cooler. If bridge designers forget this fact, and design a bridge so that its joints exactly mate in the summer, then it will fall down in the winter. We would like to understand why materials expand, and by how much.

Let's look at the potential of an atomic bond, and see what it implies about thermal expansion. In Section 4.3.1, we approximated a bond as a spring. We can use this approximation to estimate a typical **thermal-expansion coefficient**—the fractional length change per unit change in temperature. At a temperature T , the bond spring vibrates with average energy kT (actually, $kT/2$, but we ignore the 2). How far does that energy stretch the bond? The bond energy is, in this approximation,

$$E(r) \sim -E_0 + \frac{1}{2}k_s(r - a)^2, \quad (5.1)$$

where E_0 is the bond energy, k_s is the spring constant, a is the equilibrium bond length, and r is the bond length. We can factor out E_0 :

$$\frac{E(r)}{E_0} \sim -1 + \frac{k_s(r - a)^2/2}{E_0}. \quad (5.2)$$

We can write $E_0 = fk_s a^2/2$, where f is a dimensionless constant of order unity. Then

$$\frac{E(r)}{E_0} \sim -1 + \left(\frac{r}{a} - 1\right)^2 f. \quad (5.3)$$

To simplify the algebra, let's do the calculation in a dimensionless unit system. We measure energy in units of E_0 and distance in units of a . Then

$$E(r) \sim -1 + (r - 1)^2 f. \quad (5.4)$$

For now, let's simplify by assuming that $f = 1$. Then

$$E(r) \sim -1 + (r - 1)^2. \quad (5.5)$$

If we define $x \equiv r - 1$, so x is the change in length (which is what we are interested in), then the change in energy is $\Delta E(x) \sim x^2$. To find x , we equate $\Delta E(x)$ to T (we use also units in which $k \equiv 1$, so temperature is energy). We find that $x^2 \sim T$, or $x \sim \sqrt{T}$. [In normal units, this distance is $x_{\text{normal}} \sim a\sqrt{kT/E_0}$, but that value is irrelevant for the coming surprise.] How far does the bond shrink? It shrinks exactly the same distance as it expands, because a spring potential is symmetric. (Even if we had retained f , this conclusion would still hold.) So the average separation in a spring potential is independent of temperature: The thermal-expansion coefficient is zero!

We know that this result is not correct. We made too drastic an approximation when we replaced the bond with a spring. An actual interatomic potential is asymmetric, as shown in the cartoon potential-energy curve of Figure 5.1. The asymmetry produces a nonzero thermal-expansion coefficient.

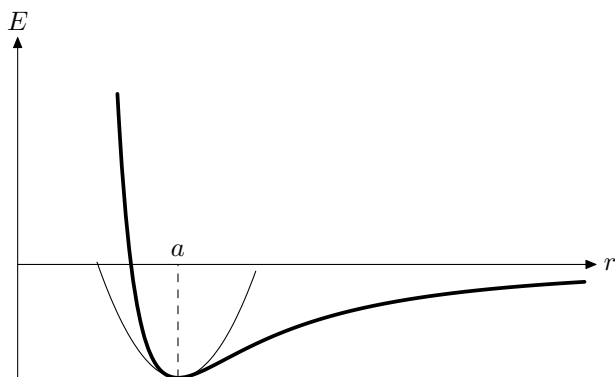


Figure 5.1. Potential energy versus separation. The potential energy, E , is plotted on the vertical axis (in arbitrary units); the equilibrium spacing, a , is marked on the horizontal axis. At higher energies (higher temperatures), the curve skews toward positive r , making the material expand as temperature rises. The parabola, although an adequate approximation to $E(r)$ for small deviations from a , gives a totally incorrect thermal-expansion coefficient: zero.

A more accurate approximation to E includes terms up to x^3 :

$$\Delta E(x) \sim x^2 + \beta x^3, \quad (5.6)$$

where β is a constant that measures the asymmetry of the potential. (We do not have to include the x^4 contribution, because it is symmetric, and therefore contributes no average length change). With the more refined bond energy (5.6), how much does the bond stretch or shrink? The change in energy is

$$\Delta E(x) \sim x^2 + \beta x^3. \quad (5.7)$$

We equate this energy to T :

$$x^2 + \beta x^3 = T. \quad (5.8)$$

If x is small, we can solve this equation by successive approximation. Can we assume that x is small? Although the analysis using a spring potential does not explain thermal expansion, it does give us the estimate that $x \sim \sqrt{T}$ (in dimensionless units). A typical T is 0.025 eV; a typical covalent bond energy is $E_0 \sim 2.5$ eV, so in dimensionless units, $T \sim 0.01$ and $x \sim 0.1$. For order-of-magnitude calculations, $x \sim 0.1$ is small, and we can solve (5.8) by successive approximations. For the first approximation, we ignore the βx^3 term, and recover the old result: $x_1 \sim \pm\sqrt{T}$; the plus sign represents bond stretching, the minus sign represents bond shrinking. For the next approximation, we replace βx^3 by βx_1^3 :

$$x^2 \pm \beta T^{3/2} = T. \quad (5.9)$$

The solution is

$$x = \pm\sqrt{T \mp \beta T^{3/2}}. \quad (5.10)$$

(We have to carefully distinguish \pm from \mp , by carefully tracking bond stretching and bond shrinking. For expansion, take the top sign; for contraction, take the bottom sign. Remember that β will be negative for most potentials, such as for the potential in Figure 5.1.) The average x is (if we neglect a factor of 2)

$$x_{\text{avg}} = \sqrt{T - \beta T^{3/2}} - \sqrt{T + \beta T^{3/2}}. \quad (5.11)$$

We factor out \sqrt{T} and use the approximation that $\sqrt{1+h} \sim 1 + h/2$:

$$\begin{aligned} x_{\text{avg}} &= \sqrt{T} \left(\sqrt{1 - \beta T^{1/2}} - \sqrt{1 + \beta T^{1/2}} \right) \\ &\approx \sqrt{T} \left((1 - \beta T^{1/2}/2) - (1 + \beta T^{1/2}/2) \right) \\ &= -\beta T \end{aligned} \quad (5.12)$$

The thermal-expansion coefficient is dx_{avg}/dT , which is $-\beta$. What is the value in a normal unit system? Let's restore normal units to $x_{\text{avg}} = -\beta T$. The quantity x_{avg} is really x_{avg}/a . The temperature T is really the energy kT , and the energy kT is really kT/E_0 . So (5.12) becomes $x_{\text{avg}}/a \sim -\beta kT/E_0$. If we had carried the f through, then we would find

$$\frac{x_{\text{avg}}}{a} \sim -\frac{\beta}{f} \frac{kT}{E_0}. \quad (5.13)$$

The thermal-expansion coefficient is the fractional change in length per change in temperature:

$$\alpha \equiv \frac{d(x_{\text{avg}}/a)}{dT} \sim -\frac{\beta}{f} \frac{k}{E_0}. \quad (5.14)$$

The Boltzmann constant is roughly $k \sim 10^{-4}$ eV K⁻¹; a typical bond energy is $E_0 \sim 3$ eV. For many interatomic potentials, $\beta \sim 1$. For example, for the potential $E(r) = 4r^{-4} - 4r^{-2}$ (in dimensionless units), $\beta \sim 2$. With $\beta \sim 1$, a typical thermal expansion coefficient is $3 \cdot 10^{-5}$ K⁻¹.

Table 5.1 contains thermal-expansion data for various materials. Our estimate seems to be an overestimate, although we can partially explain the trends. For example, stiff materials such as quartz and diamond have a particularly low α . Stiff materials are stiff, because their bonds are strong, so α , which is inversely proportional to bond strength, should be low. (Quartz has silicon-oxygen bonds, and diamond has carbon-carbon bonds; both bonds are strong.) Perhaps quartz and diamond also have particularly symmetric potentials (low β), which would also lower the expansion coefficient.

<i>Substance</i>	$\alpha (10^{-5} \text{ K}^{-1})$
Cu	1.6
Si	0.5
Quartz	0.05
Diamond	0.1
Wood	
along gr.	0.5
against gr	5
Glass	0.9
Pyrex	0.3
Vycor	0.08

Table 5.1. Thermal-expansion coefficients (at room temperature).

5.2 Phase changes

<i>Substance</i>	ϵ_c (eV)	$T_{\text{vap}} (K)$		
		<i>actual</i>	<i>A</i>	<i>B</i>
Water	0.50	373	577	533
NH ₃	0.31	240	374	347
HCl	0.21	188	263	244
O ₂	0.071	90	99	93
Au	3.35	3130	3293	3022
Xe	0.13	165	171	159
He	0.00086	4.2	2.2	2.5
Hg	0.61	630	691	638
N ₂	0.058	77	83	78

Table 5.2. Cohesive energy, ϵ_c , per atom or molecule; actual and predicted boiling temperatures (at 1 atm). We estimate the cohesive energy as L_{vap}/N_A , where L_{vap} —the heat of vaporization—is taken from experimental data. We use the cohesive energy to predict the boiling temperature. Column A is the prediction from (5.24). Column B is the prediction from (5.27). Source: [38, 6-103–6-106]

5.2.1 Boiling

To vaporize, or boil, a liquid, we must supply energy to move molecules from the liquid into the gas phase. How much energy is required? Per molecule, the energy required is roughly the cohesive energy, ϵ_c . The molar heat of vaporization (or enthalpy of vaporization) is the energy required to boil one mole of a substance, and we can estimate it as

$$L_{\text{vap}} \sim \epsilon_c N_A \sim 23 \text{ kcal mole}^{-1} \left(\frac{\epsilon_c}{1 \text{ eV}} \right). \quad (5.15)$$

For water, for example, $\epsilon_c \sim 0.5 \text{ eV}$, so $L_{\text{vap}} \sim 10 \text{ kcal mole}^{-1}$. This estimate of the enthalpy leaves out a small contribution: the energy to make room in the vapor for the evaporating molecules. This energy is PV in the expression $H = E + PV$ for the enthalpy. The heat of vaporization is the change in H :

$$L_{\text{vap}} \equiv \Delta H = \Delta E + \Delta(PV). \quad (5.16)$$

As we see shortly, after we estimate the boiling temperature, $\Delta(PV)$ is small compared to ΔE .

Predicting L_{vap} from ϵ_c is straightforward. Let's instead use L_{vap} —for which accurate experimental data are available—to determine ϵ_c (the second column of Table 5.2). Now we have a microscopic quantity, for which we can make order-of-magnitude pictures. What can we do with it? An essential feature of boiling is the boiling temperature. So let's try to predict the boiling temperature using the cohesive energy. We first define boiling temperature. As a substance heats up, more molecules leave the liquid and become vapor; the pressure of the vapor increases. The boiling temperature is the temperature at which the vapor pressure reaches atmospheric pressure (at sea level, roughly $10^6 \text{ dyne cm}^{-2}$, or 1 atm). Atmospheric pressure is arbitrary; it is less on Mount Everest, for example, and far less on Mars. It is unrelated to the properties of the substance being boiled—although it helps to determine the boiling temperature. A more general question is how boiling temperature depends on atmospheric pressure. You can generalize the methods that we introduce to answer this question.

5.2.1.1 SIMPLEST MODEL. For our first guess for the boiling temperature, we convert the cohesive energy into a temperature using Boltzmann's constant: $T_{\text{vap}} = \Pi \epsilon_c / k$, where Π is a dimensionless constant. The conversion factor k is $1 \text{ eV} \simeq 10^4 \text{ K}$ (accurate to 20 percent). The data in Table 5.2 shows how inaccurate this guess is. For example, for water, the predicted boiling temperature would be 5000 K, instead of 373 K: Even on Mercury, oceans would freeze solid. Or, for gold, the predicted temperature would be 30,000 K instead of $\sim 3000 \text{ K}$. If we by fiat insert a factor of 10, and assert that

$$T_{\text{vap}} \sim \frac{\epsilon_c}{10k}, \quad (5.17)$$

then our prediction would be reasonably accurate for both gold and water. If T_{vap} , ϵ_c , and k are the only relevant variables in this situation, then there is some constant Π such that $T_{\text{vap}} = \Pi \epsilon_c / k$. A few minutes with the data in Table 5.2 will convince you that there is no such constant, although (5.17) provides a reasonable approximation. Using this approximation, we can check that we were justified in neglecting the PV term in the heat of vaporization (5.16). From the ideal gas law, for one molecule in a gas at the boiling temperature, $PV = kT_{\text{vap}}$. Using (5.17), we find that PV is $\epsilon_c/10$. For order-of-magnitude purposes, PV is small compared to ϵ_c .

5.2.1.2 VOLUME CONTRIBUTION TO THE ENTROPY: PART 1. Now let's explain the approximate factor of 10. The factor arises because we left out **entropy**, which is a fancy word for counting; the meaning of entropy will become clear as we calculate it. At the boiling point, vapor is in equilibrium with liquid: A molecule is equally happy in either phase. In

thermodynamic language, the Gibbs' free energy is the same in the two phases. Instead of analyzing the problem using thermodynamic variables, which obscure the simplicity of the mechanism, we analyze the feelings of one molecule deciding where to go (the statistical-mechanics approach). If we consider only energy, the molecule is more likely to be in the liquid than in the vapor, by the Boltzmann factor $e^{\epsilon_c/kT_{\text{vap}}}$. So we have deduced that the liquid does not boil at any temperature. However, we must also include the entropy; that is, we must count the number of accessible states in the vapor and in the liquid. We could, using quantum mechanics, compute the number of accessible states in the liquid; compute it in the solid; and then take the ratio. Instead, we estimate the ratio directly, without stopping at the waystation of quantum mechanics.

We assume that the vapor is an ideal gas, so that $PV = NkT$. At atmospheric pressure, and at the boiling temperature, one molecule is free to wander in a volume $V_{\text{gas}} = kT_{\text{vap}}/P_0$, where $P_0 = 1$ atm. In the liquid, the free volume for one molecule is a factor times a^3 , but the factor is difficult to determine, because liquids are poorly understood. One simple model of a liquid is a solid with some atoms replaced by holes. How many holes should there be? If no atom is surrounded by holes, then the substance is a solid, not a liquid. If each atom is surrounded by six holes (all its neighbors are holes), then the substance is a gas. A reasonable compromise is that, on average, one or two holes surround each atom. Each hole is shared among its six neighboring atoms (if none of them are replaced by holes), so the free volume per atom is perhaps $a^3/6$ or $a^3/3$. Let's be more general and say that the free volume is $V_{\text{liquid}} \sim (fa)^3$, where $0 < f < 1$.

The ratio of allowed volumes is

$$r_v \equiv \frac{V_{\text{gas}}}{V_{\text{liquid}}} \sim \frac{kT_{\text{vap}}}{P_0 a^3 f^3}. \quad (5.18)$$

This ratio is $\gg 1$, as we will find out in a moment; therefore this entropy factor encourages the molecule to move to the vapor. This factor is independent of temperature, whereas the Boltzmann factor depends strongly on temperature. The factors balance at the boiling temperature:

$$e^{\epsilon_c/kT_{\text{vap}}} = r_v, \quad (5.19)$$

or

$$\frac{\epsilon_c}{kT_{\text{vap}}} = \log r_v = \log \left(\frac{kT_{\text{vap}}}{P_0 a^3 f^3} \right). \quad (5.20)$$

[All logs are natural logs.] This equation is transcendental; we can solve it by successive approximation, or by drawing a graph.

<i>Var.</i>	<i>Name</i>	<i>Value</i>
E	internal energy	
H	enthalpy	$E + PV$
F	Helmholtz free energy	$E - TS$
G	Gibbs free energy	$H - TS$

Table 5.3. *Thermodynamic energies.*

Let's first connect the result (5.20) to the (perhaps more familiar) thermodynamic quantities of Table 5.3. We are neglecting the PV term, so the internal energy and the enthalpy are, in this approximation, identical, as are the Gibbs and the Helmholtz free energies. With T constant, which is what equilibrium means, $\Delta F = \Delta E - T\Delta S$. When the liquid and gas are in equilibrium, the free energy change going from liquid to gas is zero, so $\Delta E = T\Delta S$. Matching this equation against (5.20), we find that ϵ_c corresponds to E iff $\Delta S = k \log r_v$. Here we see the value of the free energy. It is the energy with an adjustment TS ; the adjustment includes the counting of states.

Before solving by successive approximations, we should get a qualitative feel for the sizes of the terms in (5.20), and see whether the temperatures that it predicts are reasonable. According to (5.20), the boiling temperature is given by

$$T_{\text{vap}} \sim \frac{\epsilon_c}{k \log r_v}. \quad (5.21)$$

So let's evaluate r_v for a couple substances using (5.18) and then compute $\log r_v$. For water (with $f = 1$), we get

$$r_v \sim \frac{1.4 \cdot 10^{-16} \text{ erg K}^{-1} \times 374 \text{ K}}{10^6 \text{ dyne cm}^{-2} \times (3 \cdot 10^{-8} \text{ cm})^3} \sim 200, \quad (5.22)$$

and $\log r_v \sim 5$. For gold, we get $\log r_v \sim 7.4$. So we can explain most of the factor of 10 in our first guess (5.17): It comes from entropy.

The first step in solving a transcendental equation with a (natural) logarithm is to assume that **any logarithm is 10**. This rule is due to Fermi, who was a master of order-of-magnitude physics. A log could also be -10 , although not in this case because $\epsilon_c/kT_{\text{vap}}$ is positive. At first glance, this rule of 10 seems ridiculous, but it has the following justification. Initially, most order-of-magnitude arguments assume power laws for the unknown functional forms; this assumption excludes logarithms as solutions. If there is a logarithm in the solution, it often arises in a later, more refined stage of the analysis. In what circumstances would we refine the analysis? Only when the original analysis was at least moderately accurate. If it is wildly inaccurate, then we would throw out the whole analysis and start again. If it is extremely accurate, then a refinement is not likely to introduce a new function, such as a logarithm. So, if a logarithm arises, we can be reasonably certain that the original estimate was only moderately accurate, and that the logarithm introduces a correction factor of at least 5, perhaps even as high as 20 or 30. We can further narrow this range. The logarithm takes a dimensionless argument. How large can a dimensionless argument become? For example, Avogadro's number, $N_A \sim 6 \cdot 10^{23}$, is huge, and $\log N_A \sim 55$. However, this example is not relevant, because Avogadro's number is not dimensionless. Its units are mole^{-1} , and the mole is an arbitrary unit whose size depends on the size of the gram. A dimensionless argument must be the ratio of two quantities with the same units; in this boiling-temperature problem, the two quantities are V_{liquid} and V_{gas} . In other problems, they might be two characteristic lengths or energies. The ratio of two characteristic quantities is almost never larger than $r_v = 10^8$, in which case $\log r_v \sim 20$. So the logarithm typically lies between 5 and 20; a reasonable starting estimate is 10. With this starting estimate, we recover (5.17).

Starting with $\log(\cdot) \sim 10$, let's predict T_{vap} for water. The argument of the logarithm is

$$\frac{kT_{\text{vap}}}{\epsilon_c} \frac{\epsilon_c}{P_0 a^3 f^3}. \quad (5.23)$$

To continue, we need a value for the fudge factor f . Should it be 0 (the substance is virtually a solid) or 1 (the substance is virtually a gas)? It's difficult to decide; both values are extreme, but we have no clear argument favoring one extreme or the other. In such situations of ignorance, we choose the middle value $f = 0.5$; this rule is the **rule of one-half**. We can estimate the spacing a from the density and molar mass, but for simplicity we use our usual estimate $a \sim 3 \text{ \AA}$. Then $P_0 a^3 f^3 \sim 3.3 \cdot 10^{-6} \text{ eV}$. For convenience, we write F for $\log r_v$; so $F = \epsilon_c / kT_{\text{vap}}$. In terms of F , the transcendental equation is

$$F = \log \left(F^{-1} \frac{\epsilon_c}{3.3 \cdot 10^{-6} \text{ eV}} \right). \quad (5.24)$$

For water, $\epsilon_c \sim 0.5 \text{ eV}$. We start with $F_0 = 10$. The next iteration produces

$$F_1 = \log (1.5 \cdot 10^4) \simeq 9.62. \quad (5.25)$$

Because $F_1 \simeq F_0$, we do not need to compute F_2 ; we might as well use $F = 10$. Then $T_{\text{vap}} \sim \epsilon_c / 10k \sim 576 \text{ K}$. On the Celsius scale, the error in this estimate is over 100 percent. However, the Celsius scale is irrelevant for physics. It has an arbitrary offset; why should the universe care about the boiling temperature of water? On the Kelvin scale, which has a natural definition for the zero point, the error is ~ 50 percent.

5.2.1.3 VOLUME CONTRIBUTION TO THE ENTROPY: PART 2. Can we reduce it? The most arbitrary input was the factor $f = 0.5$. If we can understand its origin, we can choose it more accurately. It is a dimensionless measure of the molecular vibration amplitude. Let's estimate this amplitude. The vibrations are thermal, with a typical energy at the boiling temperature of $kT_{\text{vap}}/2$. In the model for surface tension in Section 4.3.3, each molecule had six neighbors, and each bond had energy $E_{\text{bond}} \sim \epsilon_c/3$. When we estimated thermal expansion in Section 5.1, we made a spring model in which the vibration energy (the change in bond energy relative to the equilibrium bond energy) is $E_{\text{bond}}(x/a)^2$, where x is the vibration amplitude. Let's solve for x as a function of kT_{vap} . The bond vibration gets $kT_{\text{vap}}/2$ from thermal motion, so

$$\frac{kT_{\text{vap}}}{2} \sim \frac{\epsilon_c}{3} \left(\frac{x}{a} \right)^2. \quad (5.26)$$

The solution is $x \sim a \sqrt{1.5kT_{\text{vap}}/\epsilon_c}$. If we take x as an estimate for fa , then we find that $f \sim \sqrt{1.5kT_{\text{vap}}/\epsilon_c}$. The factor $kT_{\text{vap}}/\epsilon_c$ is familiar: It is F^{-1} , which is roughly 0.1. Therefore, $f \sim 0.4$. If we use this f in the successive-approximation procedure for T_{vap} , we get $T_{\text{vap}} \sim 540 \text{ K}$ for water. This new T_{vap} gives us a new estimate for F , and thus a new estimate for f , which changes T_{vap} . We are solving the equation

$$F = \log \left(F^{-1} \frac{\epsilon_c}{2.6 \cdot 10^{-5} \text{ eV} \times f^3} \right), \quad (5.27)$$

with f itself given in terms of F :

$$f = \sqrt{1.5F^{-1}}. \quad (5.28). (5.29)$$

We can continue with further iterations; the succeeding iterations converge to 533 K. This value is in error by 45 percent; our effort has not reduced the error significantly.

5.2.1.4 ROTATIONAL CONTRIBUTION TO THE ENTROPY. Before despairing, we should check our predictions on other, less peculiar substances; water has many strange properties. The prediction is accurate for many substances, as you can see from column B of Table 5.2. The exceptions are water, ammonia, and hydrochloric acid. They share a feature, hydrogen bonds, and a common discrepancy, a lower boiling temperature than we expected. What effect do hydrogen bonds have? Let's see what effect they should have to correct our predictions, and see whether we can find a physically justify the effect using hydrogen bonds. We predicted the boiling temperature by calculating, or estimating, the entropy change. If our prediction for T_{vap} is too high, then we must have underestimated the entropy increase from liquid to gas. What physical effect could add to the entropy increase? Because our prediction is accurate for many substances, with widely differing bond types, our volume estimate in the liquid is probably reliable, even for hydrogen-bonded substances. So hydrogen bonds must restrict the freedom in a way that we have not yet included. They can do so if they have a direction—if they are dipolar. And they are; hydrogen bonds are like weak covalent bonds, so they have a preferred direction. In a water molecule, the electron around hydrogen is dragged toward the oxygen, so the oxygen end of the molecule is negatively charged, and the hydrogen end is positively charged. Thus hydrogen ends tend to attract oxygen ends. The bond is not as strong as a covalent bond, because the electron is not dragged all the way to the oxygen.

In neglecting the rotational contribution to the entropy, we implicitly, and incorrectly, assumed that a molecule in the liquid could rotate as freely as a molecule in the gas; hydrogen bonds prevent this easy rotation in the liquid. We need to compute a new ratio,

$$r_{\text{rot}} \equiv \frac{\text{allowed solid angle of rotation in gas}}{\text{allowed solid angle of rotation in liquid}}, \quad (5.30)$$

and include $k \log r_{\text{rot}}$ in the entropy. The solid angle in the gas is 4π . To compute the liquid solid angle, we estimate the vibration angle of a water molecule, using the following crude model of hydrogen bonding. We concentrate the cohesive energy into a single hydrogen bond (the energy is actually spread over many hydrogen bonds). The resulting bond energy is $E_0 \sim 2\epsilon_c$. Let's also assume that the bond is well approximated as the interaction of two dipoles. Keep one dipole (one molecule) fixed. Through what angle, θ , does the unfixed molecule rotate, because of thermal motion? For small θ , the dipole–dipole potential can be written

$$U(\theta) \sim -E_0(1 - \theta^2). \quad (5.31)$$

To find θ , we equate $\Delta U(\theta)$ to kT_{vap} , and find $E_0\theta^2 \sim kT_{\text{vap}}$, or $\theta^2 \sim kT_{\text{vap}}/2\epsilon_c$, which is $0.5F^{-1}$. The solid angle is roughly $\pi\theta^2$, so the ratio of solid angles is

$$r_{\text{rot}} \sim \frac{4\pi}{\pi 0.5F^{-1}} \sim 8F. \quad (5.32)$$

This result has a strange feature: The hydrogen-bond strength appears only indirectly, through F . The quantity F remains roughly 10, so the hydrogen bond strength hardly alters the rotational correction factor. This invariance has to be wrong. To see why, consider a nonpolar substance; we can say that it has hydrogen bonds, but that they are extremely weak. Should we then add $k \log r_{\text{rot}}$ to the entropy? If we do, we spoil the excellent agreement

between theory and experiment for those substances. There's a way out of our dilemma: This estimate of r_{rot} is valid for only small θ . If the hydrogen bond is too weak, then we cannot rely on it; in such a case, the available solid angle in the liquid is 4π , and $r_{\text{rot}} \sim 1$.

Reassured that our result is not nonsense, we use it to estimate the boiling temperature of water. For water, the preceding theories gave $F \sim 10$; more precisely, successive approximation applied to (5.27) gave $F \sim 10.9$. So $\log r_{\text{rot}} \sim 4.46$, and we should adjust F to $10.9 + 4.46 = 15.36$ (because F is the entropy without Boltzmann's constant). Then we recalculate r_{rot} , and from that, recalculate F , and so on; the procedure quickly converges to $F = 15.7$; the predicted boiling temperature is 368 K. Our prediction is in error by only 1 percent! You are invited to make the calculations for the other hydrogen-bonded substances.

5.2.2 Melting

The **heat of fusion**—the energy released in freezing—is more difficult to estimate than the heat of vaporization is. We have a clear picture of what happens when a liquid boils: Molecules leave the liquid, and join the vapor; boiling destroys the short-range order in the liquid. What happens when a solid melts? The short-range order does not change significantly; atoms in a solid are closely packed, as are atoms in a liquid. Only the weaker long-range order changes; the structured lattice of the solid becomes the chaos of the liquid. This change is difficult to quantify. We can make an attempt by using the hole model to estimate the entropy change.

5.2.2.1 ENTROPY OF FUSION OR MELTING. We assume that, at the melting point, every atom is surrounded by 1 hole: We replace one-sixth of the atoms with holes. Let there be N atoms. What is the entropy of this configuration? Equivalently, how many arrangements of $(5/6)N$ atoms and $N/6$ holes are there? The entropy is the logarithm of this number (times Boltzmann's constant). For generality, let β be the fraction $1/6$. There are

$$W = \binom{N}{\beta N} = \frac{N!}{(\beta N)!((1 - \beta)N)!}$$

possible arrangements. We have used the symbol W for the number of states in honor of Boltzmann, whose epitaph reads $S = k \log W$. To compute S , we need $\log W$:

$$\log W = \log N! - \log\{(\beta N)!\} - \log\{((1 - \beta)N)!\}. \quad (5.33)$$

For large N , we can use Stirling's formula, that $\log x! \sim x \log x$. A more accurate statement is that $\log x! \sim x \log x - x$, but the $-x$ term always cancels in the log of a binomial coefficient. Then

$$\log W \sim N \log N - N\beta \log(\beta N) - N(1 - \beta) \log\{(1 - \beta)N\}. \quad (5.34)$$

Many terms cancel after we expand the logarithms. What remains is

$$\log W \sim -N(\beta \log \beta + (1 - \beta) \log(1 - \beta)). \quad (5.35)$$

The entropy per atom is

$$s = \frac{k \log W}{N} \sim -\beta \log \beta - (1 - \beta) \log(1 - \beta). \quad (5.36)$$

For $\beta = 1/6$, the entropy of liquid is $s_{\text{liquid}} \approx 0.5k$ (per atom). In this model, a solid is a liquid with $\beta = 0$, so $s_{\text{solid}} = 0$. Then $\Delta s \approx 0.5k$. Experimentally, $\Delta s \sim k$ is more accurate; the hole model accounts for only one-half of the entropy change. As we noted before, liquids are poorly understood. In what follows, we use $\Delta s \sim k$, or, per mole, $\Delta S \sim R$.

5.2.2.2 MELTING TEMPERATURE. The entropy is difficult to calculate, and the melting temperature even harder. T_{melt} must be less than the boiling temperature, but how much less? We make the simplest assumption, that $T_{\text{melt}} \sim T_{\text{vap}}$. From the entropy and the melting temperature, we can estimate the heat of fusion:

$$L_{\text{fus}} \sim T_{\text{melt}} \Delta S \sim RT_{\text{melt}} \sim RT_{\text{vap}}. \quad (5.37)$$

Equation (5.17) gives us an estimate for L_{vap} per molecule: $L_{\text{vap}} \sim 10kT_{\text{vap}}$. Per mole,

$$L_{\text{vap}} \sim 10RT_{\text{vap}}. \quad (5.38)$$

When we substitute this estimate into (5.37), we get the estimate

$$L_{\text{fus}} \sim \frac{L_{\text{vap}}}{10}. \quad (5.39)$$

Then, using (5.15) with $\epsilon_c \sim 0.5 \text{ eV}$ (for water), we find the estimates

$$\begin{aligned} L_{\text{fus}} &\sim 10^3 \text{ cal mole}^{-1}, \\ L_{\text{vap}} &\sim 10^4 \text{ cal mole}^{-1}. \end{aligned} \quad (5.40)$$

Table 5.4 compares heats of fusion and vaporization.

<i>Property</i>	L_{vap} ($10^4 \text{ cal mole}^{-1}$)	L_{fus} ($10^3 \text{ cal mole}^{-1}$)
Water	1.0	1.5
Hg	1.5	0.6
Au	8	10

Table 5.4. Comparison of heats of fusion and of vaporization. The estimate in (5.40) is reasonably accurate. That rule compares the second and the third columns. We can also compare values within a column, and try to understand the variation. Water and mercury are liquids at room temperature, so their bonds are weak compared to the bonds in gold, which is a solid at room temperature. Not surprisingly, gold has higher heats of vaporization and of fusion.

5.3 Specific heat

How much energy does it take to heat water to bath temperature? How many days of solar heating can the oceans store? The answers to these questions depend on the energy that the substance stores per unit temperature change: the **specific heat**. We approximate this quantity for metals and for insulators (dielectrics).

Before thinking about the physics of specific heats, we make the usual dimensional estimate. The units of specific heat are

$$[\text{specific heat}] = \frac{\text{energy}}{\text{temperature} \times \text{amount of substance}}. \quad (5.41)$$

The amount can be whatever size is convenient: one mole, one molecule, one gram, and so on. For our dimensional estimate, we choose the molecule, because it is a natural size: It involves the fewest arbitrary parameters. A mole, or a gram, for example, depends on human-chosen sizes. We already know one quantity with units of energy per temperature: the Boltzmann constant. So our first estimates are

$$\begin{aligned}\frac{\text{specific heat}}{\text{molecule}} &\sim k, \\ \frac{\text{specific heat}}{\text{mole}} &\sim kN_A = R.\end{aligned}\tag{5.42}$$

Table 5.5 lists the specific heat of various substances.

We now estimate the missing constant of proportionality using physical reasoning.

<i>Substance</i>	c_p/k
I ₂	4.4
Cl ₂	4.1
O ₂	3.5
N ₂	3.5
Ni	3.1
Au	3.1
Zn	3.1
Fe	3.0
Xe	2.5
He	2.5
C (diamond)	
0 °C	0.6
223 °C	1.6
823 °C	2.6

Table 5.5. *Specific heats at constant pressure. All data are for room temperature (unless otherwise noted) and atmospheric pressure. We have listed the specific heats in dimensionless form—in units of k per atom or molecule—because we already know that the specific heat must contain a factor of k . Source: [16, p. 155].*

5.3.1 Ideal gases

We first study ideal gases, because they illustrate the basic ideas of thermal modes, the importance of quantum mechanics, and the difference between constant-pressure and constant-volume calculations. We choose the simplest ideal-gas atom: helium. Hydrogen, though containing fewer protons and electrons per atom, is diatomic: It exists as H₂. Helium, an inert gas, is monoatomic, and so more suited for beginning our study.

5.3.1.1 HELIUM. What is the specific heat of helium? Imagine a collection of helium atoms. In three dimensions, it has three degrees of freedom, one for each dimension. A degree of freedom corresponds to a quadratic contribution to the energy of a molecule (here, quadratic

means quadratic in a position or velocity coordinate). The translational energy is $m(v_x^2 + v_y^2 + v_z^2)/2$; each term corresponds to a degree of freedom. At temperature T , each degree of freedom contributes energy $kT/2$ to the internal energy of the atom; so the internal energy per atom is $u = 3kT/2$, and the specific heat per atom is $u/T = 3k/2$. The $kT/2$ -per-degree-of-freedom rule holds for only classical degrees of freedom; the meaning of this distinction will become clear when we discuss nitrogen in (5.3.1.2). The specific heat should therefore be $1.5k$. However, the value given in Table 5.5 is roughly $2.5k$. We somehow neglected a contribution of kT per atom. What went wrong?

Specific heat measurements are of two types: constant pressure (the more common) and constant volume, depending which quantity is held fixed as the temperature changes. If the volume of the gas is held constant, then the specific heat is indeed $1.5k$; this specific heat is denoted c_v . However, if the pressure is constant, we must include another contribution. Consider N atoms of the gas in a container of volume V . To measure the specific heat, we increase the temperature by ΔT . If we leave the volume unchanged, the pressure increases by $\Delta P = nk\Delta T$, where $n = N/V$ is the number density of atoms. To keep the pressure constant, we must increase the volume by the same fraction that T increased: by $\Delta T/T$. So $\Delta V = \Delta T(V/T)$. Let A be the surface area of the box. Then $A dx = \Delta V$, where dx is how far the walls move. As the walls recede, the gas does work $F dx = P A dx = P \Delta V$. Thus, the heat required to raise the temperature must include $PV(\Delta T/T)$, which, for an ideal gas, is $Nk\Delta T$. This energy contributes k to the specific heat per atom. The specific heat at constant pressure is therefore $c_p = c_v + k$, which explains the formerly mysterious $c_p = 2.5k$. From the two specific heats, we can form a dimensionless group:

$$\gamma \equiv \frac{c_p}{c_v}, \quad (5.43)$$

which is $5/3$ for helium (and any other monoatomic ideal gas).

The value of γ determines how the temperature changes in an **adiabatic** change of state. An adiabatic change of state is a transformation that is rapid enough so that no heat flows to or from the substance. Such changes include the change that a parcel of air flowing up the side of a mountain experiences. As the parcel rises, it moves into regions with lower pressure, so it expands. This expansion is adiabatic; in an adiabatic expansion, the relation between pressure and volume is $PV^\gamma = \text{constant}$. So the value of γ determines how P and V change, which determines, via the ideal-gas law, how T changes. For air, $\gamma \sim 7/5$, as we see when we estimate c_p for nitrogen (the main constituent of air).

5.3.1.2 NITROGEN. Nitrogen is more complicated than helium, because it has two atoms in a molecule. How many degrees of freedom does the molecule have? Its center of mass can translate in three dimensions, which gives three degrees of freedom. The bond can vibrate, which could give *two* degrees of freedom, one for the potential energy in the bond, one for the relative motion of the atoms; it actually gives none, for reasons that will become clear shortly. The molecule can rotate about any of three axes; three axes could contribute three degrees of freedom (but actually contribute only two, as we see shortly). The total is five degrees of freedom (if we count none for vibration and two for rotation). Then $u = 5kT/2$, so $c_v = 5k/2$ and $c_p = 7k/2$, which agrees with the tabulated value.

What justifies neglecting the two vibrational degrees of freedom, and one of the rotational degrees of freedom? Each evidently contributes much less energy than $kT/2$. The reason why is quantum mechanics. The $kT/2$ rule is valid only for classical degrees of freedom. Consider, for example, a rotational degree of freedom (or mode). What makes it classical? In an exact analysis, there are only certain allowed rotations: Rotation is quantized. When does the quantization become irrelevant? Suppose that the spacing of the rotational energy levels is ϵ ; then mode is classical if $kT \gg \epsilon$. In that case, the energy levels are effectively continuous, and we can treat the mode with classical statistical mechanics. What happens if $kT \ll \epsilon$? Then the mode is mostly in the lowest energy level; the probability that the mode is in a higher level is roughly the Boltzmann factor $e^{-\epsilon/kT}$, which is minute. The average energy in the mode is roughly $\epsilon e^{-\epsilon/kT}$, which is $\ll kT/2$ when $kT \ll \epsilon$. Let's estimate ϵ for bond vibrations in nitrogen. We first estimate the bond-vibration frequency ω . The spring constant of the bond is roughly given by $k_s a^2/2 \sim E_0$. The vibration frequency is $\omega \sim \sqrt{2k/M}$, where M is the mass of a nitrogen atom. (We are being careful with factors of 2 in this derivation, because the result is equivocal if we are not careful, as you can check by neglecting the factors of 2.) So

$$\omega \sim 2\sqrt{\frac{E_0}{Ma^2}}. \quad (5.44)$$

In a spring, the energy-level spacing is $\epsilon = \hbar\omega$, so

$$\epsilon \sim 2\sqrt{\frac{\hbar E_0}{Ma^2}}. \quad (5.45)$$

To evaluate the spacing, we introduce c^2/c^2 inside the square root:

$$\epsilon \sim 2\sqrt{\frac{(\hbar c)^2 E_0}{(Mc^2)a^2}}. \quad (5.46)$$

The nitrogen molecule contains a triple bond, so E_0 is substantial, perhaps 7 eV; and a is short, perhaps 1 Å (triple bonds are shorter than double bonds, which are shorter than single bonds). The mass of a nitrogen atom is $M \sim 14m_p$. So

$$\epsilon \sim 2\sqrt{\frac{4 \cdot 10^6 \text{ eV}^2 \text{ Å}^2 \times 7 \text{ eV}}{14 \cdot 10^9 \text{ eV} \times 1 \text{ Å}^2}}, \quad (5.47)$$

because $\hbar c \sim 2000 \text{ eV Å}$ and $m_p c^2 \sim 10^9 \text{ eV}$. After the dust clears, we find $\epsilon \sim 0.1 \text{ eV}$. The thermal energy is $kT \sim 0.025 \text{ eV}$, so $kT \ll \epsilon$: The vibration mode is not classical. We say that the vibrations are *frozen out*: Room temperature is cold compared to their intrinsic energy (or temperature) scale. A similar calculation shows that rotation about the interatomic axis is also frozen out. So quantum mechanics explains why $c_v \sim 5k/2$. Then c_p should be $7k/2$, and it is.

5.3.1.3 IODINE. Iodine is also diatomic, but it is more massive than nitrogen (larger M); and its bond is weaker (lower E_0) and longer (larger a). We estimate ϵ_I by scaling the result for nitrogen. Relative to nitrogen, the mass of an iodine atom is larger by roughly a factor of 10; let's say that the bond is weaker by a factor of 2, and longer by a factor of 2. Then $\epsilon_I \sim \epsilon_N/6 \sim 0.015 \text{ eV}$. Now $kT > \epsilon$, so the two bond degrees of freedom unfreeze. (The interatomic-axis rotational mode is still frozen out.) We therefore expect that, for iodine, $c_v \sim 7k/2$, and $c_p \sim 9k/2$. We are not disappointed: The tabulated value is $c_p \sim 4.4k$.

5.3.2 Liquids and gases

We now estimate specific heats of liquids and solids. Liquids and solids hardly change volume, even if the volume is not held constant, so $c_p \simeq c_v$. If each molecule sits in a three-dimensional harmonic potential produced by the rest of the lattice, each molecule has 3 potential-energy degrees of freedom. Combining them with the three translational degrees of freedom produces 6 total degrees of freedom. The energy per molecule is therefore $3kT$, and $c_p \sim 3k$. The specific heat per mole is

$$C_p \sim 3R \sim 24 \frac{\text{J}}{\text{mole K}} \sim 6 \frac{\text{cal}}{\text{mole K}}, \quad (5.48)$$

which is a useful number to remember. This value is the lattice specific heat. The prediction is quite accurate for most of the solids and liquids (Ni, Au, Zn, and Fe) in Table 5.5. For diamond, the prediction is accurate at only high temperature. As in Section 5.3.1.2, we have an example of frozen modes. The bonds in diamond are extremely strong, and therefore have a high vibration frequency, and large level spacing. At room temperature, the thermal energy is not large compared to the energy-level spacing, so most bond vibration is frozen out. Only at $T \sim 800^\circ\text{C}$ is kT large enough to make bond vibration a classical mode; at that temperature, c_p is $2.6k$, which is close to our prediction of $3k$. The other curious fact is that nickel, gold, and zinc have c_p slightly greater than $3k$! These substances are all metals; we neglected a contribution to the specific heat that occurs in metals: the specific heat of the electrons. In metals, the electrons are not bound to an atom, but are free to move through the lattice; a small fraction of the electrons can store thermal energy. Electrons therefore make a small contribution to the specific heat. In Section 5.6.2, we explain why the fraction of contributing electrons is small.

5.4 Thermal diffusivity of liquids and solids

We would like to estimate the rate of heat transport—the thermal conductivity. One piece in that calculation is the specific heat: how much heat one molecule, or one mole, stores. The other piece, which is the topic of this section, is the **thermal diffusivity**. The thermal diffusivity determines how rapidly heat spreads—say, into the center of a turkey, where it denatures (cooks) the proteins. Heat is the vibration of atoms. In a solid, the atoms are confined in a lattice, and the vibrations can be represented as combinations of many sound waves. More precisely, the waves are phonons, which are sound waves that can have polarization (just as light waves have polarization). Heat diffusion is the diffusion of phonons.

Let's estimate how long it takes for phonons (or heat) to diffuse a macroscopic distance L . Phonons act like particles: They travel through lattice, bounce off impurities, and bounce off other phonons. The phonon mean free path λ measures how far a phonon travels before bouncing (or scattering), and then heading off in a random direction. In a solid without too many defects, at room temperature, typically $\lambda \sim 10 \text{ \AA}$. After a scattering, the phonon heads off in a random direction; so at each scattering, the phonon takes one step of a random walk with step size λ . After N steps, it has gone an average distance $\lambda\sqrt{N}$; this square-root dependence is characteristic of random walks. See the classic book *Random Walks in Biology* [2], which has an excellent treatment of random walks, illustrated with fascinating examples. So it needs L^2/λ^2 steps to travel a distance L . How long does each step take? A

typical phonon velocity is given by the velocity of sound c_e (the e indicates elastic wave); for most solids or liquids, $c_e \sim 3 \text{ km s}^{-1}$, as we found in Section 4.3.1. (We treat a liquid as a disordered solid—as a solid with low λ .) The time between scatterings is λ/c_e . So the time τ to travel a distance L is

$$\tau \sim \frac{L^2}{\lambda^2} \frac{\lambda}{c_e} = \frac{L^2}{\lambda c_e}. \quad (5.49)$$

The factor in the denominator is a characteristic of the substance, and—with a magic factor of one-third—is the thermal diffusivity:

$$\kappa \sim \frac{1}{3} \lambda c_e \sim \frac{1}{3} \times 10^{-7} \text{ cm} \times 3 \cdot 10^5 \text{ cm s}^{-1} \sim 10^{-2} \text{ cm}^2 \text{ s}^{-1}. \quad (5.50)$$

The magic one-third comes from doing honest physics—from solving messy differential equations, from worrying about boundary conditions, from not treating every object as a sphere, and so on. In terms of our new constant κ , the diffusion time is

$$\tau \sim L^2/\kappa. \quad (5.51)$$

Let's look at two examples of (5.51).

Example 5.2 Cooking a turkey

We apply our knowledge to Thanksgiving, an American holiday in which large turkeys are cooked, and many people give thanks once all the turkey leftovers are finished. How long does it take to cook a $R \sim 20 \text{ cm}$ turkey?

$$\tau_{\text{turkey}} \sim \frac{(20 \text{ cm})^2}{10^{-2} \text{ cm}^2 \text{ s}^{-1}} \sim 10^5 \text{ s} \sim 0.5 \text{ days}. \quad (5.52)$$

Start cooking early in the morning! A similar example that you can try is to predict the cooking time for an egg.

This estimate ignored an important parameter: the oven temperature. We showed that diffusion equalizes the central temperature to the oven temperature after roughly 0.5 days. However, the inside must still cook (the proteins must denature) after reaching this temperature. So now we have the explanation of why the oven must be hot: A cold oven would not cook the meat, even after the meat reached oven temperature. What oven temperature is hot enough? Protein physics is complicated; we cannot expect to estimate the temperature from theory, so we rely on experiment. We know from cooking meals that a thin piece of meat next to a hot skillet (perhaps at 200°C) cooks in a few minutes; we also know that if the skillet is at 50°C , the meat will never cook (50°C is not much hotter than body temperature). So if we set the oven to 200°C , the turkey will cook in a few minutes after attaining oven temperature; most of the cooking time is spent attaining oven temperature, which is the assumption that we made when we estimated the cooking time as L^2/κ .

Example 5.3 Cooling the moon

How long does it take for the moon ($R \sim 1.7 \cdot 10^8$ cm) to cool? For the moon, we find that

$$\tau_{\text{moon}} \sim \frac{(1.7 \cdot 10^8 \text{ cm})^2}{10^{-2} \text{ cm}^2 \text{ s}^{-1}} \sim \frac{3 \cdot 10^{16}}{10^{-2}} \text{ s} \sim 3 \cdot 10^{18} \text{ s} \sim 10^{11} \text{ yr}. \quad (5.53)$$

[Another useful number: $1 \text{ yr} \sim 3 \cdot 10^7 \text{ s}$.] The universe has been around for only 10^{10} yr; so why is the moon cold already, especially if heat is generated at its core by radioactive heating? Because of solid-state **convection**: Rock flows slowly (on a billion-year timescale), and transports heat to the surface much faster than molecular motions could.

In this moon example, we see one use of order-of-magnitude physics. We could—with lots of computer time—solve the diffusion equation for a mostly-spherical body like the moon. We might find that the more accurate cooling time is $6.74 \cdot 10^{10}$ yr. What would that result gain us? The more accurate time is still far longer than the age of the universe; we would still conclude that convection must be responsible. Order-of-magnitude analysis allow you to determine quickly which approaches are worth pursuing; you can then spend more time refining productive approaches, and less time chasing unproductive ones.

5.5 Diffusivity and viscosity of gases

Before estimating the thermal conductivity of liquids and solids, we pause to harvest results that follow easily from our understanding of the mechanism of diffusivity: We estimate the thermal diffusivity and the viscosity of a gas. In a solid or liquid, the atoms themselves do not transport the heat that they store; phonons, which are much more mobile than the atoms, do. In a gas, the analogue of phonons is sound waves, but sound waves do not carry the heat in a gas; the atoms (or molecules) themselves carry the heat. We can estimate the thermal diffusivity using (5.50), but with parameters for a gas:

$$\kappa_{\text{gas}} \sim \frac{1}{3} \lambda v. \quad (5.54)$$

Now, λ is the mean free path of a gas molecule, and v is the velocity of the molecule. The velocity we can estimate easily—it is the thermal velocity. The mean free path is given by $\lambda \sim (n\sigma)^{-1}$, where n is the number density of molecules, and σ is the cross-sectional area of a molecule. This result is dimensionally correct (a useful sanity check), and it has the right quantities downstairs: Larger molecules scatter after a shorter distance than smaller molecules do, and denser gases cause a molecule to scatter sooner than a less dense gas does. So

$$\kappa_{\text{gas}} \sim \frac{1}{3} \frac{v}{n\sigma}. \quad (5.55)$$

For air at room temperature and atmospheric pressure, the cross section is $\sigma \sim a^2$, where $a \sim 3 \text{ Å}$; the number density is $n \sim 3 \cdot 10^{19} \text{ cm}^{-3}$; and the thermal velocity is $v \sim 3 \cdot 10^4 \text{ cm s}^{-1}$

(roughly the speed of sound). Then

$$\begin{aligned}\kappa_{\text{gas}} &\sim \frac{1}{3} \frac{3 \cdot 10^4 \text{ cm s}^{-1}}{3 \cdot 10^{19} \text{ cm}^{-3} \times 10^{-15} \text{ cm}^2} \\ &\sim 0.3 \text{ cm}^2 \text{ s}^{-1}.\end{aligned}\tag{5.56}$$

This result is too large by a factor of only 1.5: the measured value is $\kappa \sim 0.2 \text{ cm}^2 \text{ s}^{-1}$. Most of the error is in our estimate of σ .

What about kinematic viscosity, ν ? Thermal diffusivity is the heat- or energy-diffusion coefficient, and viscosity is the momentum-diffusion coefficient. In a gas, momentum and energy are transported by the same mechanism: molecular motion. So we estimate that $\nu \sim \kappa$, and we are rewarded: $\nu \sim 0.15 \text{ cm}^2 \text{ s}^{-1}$. Kinematic viscosity and thermal diffusivity have the same units. A natural dimensionless group is therefore

$$Pr \equiv \frac{\nu}{\kappa},\tag{5.57}$$

which is the Prandtl number. It is close to unity in gases, and even in many solids and liquids (for water, it is 6).

5.6 Thermal conductivity

Why can people (who are not too confident) walk on hot coals without getting burned? How long does the 3-inch layer of frost in the freezer take to melt? The answers to these questions depend on the **thermal conductivity**—how fast a material can transport heat. Now that we have estimated the specific heat and the thermal diffusivity, we can estimate the thermal conductivity.

5.6.1 Dielectrics

We first study the simpler case of dielectrics, where we can neglect the electronic contribution to thermal conductivity. The thermal conductivity is defined as the constant K in the heat-flow equation

$$F = -K \frac{dT}{dx},\tag{5.58}$$

where F is the heat flux (power per area), and T is the temperature. We can estimate the value of the constant using our understanding of specific heat and thermal diffusivity. Consider a rectangle of material, with length x and cross-sectional area A . Heat one half (measuring along the length) to temperature T , and start the other at $T = 0$. One half contains a quantity of heat $H \sim C'_p T x A$, where C'_p is the heat capacity per volume. After a time $\tau \sim x^2/\kappa$, the temperature of the two halves equalizes (because of diffusion); thus, heat must flow from the hot to the cold half at a rate $\dot{H} \sim H/\tau \sim C'_p A T \kappa/x$. The heat flux is $F \sim \dot{H}/A$, or $F \sim C'_p \kappa T/x$. Now we match terms with the defining equation (5.58); dT/dx corresponds to T/x , so $C'_p \kappa$ corresponds to K . In fact, this relation is exact: $K = C'_p \kappa$

The specific heat per volume is C_p/V_{mol} , where V_{mol} is the molar volume. Therefore,

$$K = \frac{C_p \kappa}{V_{\text{mol}}}.\tag{5.59}$$

In our order-of-magnitude world, the volume of one molecule is a^3 , where $a \sim 3 \text{ \AA}$. So $V_{\text{mol}} \sim N_A a^3$, which is $16 \text{ cm}^3 \text{ mole}^{-1}$. We also substitute κ from (5.50), and C_p from (5.48), and we find

$$K \sim \frac{6 \text{ cal}}{\text{mole K}} \times 10^{-2} \text{ cm}^2 \text{ s}^{-1} \times \frac{\text{mole}}{16 \text{ cm}^3} \sim 10^{-2} \frac{\text{cal}}{\text{cm s K}}. \quad (5.60)$$

We neglected a factor of 2.5 so that the answer comes out in a round number. This value includes only the lattice transport of heat; it applies to dielectrics, but not to metals, where the electrons transport most of the heat. Table 5.6 contains thermal conductivities for dielectrics.

<i>Substance</i>	K $\left(10^{-3} \frac{\text{cal}}{\text{cm s K}}\right)$
Water	5
Glass	3
Rock	2
Wood	0.3
Asbestos	0.2

Table 5.6. Thermal conductivities for common dielectrics. Note that we chose the typical dielectric thermal conductivity as the unit size: We **scaled** the value relative to a reasonable estimate. Scaled means that we measure the physical quantities in units that we expect to be reasonable: in units that should yield a number near unity.

5.6.2 Metals

Metals feel colder to the touch than, say, concrete does because metals transport heat faster than concrete does. If lattice conduction is the only factor responsible for heat conduction, then metals and concrete should not have such a great disparity in conductivity. That they do is a consequence of the large contribution to the thermal conductivity from free electrons in a metal. [A related example: A slab of granite feels much colder than a slab of cork does. Why does granite carry heat more efficiently than cork, even though both substances are dielectrics, and therefore carry heat via lattice conduction? The answer is that granite is much denser than cork, so its specific heat per volume is much larger than cork's.]

Let's estimate the ratio of electronic to lattice thermal conductivity. Because the conductivity factors into diffusivity \times specific heat, we first estimate the ratio of electronic to lattice diffusivities. Diffusivity is

$$\text{velocity} \times \text{mean free path}, \quad (5.61)$$

so we subdivide the estimate of the diffusivity ratio into estimates of the velocity ratio and of the mean-free-path ratio.

Phonons move at speeds similar to the speed of sound—a few kilometers per second. Electrons move at the **Fermi velocity** v_F . We have seen this concept in disguise when we studied white dwarfs in Section 4.4. In a white dwarf, the electrons are confined by gravity; in a metal, they are confined by electrostatics (a metal is like a giant hydrogen atom). In both systems, the electrons are confined, which gives them a velocity. We could modify the white-dwarf analysis to account for these changes; instead, we just redo the relevant part. Let n_e be the number density of free electrons in the metal. Then $\Delta p \sim \hbar/\Delta x \sim \hbar n_e^{1/3}$, where Δp is the momentum uncertainty produced by confinement, and Δx is the position uncertainty. The velocity is the Fermi velocity: $v_F \sim \hbar n_e^{1/3}/m_e$. A typical metal has two free electrons per atom, and an interatomic spacing of $a \sim 3 \text{ \AA}$. So $n_e \sim 2/a^3 \sim 10^{23} \text{ cm}^{-3}$. Then $v_F \sim 8 \text{ cm s}^{-1} = 1000 \text{ km s}^{-1}$. This velocity is much faster than a typical elastic wave. So the velocity ratio is ~ 300 (electrons move faster).

Electrons also have a much longer mean free path than do phonons; in copper, $\lambda_e \sim 100a$, where $a \sim 3 \text{ \AA}$ is a typical lattice spacing. The phonon mean free path is $\lambda_p \sim 10 \text{ \AA}$. So the mean-free-path ratio is ~ 30 , and the diffusivity ratio is $300 \times 30 \sim 10^4$ in favor of the electrons.

We now have to compute the specific-heat ratio. Here, the phonons win, because only a small fraction of the electrons carry thermal energy. To understand why, we have to improve the Fermi-sphere argument that we used to compute v_F . A more accurate picture is that the metal is a three-dimensional potential well (the electrons are confined to the metal). There are many energy levels in the well, which we can label by the momentum of the electrons that occupy it. The free electrons fill the lowest levels first (two per level); the Pauli principle forbids more than two electrons from being in the same energy level. The electrons in the highest levels have velocity v_F (that part of the previous argument was correct), or momentum $m_e v_F$. As vectors, the momenta of the highest-energy electrons are uniformly distributed over the surface of a sphere in momentum space; the sphere has radius $m_e v_F$, and is called the Fermi sphere. How is this sphere relevant to specific heat? To carry thermal energy, an electron has to be able to absorb and deposit energy; the typical package size is kT . Consider an electron that wants to absorb a thermal-energy package. When it does so, it will change its energy—it will move outward in the Fermi sphere. But can it? If it is in most of the sphere, it cannot, because the sphere is packed solid—all interior levels are filled. Only if the electron is near the surface of the sphere can it absorb the package. How near the surface must it be? It must have energy within kT of the Fermi energy \mathcal{E}_F (the Fermi energy is $\sim m_e v_F^2$). The fraction of electrons within this energy range is $f \sim kT/\mathcal{E}_F$. Typically, $kT \sim 0.025 \text{ eV}$ and $\mathcal{E}_F \sim \text{few eV}$, so $f \sim 10^{-2}$. This fraction is also the specific-heat ratio. To see why, consider the case where $f = 1$ (every electron counts). Each atom contributes $3k$ to the specific heat, and each electron contributes one-half of that amount— $3k/2$ —because there is no spring potential for the electrons (they contribute only translational degrees of freedom). The number of free electrons is typically twice the number of atoms, so the total electron and phonon contributions to specific heat are roughly equal. When $f \sim 10^{-2}$, the contributions have ratio 10^{-2} .

Now we have all the pieces. The conductivity ratio is

$$\frac{K_{\text{metal}}}{K_{\text{dielectric}}} \sim 10^4 \times 10^{-2} = 10^2. \quad (5.62)$$

We use (5.60) to estimate $K_{\text{dielectric}}$, and find that

$$K_{\text{metal}} \sim 1 \frac{\text{cal}}{\text{cm s K}}. \quad (5.63)$$

Note that this estimate is for a good conductor, such as copper, with a large electron mean free path. Table 5.7 contains data on the thermal conductivities of common metals. Our estimate in (5.63) is quite accurate for copper—more accurate than it ought to be given the number of approximations that we made to derive it.

<i>Substance</i>	$K \left(\frac{\text{cal}}{\text{cm s K}} \right)$
Al	0.5
Cu	1.0
Fe	0.2
Hg	0.02
W	0.5

Table 5.7. *Thermal conductivities for common metals at room temperature. Metals should have thermal conductivities of roughly $1 \text{ erg cm}^{-1} \text{ s}^{-1} \text{ K}^{-1}$, so we measure the actual values in that unit. We are being gentle with our neural hardware, giving it the kind of numbers that it handles with the least difficulty: numbers close to 1. Copper has a large thermal conductivity, because the electron mean free path is long (which is the same reason that copper has a high electrical conductivity). Mercury has a low thermal conductivity, because it is a liquid at room temperature, so the electron mean free path is short (only one or two lattice spacings).*

5.7 What you have learned

- *The method of successive approximation:* How to solve complicated equations one step at a time. We saw two examples of this method: in estimating the thermal-expansion coefficient, and in solving for the boiling point as a function of cohesive energy.
- *How to handle logarithms:* Every (natural) log is 10. This rule is often helpful for beginning a successive-approximation solution.
- *The microscopic basis of thermal diffusivity and viscosity:* Particles (or phonons) move in steps whose size is equal to the mean free path, λ . The particles' velocity v determines the time to take one step, and therefore the diffusion constant, which is $\kappa \sim v\lambda$.

6 WAVES

Ocean covers most of the earth, and waves roam most of the ocean. What makes waves move, and what determines their speed? We now apply and extend our techniques of approximation to answer such questions, as we study water waves. The theme of this chapter is: Consider limiting cases.

6.1 Dispersion relations

We study waves using **dispersion relations**. A dispersion relation states what values of frequency and wavelength a wave can have; see for example, Crawford's book [8] for more information on dispersion relations. In their usual form, dispersion relations connect frequency ω , and wavenumber k . Wavenumber is defined as $2\pi/\lambda$, where λ is the wavelength. As an example, for electromagnetic waves in a vacuum, the frequency and wavenumber are related by the dispersion relation $\omega = ck$, which states that waves travel at velocity $\omega/k = c$, independent of frequency. Dispersion relations contain a vast amount of information about waves. They tell us, for example, how fast crests and troughs travel: the **phase velocity**. They tell us how fast wave packets travel: the **group velocity**. They tell us how these velocities depend on frequency: the **dispersion**. And they tell us the rate of energy loss: the **attenuation**.

You usually find a dispersion relation by solving a **wave equation**, which is an unpleasant partial differential equation. For water waves, you obtain a wave equation by linearizing the Euler equation of hydrodynamics. This procedure is mathematically involved, particularly in handling the boundary conditions. Being impatient and lazy, we instead derive dispersion relations with dimensional analysis, then complete and complement the derivation with physical arguments. We cannot evaluate the dimensionless constants with our sloppy methods, but the beauty of studying waves is that these constants are often unity.

How can we connect frequency and wavenumber? The two quantities have dimensions $[T]^{-1}$ and $[L]^{-1}$, respectively. As glue, we need to include more variables, with various dimensions. So, as usual, we consider what physical properties of the system determine wave behavior, and thereby form a set of relevant variables. Waves on the open ocean are different from waves in your bathtub, presumably because of the difference in the depth of water, which we call h . The width of the tub or ocean could matter, but we neglect such effects, and consider waves that move in only one dimension, perpendicular to the width.

To determine what other variables are important, we note use the principle that waves are like springs, because *every physical process contains a spring*. This blanket statement cannot be strictly correct. However, as a broad generalization, it is useful. We can get a more precise idea of when this assumption is useful by considering the characteristics of spring motion. First, springs have an equilibrium position. If your system has an undisturbed, resting state, consider looking for a spring. For example, for waves on the ocean, the undisturbed state is a calm ocean. For electromagnetic waves—springs are not confined

to mechanical systems—the resting state is an empty vacuum, with no radiation. Second, springs oscillate. In mechanical systems, oscillation depends on inertia carrying the mass beyond the equilibrium position. Equivalently, it depends on kinetic energy turning into potential energy, and vice versa. Water waves store potential energy in the disturbance of the surface, and kinetic energy in the motion of the water. For electromagnetic waves, energy sloshes (oscillates) between the electric field and the magnetic field. A magnetic field can be generated by moving charges, so we call the magnetic field the reservoir of kinetic (motion) energy. An electric field can be generated by static charges, so we call the electric field the reservoir of potential energy. With these identifications, the electromagnetic field is a set of springs, one for each frequency. These examples are positive examples. A negative example—a marble oozing its way through corn syrup, which we studied in Section 3.3.2—illustrates when springs are not present. The marble moves so slowly that the kinetic energy of the corn syrup, and therefore its inertia, is miniscule and irrelevant. In such a system, there is no reservoir of kinetic energy, so a spring is not present.

We now return to the problem of waves, which have the necessary reservoirs. In its lowest-energy state, the surface of the water is level; deviations from flatness—waves—are opposed by a restoring force. In a real spring, the restoring force comes from interatomic bonds. Once we know the origin of a spring constant, we can calculate it; from the spring constant, we can calculate the oscillation frequency. Our model that waves are springs suggests that we study the restoring force in a wave. Distorting the surface is like stretching a rubber sheet; the surface tension of water opposes the distortion. Distorting the surface also requires raising the average water level, a change that gravity opposes. The restoring force is therefore a combination of gravity and surface tension. In our list of variables, then, we include surface tension, γ , and gravity, g . We consider how gravity (or surface tension) restores the water level when we introduce a simple physical model in Section 6.2.1.

In a wave, like in a spring, the restoring force fights inertia, represented here by the fluid density ρ . Gravity does not care about density, because gravity's stronger pull on denser substances is exactly balanced by their greater inertia. This exact cancellation is a restatement of the **equivalence principle**, on which Einstein based his theory of general relativity. In pendulum motion (Section 3.2), the mass of the bob drops out of the final solution for the same reason. However, surface tension does not grow stronger in proportion to density. So we expect ρ to determine the properties of waves in which surface tension provides the restoring force. We therefore include ρ in the list of variables. For concreteness, we often refer in this chapter to water waves; the results are general, and work for any fluid. We pretend that waves are lossless, and exclude viscosity from the set of variables. For simplicity, we also exclude c_s , the speed of sound in the medium (usually water); with this approximation, we ignore sound waves. We also exclude the wave amplitude ξ from our dimensional analysis by assuming that ξ is small compared to the other length scales (wavelength and depth): We study only small-amplitude waves. Table 6.1 shows the list of variables that we have collected.

These six variables consist of three fundamental dimensions. We therefore hunt for $6 - 3 = 3$ dimensionless groups. One group is easy: k is an inverse length, and h is a length, so we can form

$$\Pi_1 \equiv kh. \quad (6.1)$$

This group is the dimensionless depth of the water: $\Pi_1 \ll 1$ means shallow and $\Pi_1 \gg 1$ means deep water. A second dimensionless group comes from gravity. Gravity, represented by g , has the same units as ω^2 , except for a factor of length. The wavenumber fixes this deficit:

$$\Pi_2 = \frac{\omega^2}{gk}. \quad (6.2)$$

Notice that this group does not contain density. If we exclude surface tension, then Π_1 and Π_2 are the only dimensionless groups that we can form; without surface tension, the waves propagate because of gravity alone. The equivalence principle tells us that the way in which gravity affects motion is independent of density. Therefore, density cannot—and does not—appear in either group.

Now we consider what happens when we allow surface tension back into the playpen of dimensionless variables. It must belong in the third group. Knowing only that γ belongs to Π_3 , we still have great freedom in choosing the form of Π_3 . We choose it so that it measures the relative importance of gravity and surface tension in determining the wave motion. Surface tension has units of force per length, whereas ρg has units of force per volume, or force per length cubed. So the ratio $\gamma/\rho g$ has units of length squared. We can neutralize the two excess lengths with two factors of k :

$$\Pi_3 \equiv \frac{\gamma k^2}{\rho g}. \quad (6.3)$$

We could also have built this group without using gravity, choosing $\Pi_3 \equiv \gamma k^3/\rho\omega^2$. There are many possibilities for dimensionless groups. The choices that we made, which emphasize gravity over surface tension, are convenient for studying waves restored by gravity; most ocean waves are of that ilk.

<i>Var.</i>	<i>Dimensions</i>	<i>Description</i>
ω	$[\text{T}]^{-1}$	angular frequency
k	$[\text{L}]^{-1}$	wavenumber
g	$[\text{L}][\text{T}]^{-2}$	acceleration due to gravity
h	$[\text{L}]$	depth of water
ρ	$[\text{M}][\text{L}]^{-3}$	density
γ	$[\text{M}][\text{T}]^{-2}$	surface tension

Table 6.1. Variables that determine the behavior of water waves.

We want to solve for frequency, ω , as a function of wavenumber, k . We could also solve for k versus ω , but the relations for phase and group velocity are simpler with ω as a function of k . Only the group Π_2 contains ω , so we write the most general dimensionless relation as

$$\Pi_2 = f(\Pi_1, \Pi_3), \quad (6.4)$$

or

$$\frac{\omega^2}{gk} = f\left(kh, \frac{\gamma k^2}{\rho g}\right). \quad (6.5)$$

Solving for ω^2 , we find that

$$\omega^2 = gk \cdot f(kh, \frac{\gamma k^2}{\rho g}). \quad (6.6)$$

This relation is valid for waves in shallow or deep water; for waves propagated by surface tension or by gravity; and for waves in between. Figure 6.1 shows how the two groups Π_1 and Π_3 divide the world of waves into four regions. Figure 6.2 shows how wavelength and depth partition the world, and gives examples of different types of waves.

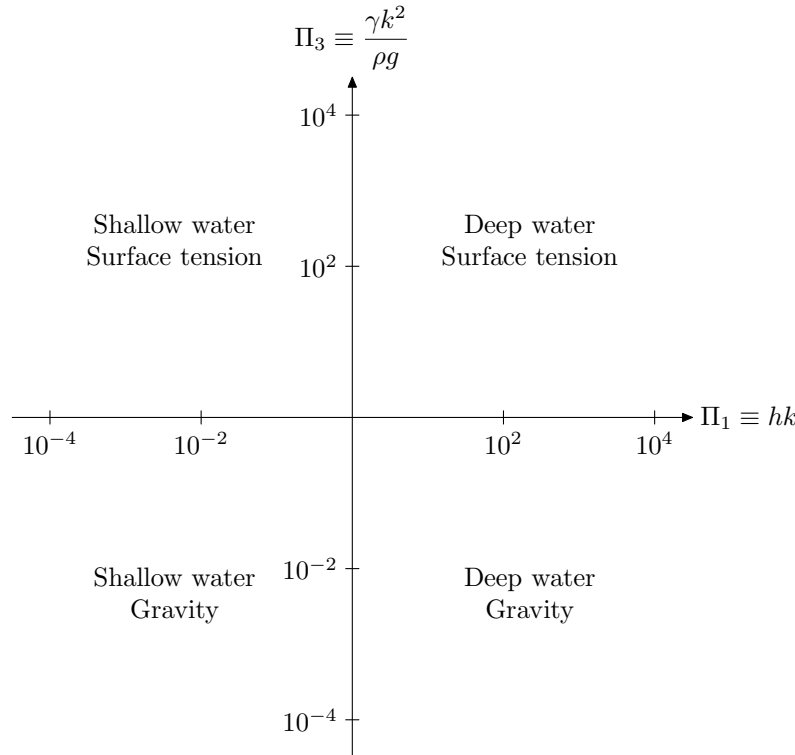


Figure 6.1. Map of the world of waves. The dimensionless groups Π_1 and Π_3 partition the world into four regions. We study them in turn, and combine our analyses to understand the whole world (of waves). The group Π_1 measures the depth of the water: Are we in a puddle or an ocean? The group Π_3 measures the relative contribution of gravity and surface tension: Are the waves ripples or gravity waves?

The division into deep and shallow water (left and right) follows from the interpretation of $\Pi_1 = kh$ as dimensionless depth. The division into surface-tension- and gravity-dominated waves (top and bottom) is more subtle. You can understand the division by looking at the form of Π_3 , which is $\gamma k^2 / \rho g$. Large g , or small k , result in the same consequence: small Π_3 . Therefore, the physical consequence of longer wavelength is similar to that of stronger gravity: Longer-wavelength waves are gravity waves. The large- Π_3 portion of the world (top) is therefore labeled with surface tension.

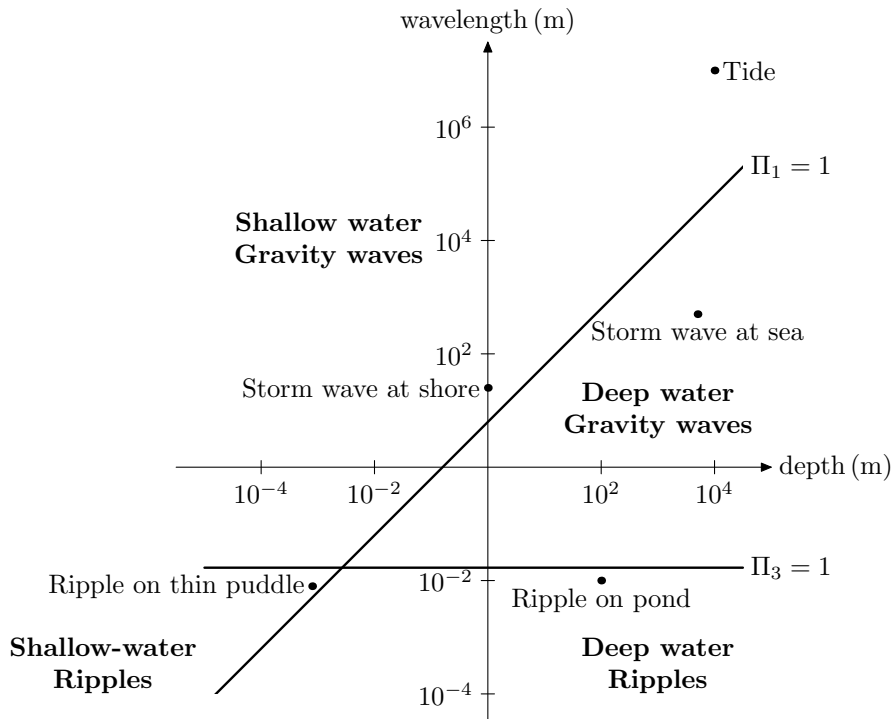


Figure 6.2. Water waves on the earth. Here, we partition the world using the familiar variables of depth and wavelength for the axes. The thick dividing lines are based on the dimensionless groups $\Pi_1 = hk$ and $\Pi_3 = \gamma k^2 / \rho g$. Each region contains one or two examples of its kind of waves. With $g = 1000 \text{ cm s}^{-1}$ and $\rho \sim 1 \text{ g cm}^{-3}$, the border wavelength between ripples and gravity waves is just over $\lambda \sim 1 \text{ cm}$ (the horizontal, $\Pi_3 = 1$ dividing line).

We do not as yet know the form of the magic function f in any region, let alone for all regions. To determine its form, and to understand its consequences, we make a simple physical model, by **considering limiting cases**. Like a jigsaw-puzzle-solver, we work out the corners of the world first, because the physics is simplest there. Then, we connect the solutions along the edges, where the physics is the almost as simple. Finally, we crawl our way inward to assemble the complicated, more complete solution. This chapter is a large example of the divide-and-conquer approach to solving problems, where limiting cases are the pieces.

6.2 Deep water

We first study deep water, where $kh \gg 1$, as shaded in the map in Figure 6.3. Deep water is defined as water sufficiently deep that waves cannot feel the bottom of the ocean. How deep do waves' feelers extend? The only length scale in the waves is the wavelength, $\lambda = 2\pi/k$. We expect, therefore, that the feelers extend to a depth $d \sim 1/k$ (as always, we neglect constants, such as 2π). We can justify this guess by using Laplace's equation, which is the spatial part of the wave equation. Suppose that the waves are periodic in the x direction, and z measures depth below the surface, as shown in Figure 6.4. Then, Laplace's equation

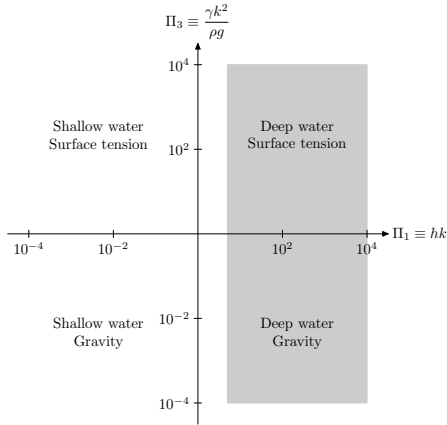


Figure 6.3. Deep-water portion of the world of waves.

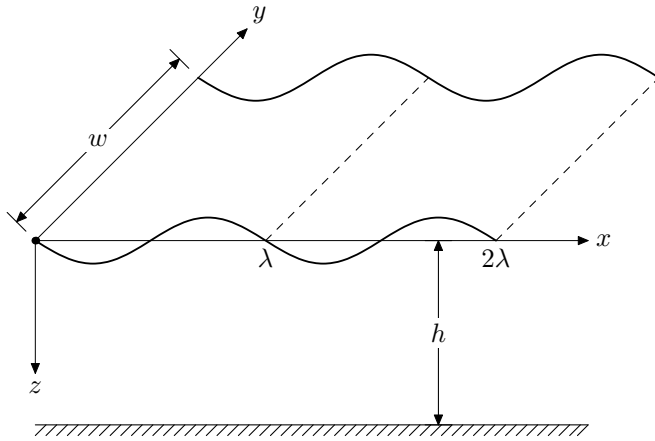


Figure 6.4. Piece of an infinite wave. The wave, with a sinusoidal profile, has wavelength λ . The water has depth h , and we study a width w in the y direction.

becomes

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial z^2} = 0, \quad (6.7)$$

where ϕ is the velocity potential. It's not important what exactly ϕ is. You can find out more about it in a fluid mechanics textbook. Our favorite is Faber's *Fluid Dynamics for Physicists* [13]; Lamb's *Hydrodynamics* [35] is a classic, but more difficult. For this argument, all that matters is that ϕ measures the effect of the wave; that where $\phi = 0$, there's no wave; and that ϕ satisfies Laplace's equation. The wave is periodic in the x direction, with a form such as $\sin kx$. We take

$$\phi \sim Z(z) \sin kx. \quad (6.8)$$

The function $Z(z)$ measures how the wave decays with depth.

The second derivative in x brings out two factors of k , and a minus sign:

$$\frac{\partial^2 \phi}{\partial x^2} = -k^2 \phi. \quad (6.9)$$

So that ϕ satisfies Laplace's equation, the z -derivative term must contribute $+k^2 \phi$. Therefore,

$$\frac{\partial^2 \phi}{\partial z^2} = k^2 \phi, \quad (6.10)$$

so $Z(z) \sim e^{\pm kz}$. The physically possible solution—the one that does not blow up exponentially at the bottom of the ocean—is $Z(z) \sim e^{-kz}$. Therefore, relative to the effect of the wave at the surface, the effect of the wave at the bottom of the ocean is $\sim e^{-kh}$. When $kh \gg 1$, the bottom might as well be on the moon. The dimensionless factor kh —it has to be dimensionless to sit alone in an exponent as it does—compares water depth with feeler depth $d \sim 1/k$:

$$\frac{\text{water depth}}{\text{feeler depth}} \sim \frac{h}{1/k} = hk, \quad (6.11)$$

which is the dimensionless group Π_1 .

In deep water, where the bottom is hidden from the waves, the water depth h does not affect the waves' propagation, so h disappears from the list of relevant variables. When h goes, so does $\Pi_1 = kh$. There is one caveat. If Π_1 is the only variable that contains k , then we cannot blithely discard it, just because we no longer care about h . Fortunately, k appears in $\Pi_3 = \gamma k^2 / \rho g$ as well. So it's safe to toss Π_1 for deep water. We have just argued that h is irrelevant based on a physical argument. This argument has a mathematical equivalent that we can express in the language of dimensionless groups and functions. The statement that h is large is meaningless, because h has units. The question that remains is, "large compared to what length?" When we choose $1/k$ as the standard of comparison—based on the Laplace's-equation argument—we can rephrase the meaningless " h is large" statement to " $\Pi_1 = kh$ is large." Using the Laplace's-equation argument, we drop Π_1 because h is irrelevant. Mathematically, we are assuming that the function $f(kh, \gamma k^2 / \rho g)$ from (6.6) has a limit as $kh \rightarrow \infty$.

Without Π_1 , the general dispersion relation (6.6) simplifies to

$$\omega^2 = gk f_{\text{deep}}\left(\frac{\gamma k^2}{\rho g}\right). \quad (6.12)$$

This relation describes the deep-water edge of the world of waves. The edge has two corners, labeled by whether gravity or surface tension provides the restoring force. Although we do not know the form of f_{deep} , it is a function simpler than the original f . To determine the form of f_{deep} , we partition deep-water waves into its two limiting cases: gravity waves and ripples.

6.2.1 Gravity waves on deep water

Now we specialize to regions of the wave world where water is deep and gravity is strong (Figure 6.5). This category includes most waves generated by wind, and wakes generated by ships. With gravity much stronger than surface tension, the dimensionless group $\gamma k^2 / \rho g$

limits to 0 (pretend that we cranked up g , or, by dumping soap on the water, that we turned down γ). So the general deep-water dispersion relation (6.12) simplifies to

$$\omega^2 = f_{\text{deep}}(0)gk = C_1 gk, \quad (6.13)$$

where $f_{\text{deep}}(0)$ is an as-yet-unknown constant, C_1 . [We do not know that $f_{\text{deep}}(x)$ has a limit as $x \rightarrow 0$. The slab argument, which follows shortly, shows that it does.] The constant remains unknown to our lazy methods, because the methods trade evaluation of dimensionless constants for comprehension of physics. We usually assume that such constants are unity. In this case, we get lucky: An honest calculation does produce $C_1 = 1$. Therefore,

$$\omega^2 = gk. \quad (6.14)$$

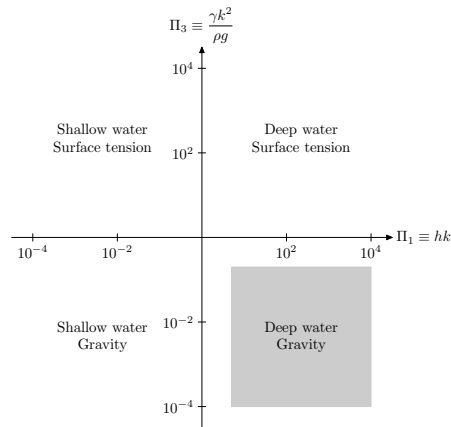


Figure 6.5. Deep-water-gravity-wave portion of the world of waves.

Such results from dimensional analysis seem like rabbits jumping from a hat. The relation (6.14) is correct, but our gut is not happy with this magical derivation; it asks, “Why is the result true?” We can understand the origin of the dispersion relation by making a simple physical model of the forces or energies that drive the waves. The first step is to understand the mechanism: How does gravity make the water level rise and fall? Taking a hint from the Watergate investigators,¹ we follow the water. The water in the crest does *not* move into the trough. Rather, the water in the crest, being higher, creates a pressure underneath it higher than that of the water in the trough, as shown in Figure 6.6. The higher pressure forces water that is underneath the crest to flow toward the trough, making the water level there rise. Like a swing sliding past equilibrium, the surface overshoots the equilibrium level, to produce a new crest, and the cycle repeats.

Our next step is to make this model quantitative, by estimating sizes, forces, speeds, and energies. In Section 2.1.4, we replaced a messy mortality curve with a more tractable

1. When the reporters Woodward and Bernstein [3] were investigating coverups during the Nixon administration, they received help from the mysterious “Deep Throat,” whose valuable advice was that they “follow the money.”

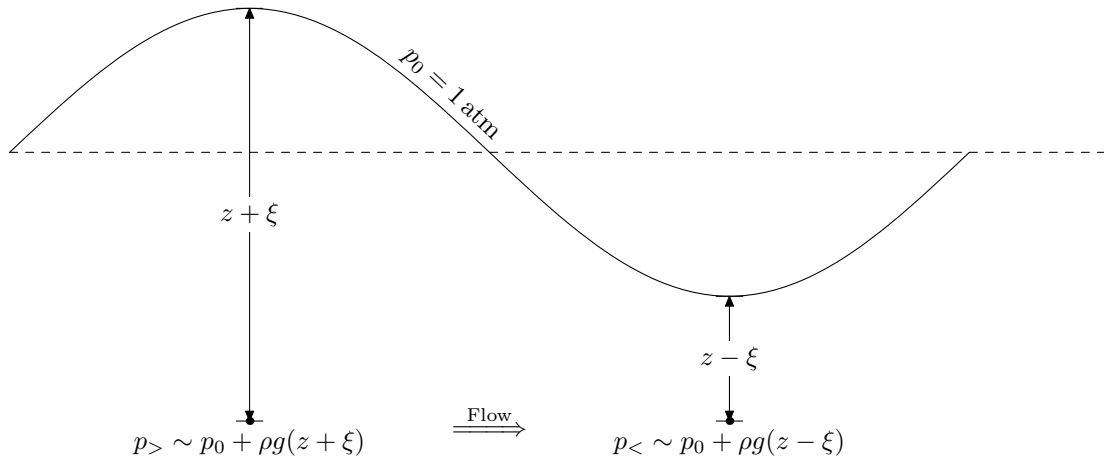


Figure 6.6. Higher pressure under the crest. The crest is higher than the trough by roughly ξ , the wave amplitude. This extra water creates a pressure underneath the crest larger than that under the trough, by roughly $\Delta p \sim \rho \xi g$. This pressure difference makes water flow from under the crest to under the trough: The wave advances.

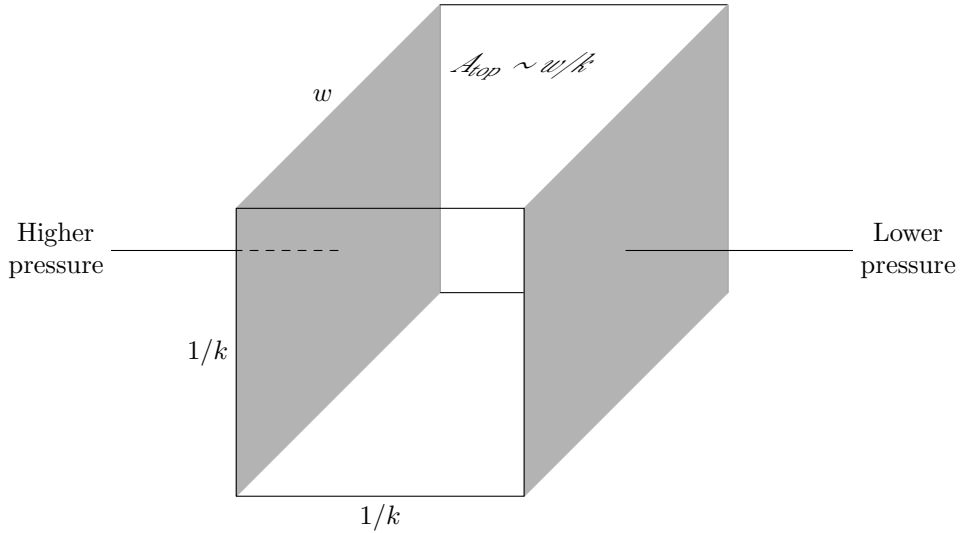


Figure 6.7. Slab of water. The pressure is higher under the crest (left shaded side) than toward the trough (right shaded side). The shaded sides have area $A \sim w/k$. From Figure 6.6, this extra pressure is $\Delta p \sim \rho \xi g$, and it acts on an area $A \sim w/k$ to produce a force $F \sim A \Delta p \sim \rho w \xi g/k$. The slab has mass $m \sim \rho w/k^2$. In the text, we use this mass to compute the acceleration and kinetic energy of the slab.

shape: a rectangle. We use the same trick repeatedly for this model. Water underneath the surface moves quickly because of the pressure gradient. Farther down, it moves more

slowly. Eventually, it does not move at all. We replace this falloff with a step function, and pretend that water down to a certain depth moves as a block, and deeper water stays still (Figure 6.7). How deep should this slab of water extend? By the Laplace-equation argument, the pressure variation falls off exponentially with depth, with length scale $1/k$. So we assume that the slab has depth $1/k$. (What choice do we have? On an infinitely deep ocean, there is no length scale other than $1/k$.) How long should the slab be? Its length should be roughly the peak-to-trough distance, because the surface height changes significantly in that distance. This distance is $1/k$. Actually, it's π/k , but we ignore the constants, such as π . All such constants combine into a constant at the end, which we cannot determine with dimensional analysis anyway, so why not discard it now? The width of the slab, w , is arbitrary. It drops out in all the analyses.

So the slab of water has depth $1/k$, length $1/k$, and width w . What forces act on it? We can determine the forces by estimating the pressure gradients. Across the width of the slab (the y direction), the water surface does not vary, so there are no pressure variations in that direction. Across the depth (the z direction), the pressure varies because of gravity, but that variation is just sufficient to prevent the slab from sinking. We care about only the pressure difference across the length. This difference depends on the height of the crest, ξ , and is $\Delta p \sim \rho g \xi$ (see Figure 6.6). This pressure difference acts on a cross-section with area $A \sim w/k$, to produce a force

$$F \sim \underbrace{w/k}_{\text{area}} \times \underbrace{\rho g \xi}_{\Delta p} = \rho g w \xi / k. \quad (6.15)$$

The slab (Figure 6.7) has mass

$$m = \rho \times \underbrace{w/k^2}_{\text{volume}}, \quad (6.16)$$

so this force produces an acceleration

$$a_{\text{slab}} \sim \underbrace{\frac{\rho g w \xi}{k}}_{\text{force}} / \underbrace{\frac{\rho w}{k^2}}_{\text{mass}} = g \xi k. \quad (6.17)$$

The factor of g says that the gravity produces the acceleration. Full gravitational acceleration is reduced by the dimensionless factor ξk , which is roughly the slope of the waves.

From the acceleration of the slab, we can calculate the acceleration of the surface. If the slab moves a distance x , it sweeps out a volume of water $V \sim xA$. This water moves under the trough, and forces the surface upward a distance V/A_{top} . Because $A_{\text{top}} \sim A$ (both are $\sim w/k$), the surface moves the same distance, x , that the slab moves. Therefore, the acceleration of the slab, a_{slab} , given in (6.17), is equal to the acceleration a of the surface:

$$a \sim a_{\text{slab}} \sim g \xi k. \quad (6.18)$$

As we discover in Section 6.3, this equivalence of slab and surface acceleration does not hold in shallow water, where the bottom at depth h cuts off the slab before $1/k$.

With this slab argument, we hope to obtain the deep-water dispersion relation that we derived by dimensional analysis in (6.14). That equation contains frequency; the acceleration relation (6.18), however, does not. So we **massage** it until ω appears. The acceleration relation contains a and ξ , whereas the dispersion relation does not. We therefore manufacture an alternative expression for the acceleration. With luck, the expression will contain ω^2 and also ξ , so that we can cancel ξ and insert ω^2 . In simple harmonic motion (springs!), acceleration is $a \sim \omega^2 \xi$, where ξ is the amplitude. In waves, which behave like springs, we can use the same expression for a . We can also derive it for waves: In time $\tau \sim 1/\omega$, the surface moves a distance $d \sim \xi$, so $a/\omega^2 \sim \xi$, or $a \sim \omega^2 \xi$. Then (6.18) becomes

$$\underbrace{\omega^2 \xi}_a \sim g \xi k, \quad (6.19)$$

or

$$\omega^2 \sim gk, \quad (6.20)$$

which is the dispersion relation (6.14).

This equation, as usual, lacks the dimensionless constant, which we assume is unity. Fortunately, an exact calculation confirms this lucky guess. Our luck suggests that our procedures for choosing how to measure the lengths were reasonable. We made two approximations:

- We replaced an exponentially falling variation in velocity potential by a step function with size equal to the length scale of the exponential decay.
- By taking the length of the slab to be $1/k$ instead of π/k , we used only 1 radian of the cycle as the characteristic length, instead of using a half cycle.

Both approximations are usually accurate in order-of-magnitude calculations. Rarely, however, you get killed by a factor of $(2\pi)^6$, and wish that you had used a full cycle instead of only 1 radian.

The derivation that resulted in (6.20) analyzed the motion of the slab using forces. We can also derive the dispersion relation using energy, by balancing kinetic and potential energy as we did for the Bohr atom in Section 4.1.2. To make a wavy surface, we require energy, as shown in Figure 6.8. The crest rises a characteristic height ξ above the zero of potential, which is the level surface. The volume of water moved upward is $\xi w/k$. So the potential energy is

$$\text{PE}_{\text{gravity}} \sim \underbrace{\rho \xi w/k}_m \times g \xi \sim \rho g w \xi^2/k. \quad (6.21)$$

The kinetic energy is contained in the sideways motion of the slab and in the upward motion of the water pushed by the slab. Because slab and surface move at approximately the same speed, the sideways and upward motions contribute similar energies. We evaluate the energy for just the sideways motion. Because we are ignoring constants, such as 2, we do not need to compute the energy contributed by the upward motion. The surface, and therefore the slab, moves a distance ξ in a time $1/\omega$, so its velocity is $\omega \xi$. The kinetic energy is

$$\text{KE}_{\text{deep}} \sim \underbrace{\rho w/k^2}_{m_{\text{slab}}} \times \underbrace{\omega^2 \xi^2}_{v^2} \sim \rho \omega^2 \xi^2 w/k^2. \quad (6.22)$$

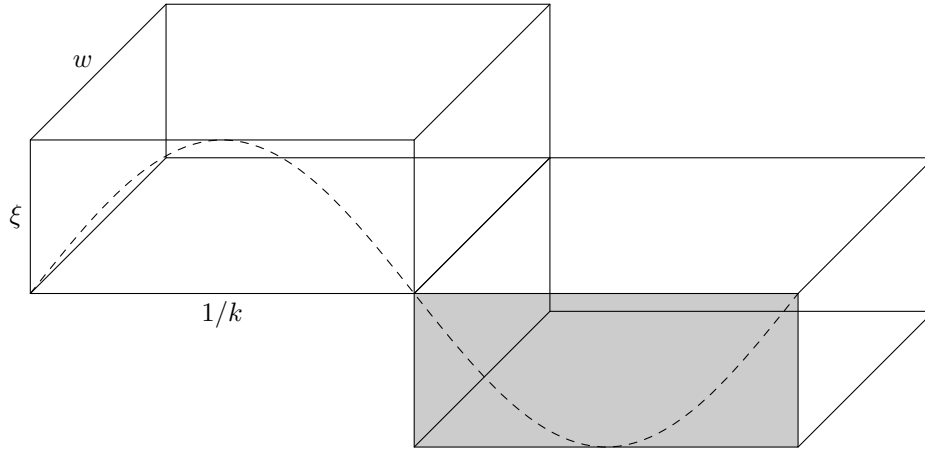


Figure 6.8. Energy required to make a wave. The shaded block of water has to be lifted up (and over, but sliding over takes no energy) a height $\sim \xi$. As explained in the text, water does not actually move from trough to crest, but assuming that it does makes a convenient picture, and does not affect the energy calculation.

We balance this energy against the potential energy (6.21):

$$\underbrace{\rho \omega^2 \xi^2 w / k^2}_{KE} \sim \underbrace{\rho g w \xi^2 / k}_{PE}. \quad (6.23)$$

Canceling the common factor of $\rho w \xi^2$, we find, once again, that

$$\omega^2 \sim gk. \quad (6.24)$$

The energy method agrees with the force method, as it should, because energy can be derived from force (by integration). The energy derivation tell us what the dimensionless group Π_2 means:

$$\Pi_2 \sim \frac{\text{kinetic energy in slab}}{\text{gravitational potential energy}} \sim \frac{\omega^2}{gk}. \quad (6.25)$$

The gravity-wave dispersion relation, $\omega^2 = gk$, is equivalent to $\Pi_2 \sim 1$, or to the assertion that kinetic and gravitational potential energy are comparable in wave motion. This rough equality is not a surprise for waves, because waves are like springs; in spring motion, kinetic and potential energies have equal averages (which is a consequence of the virial theorem, which we discussed in Section 4.2.1).

We have now derived the same dispersion relation in three ways: with dimensional analysis, and with two physical arguments applied to a simple model. Using multiple methods increases our confidence not only in our result, but also in our methods. We gain confidence in our methods of dimensional analysis, and in our slab model for waves. Were we to study nonlinear waves, for example, where we cannot assume that the wave height is infinitesimal, we would be able to use our techniques and model with more confidence.

Now that we believe the dispersion relation, let's study its consequences. We first determine the phase and group velocities. The crests move at the phase velocity: $v_{\text{ph}} = w/k$. For deep-water gravity waves, this velocity becomes

$$v_{\text{ph}} = \sqrt{\frac{g}{k}}, \quad (6.26)$$

or, using the dispersion relation to replace k by ω ,

$$v_{\text{ph}} = \frac{g}{\omega}. \quad (6.27)$$

In an infinite, single-frequency wave train, the crests and troughs move at this speed. However, a finite wave train contains a mixture of frequencies, and the various frequencies move at different speeds, as given by (6.27): Deep water is **dispersive**. Dispersion makes a finite wave train travel with the group velocity, given by $v_g = \partial w / \partial k$, as explained in *Fluid Dynamics for Physicists* [13]. Within a wave train, the crests move at the phase velocity, which can be different from the group velocity, shrinking and growing to fit under the envelope of the wave train. The group velocity is

$$v_g = \frac{\partial}{\partial k} \sqrt{gk} = \frac{1}{2} \sqrt{\frac{g}{k}} = \frac{1}{2} v_{\text{ph}}. \quad (6.28)$$

The group velocity is one-half of the phase velocity. The everyday consequence is that ship wakes trail a ship. A ship, moving with velocity v , creates gravity waves with $v_{\text{ph}} = v$. (Wind also generates waves, and the wind speed corresponds roughly to the phase velocity, as we discover shortly.) The waves combine to produce wave trains that propagate forward with the group velocity, which is only one-half of the velocity of the ship. From the ship's point of view, these gravity waves travel backward. In fact, they form a wedge, and the opening angle of the wedge depends on the factor of $1/2$.

6.2.1.1 SURFING. Let's apply the dispersion relation to surfing. Following one winter storm reported in the *Los Angeles Times*—the kind of storm that brings cries of “Surf's up”—waves arrived at Los Angeles beaches roughly every 18 s. How fast were the storm winds that generated the waves? Wind pushes crests, as long as they move more slowly than the wind. After a long-enough push, the crests move with nearly the wind speed. Therefore, the phase velocity of the waves is an accurate approximation to the wind speed. The phase velocity, from (6.27), is g/ω . In terms of the wave period T , the velocity is $v_{\text{ph}} = gT/2\pi$, so

$$v_{\text{wind}} \sim v_{\text{ph}} \sim \frac{\overbrace{10 \text{ m s}^{-2}}^g \times \overbrace{18 \text{ s}}^T}{2 \times 3} \sim 30 \text{ m s}^{-1}. \quad (6.29)$$

In units more familiar to Americans, this wind speed is roughly 60 mph—a hefty storm. The wavelength is given by

$$\lambda = v_{\text{ph}} T \sim 30 \text{ m s}^{-1} \times 18 \text{ s} \sim 500 \text{ m}. \quad (6.30)$$

The crests are separated by half a kilometer. They bunch up near shore, because they feel the bottom; this bunching is a consequence of the shallow water dispersion relation, which we discuss in Section 6.3.1.

In this same storm, the following day's waves arrived more frequently, at 17 s intervals. Let's first check that this decrease in period is reasonable. This precaution is a simple **sanity check**. If our theory is wrong about a physical effect as fundamental as this sign—whether the period should decrease or increase—then we have certainly neglected important physics. Fortunately, our theory passes the test. Group velocity varies as $1/\omega$, which is proportional to the period. The storm winds generate waves of different wavelengths and periods, and the different wavelengths sort themselves in the trip from the far ocean to Los Angeles. Longer-period waves move faster, so the 18 s waves should arrive before the 17 s waves, and they did. We can extract a surprising piece of information from the decline in the interval; we can determine the distance to the storm. In their long journey, the 18 s waves raced ahead 1 day. The ratio of speeds is

$$\frac{\text{velocity}(18 \text{ s waves})}{\text{velocity}(17 \text{ s waves})} = \frac{18}{17} = 1 + \frac{1}{17}. \quad (6.31)$$

so the race must have lasted roughly $t \sim 17 \text{ days} \sim 1.5 \cdot 10^6 \text{ s}$. The wave train moves at the group velocity, $v_g = v_{ph}/2 \sim 15 \text{ m s}^{-1}$, so the storm distance was $d \sim tv_g \sim 2 \cdot 10^4 \text{ km}$, or roughly halfway around the world (Figure 6.9)!

6.2.1.2 SPEEDBOATING. Our next application of the dispersion relation is to speedboating: How fast can a boat travel? We exclude hydroplaning boats from our analysis (even though some speedboats can hydroplane). Longer boats generally move faster than shorter boats, so it is likely that the length of the boat, l , determines the top speed. The density of water might matter. From only v (the speed), ρ , and l , however, we cannot form any dimensionless groups. So we need to add at least one more variable. Viscosity is irrelevant because the Reynolds' number (Section 3.3) for boat travel is gigantic. Even for a small boat of length 5 m, creeping along at 2 m s^{-1} ,

$$Re \sim \frac{500 \text{ cm} \times 200 \text{ cm s}^{-1}}{10^{-2} \text{ cm}^2 \text{ s}^{-1}} \sim 10^7. \quad (6.32)$$

At such a huge Reynolds' number, the flow is turbulent, and independent of viscosity. Surface tension is also irrelevant, because boats are much longer than a ripple wavelength (roughly 1 cm). Our search for new variables is not meeting with success. Perhaps gravity is relevant. From v , ρ , g , and l , we can form one dimensionless group (four variables, three dimensions), which is v^2/gl , also called the **Froude number**:

$$Fr \equiv \frac{v^2}{gl}. \quad (6.33)$$

The critical Froude number, which determines the maximum boat speed, is a dimensionless constant. As usual, we assume that the constant is unity:

$$v \sim \sqrt{gl}. \quad (6.34)$$

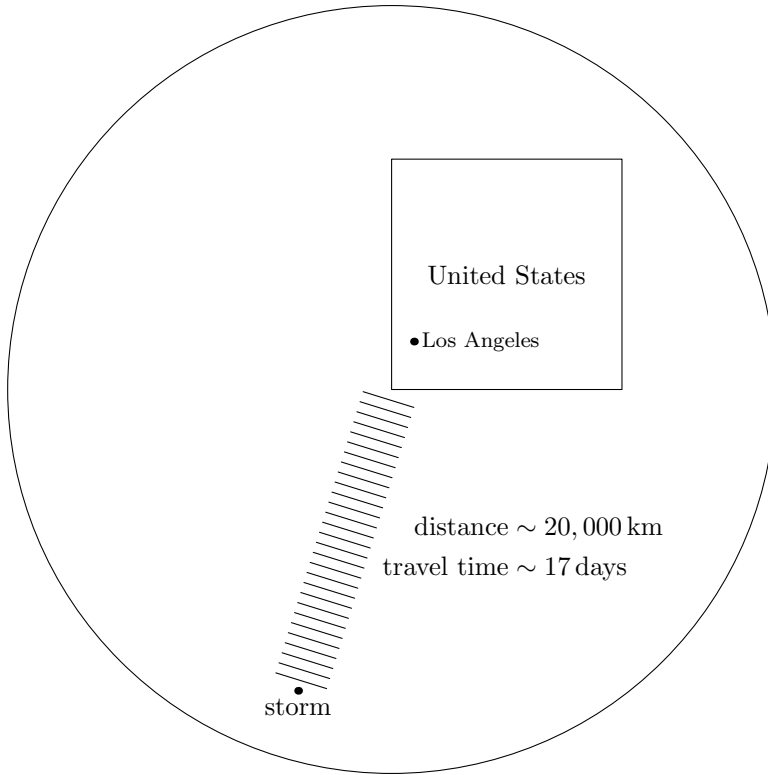


Figure 6.9. Map of the actual world (roughly drawn). A storm, almost at the south pole, generates waves that propagate outward (like ripples in a pond). Some waves make it to Los Angeles, to gladden the hearts of surfers and textbook writers.

A rabbit has just jumped out of our hat. What mechanism justifies this formula? Consider what happens when a boat plows through water, and follow the waves. The moving boat generates waves (the wake), and it rides on one of those waves. Let's analyze this bow wave: It is a gravity wave with $v_{\text{ph}} \sim v_{\text{boat}}$. Because $v_{\text{ph}}^2 = \omega^2/k^2$, the dispersion relation tells us that

$$v_{\text{boat}}^2 \sim \frac{\omega^2}{k^2} = \frac{g}{k} = g\lambda, \quad (6.35)$$

where $\lambda \equiv 1/k = \lambda/2\pi$. So the wavelength of the waves is $\sim v_{\text{boat}}^2/g$. The other length in this problem is the boat length; so we can interpret the Froude number (6.34) as

$$Fr = \frac{v_{\text{boat}}^2/g}{l} \sim \frac{\text{wavelength of bow wave}}{\text{length of boat}}. \quad (6.36)$$

Why is $Fr \sim 1$ the critical number? Interesting—and often difficult—physics occurs when a dimensionless number is near unity. In this case, the physics is as follows. The wave height changes significantly in a distance λ ; if the boat's length l is comparable to λ , then the boat rides on its own wave and tilts upward. It then presents a large cross-section to

the water, and the drag becomes huge.² So the top speed is given by

$$v_{\text{boat}} \sim \sqrt{gl}. \quad (6.37)$$

For a small motorboat, with length $l \sim 5$ m, this speed is roughly 7 m s^{-1} , or 15 mph. Boats can go faster than the nominal top speed given in (6.37), but it takes plenty of power to fight the drag, which is why police speedboats have huge engines.

We meet the Froude number in surprising places. It determines, for example, the speed at which an animal's gait changes from a walk to a trot or, for animals that do not trot, to a run. Here, it determines maximum boating speed. The Froude number is, in a sense, the ratio of potential energy to kinetic energy, as we can see by massaging v^2/gl :

$$Fr = \frac{v^2}{gl} = \frac{mv^2}{mgl} \sim \frac{\text{kinetic energy}}{\text{potential energy}}. \quad (6.38)$$

In this case, the length l is a horizontal length, so gl is not a potential energy. In the Froude number for walking speed, l is leg length, and gl is a potential energy. Then, for a human with leg length $l \sim 1$ m, the condition $Fr \sim 1$ implies that $v \sim 3 \text{ m s}^{-1}$ or 6 mph; this speed is a rough estimate for the top speed for a race walker (the world record for men's race walking is held by Bernado Segura of Mexico, who walked 20 km in 1:17:25.6, for a speed of 4.31 m s^{-1}).

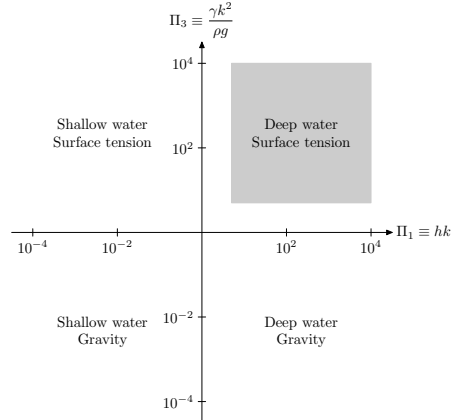


Figure 6.10. Deep-water-ripple portion of the world of waves.

6.2.2 Ripples on deep water

For small wavelengths (large k), surface tension provides the restoring force: We are now studying the shaded region in Figure 6.10. If surface tension, rather than gravity, provides the restoring force, then g drops out of the final dispersion relation. We could argue, as

2. Catamarans and hydrofoils skim the water, so this kind of drag does not limit their speed. The hydrofoil makes a much quicker trip across the English channel than the ferry makes, even though the hydrofoil is much shorter.

we did in Section 6.2.1, that g shows up in the dimensionless group $\Pi_3 \equiv \gamma k^2 / \rho g$, so Π_3 is irrelevant. In that argument, however, lies infanticide. It throws out the variable that determines the restoring force: surface tension. To retrieve the baby from the bathwater, we do not throw out $\gamma k^2 / \rho g$ directly. We instead choose the form of f_{deep} so that gravity, represented by g , vanishes from the dispersion relation.

The deep-water dispersion relation (6.12) contains one power of g in front. The argument of f_{deep} also contains one power of g , in the denominator. To cancel g , we choose f_{deep} to have the form $f_{\text{deep}}(x) \sim x$. Then $\omega^2 \sim \frac{\gamma k^3}{\rho}$. By luck, the dimensionless constant is unity, which we would have assumed anyway, so

$$\omega^2 = \frac{\gamma k^3}{\rho}. \quad (6.39)$$

Let's try to derive this relation (up to a dimensionless constant) using the slab argument.

In the slab picture, we replace gravitational by surface-tension energy, and again balance potential and kinetic energies. The surface of the water is like a rubber sheet. A wave stretches the sheet and creates area. This stretching requires energy. To estimate the energy, we first estimate the extra area that a wave of amplitude ξ and wavenumber k creates. The extra area depends on the extra length in a sine wave compared to a flat line. The typical slope in the sine wave $\xi \sin kx$ is ξk . Instead of integrating to find the arc length, we approximate the curve as a straight line with slope ξk (Figure 6.11). Relative to the level line, the tilted line is longer by a factor $1 + (\xi k)^2$, for small slopes. As before, we study

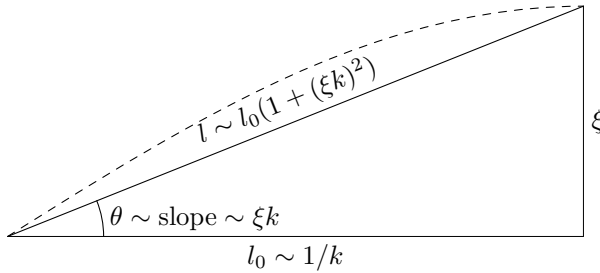


Figure 6.11. Approximating a sine wave by a straight line. We want to compute the arc length of the sine wave so that we can compute the extra surface area required to make a wave. We're too lazy to integrate to find arc length. ("We" here means "the authors"—our readers may be more diligent; the purpose of this text is to correct that habit.) So we replace the curve by a straight line with the same characteristic slope. The characteristic slope of the sine wave is ξk (which happens to be the exact slope at the origin): In a distance of roughly 1 rad—which is a length $1/k$ —it rises a height ξ . The length of the hypotenuse is $l \sim l_0 / \cos \theta$, which is roughly $l_0(1 + (\xi k)^2)$ for small slopes ξk .

a piece of a wave, with characteristic length $1/k$ in the x direction, and width w in the y direction. The extra area is

$$\Delta A \sim \underbrace{w/k}_{\text{level area}} \times \underbrace{(\xi k)^2}_{\text{fractional increase}} \sim w \xi^2 k. \quad (6.40)$$

The potential energy stored in this extra surface is

$$\text{PE}_{\text{ripple}} \sim \gamma \Delta A \sim \gamma w \xi^2 k. \quad (6.41)$$

The kinetic energy in the slab is the same as it is for gravity waves, which is given in (6.22):

$$\text{KE} \sim \rho \omega^2 \xi^2 w / k^2. \quad (6.42)$$

We balance the energies,

$$\underbrace{\rho \omega^2 \xi^2 w / k^2}_{\text{KE}} \sim \underbrace{\gamma w \xi^2 k}_{\text{PE}}, \quad (6.43)$$

and find that

$$\omega^2 \sim \gamma k^3 / \rho. \quad (6.44)$$

This dispersion relation agrees with the one that we found in (6.39) using dimensional analysis. For deep-water gravity waves, we used energy and force arguments to rederive the dispersion relation. For ripples, we have just worked out the force argument; you are invited to work out the corresponding force argument.

We have already interpreted the first two dimensionless groups: Π_1 is the dimensionless depth, and Π_2 is ratio of kinetic energy to gravitational potential energy. We described Π_3 as a group that compares the effects of surface tension and gravity. Having computed potential energy for gravity waves in (6.21), and for ripples in (6.41), we can make the comparison more precise:

$$\begin{aligned} \Pi_3 &\sim \frac{\text{potential energy in a ripple}}{\text{potential energy in a gravity wave}} \\ &\sim \frac{\gamma w \xi^2 k}{\rho g w \xi^2 / k} \\ &\sim \frac{\gamma k^2}{\rho g}. \end{aligned} \quad (6.45)$$

Alternatively, Π_3 compares $\gamma k^2 / \rho$ with g :

$$\Pi_3 \equiv \frac{\gamma k^2 / \rho}{g}. \quad (6.46)$$

This form of Π_3 may seem like a trivial revision of $\gamma k^2 / \rho g$. However, it suggests an interpretation of surface tension: that surface tension acts like an effective gravitational field with strength

$$g_{\text{surface tension}} = \gamma k^2 / \rho, \quad (6.47)$$

In a balloon, the surface tension of the rubber implies a higher pressure inside than outside. Similarly, in wave, the water skin implies a higher pressure underneath the crest, which is curved like a balloon; and a lower pressure under the trough, which is curved opposite to a balloon. This pressure difference is what a gravitational field with strength $g_{\text{surface tension}}$. We have met this idea of effective gravity already, when we studied marbles falling in corn syrup (Section 3.3.2); in that problem, we replaced g by an effective g that included the effect of buoyancy.

If we replace g in the gravity-wave potential energy (6.21) with this effective g , we get the ripple potential energy (6.41):

$$\underbrace{\rho g w \xi^2 / k}_{PE(\text{gravity wave})} \xrightarrow{g \rightarrow \gamma k^2 / \rho} \underbrace{\gamma w \xi^2 k}_{PE(\text{ripple})}. \quad (6.48)$$

The left side becomes the right side after we make the substitution above the arrow. If we make the same replacement in the gravity-wave dispersion relation (6.14), we get the ripple dispersion relation (6.39):

$$\omega^2 = gk \xrightarrow{g \rightarrow \gamma k^2 / \rho} \omega^2 = \frac{\gamma k^3}{\rho}. \quad (6.49)$$

The interpretation of surface tension as effective gravity is useful when we combine our solutions for gravity waves and for ripples, in Sections 6.2.3 and 6.3.3. We could have reversed the analysis and interpreted gravity as effective surface tension; however, gravity is the more familiar force, so we chose to use effective gravity.

As we did for gravity waves, we now use the dispersion relation to calculate phase and group velocities. The phase velocity is

$$v_{\text{ph}} \equiv \frac{\omega}{k} = \sqrt{\frac{\gamma k}{\rho}}, \quad (6.50)$$

and the group velocity is

$$v_{\text{g}} \equiv \frac{\partial \omega}{\partial k} = \frac{3}{2} v_{\text{ph}}. \quad (6.51)$$

The factor of $3/2$ is a consequence of the form of the dispersion relation: $\omega \propto k^{3/2}$; for gravity waves, $\omega \propto k^{1/2}$, and the corresponding factor is $1/2$. In contrast to deep-water waves, a train of ripples moves *faster* than the phase velocity. So, ripples steam ahead of a boat, whereas gravity waves trail behind.

Let's work out speeds for typical ripples, such as the ripples that we create by dropping a pebble into a pond. These ripples have wavelength $\lambda \sim 1$ cm, and therefore have wavenumber $k = 2\pi/\lambda \sim 6 \text{ cm}^{-1}$. The surface tension of water, which we estimated in Section 4.3.3, is $\gamma \sim 72 \text{ erg cm}^{-2}$. So the phase velocity, given by (6.50), is

$$v_{\text{ph}} = \left(\frac{\overbrace{72 \text{ erg cm}^{-2}}^{\gamma} \times \overbrace{6 \text{ cm}^{-1}}^k}{\underbrace{1 \text{ g cm}^{-3}}_{\rho}} \right)^{1/2} \sim 21 \text{ cm s}^{-1}. \quad (6.52)$$

The group velocity is $v_{\text{g}} \sim 30 \text{ cm s}^{-1}$. This wavelength of 1 cm is roughly the longest wavelength that still qualifies as a ripple, as shown in Figure 6.2; the third dimensionless group, which distinguishes ripples from gravity waves, has value

$$\Pi_3 \equiv \frac{\gamma k^2}{\rho g} \sim \frac{\overbrace{72 \text{ erg cm}^{-2}}^{\gamma} \times \overbrace{36 \text{ cm}^{-2}}^{k^2}}{\underbrace{1 \text{ g cm}^{-3}}_{\rho} \times \underbrace{1000 \text{ cm s}^{-2}}_g} \sim 2.6. \quad (6.53)$$

With a slightly smaller k , the value of Π_3 would slide into the gray zone where $\Pi_3 \approx 1$. If k were yet smaller, the waves would be gravity waves. Other ripples, with a larger k , have a shorter wavelength, and therefore move faster: 21 cm s^{-1} is roughly the minimum phase velocity for ripples. This minimum speed explains why we see mostly $\lambda \sim 1 \text{ cm}$ ripples when we drop a pebble in a pond. The pebble excites ripples of various wavelengths; the shorter ones propagate faster and the 1 cm ones straggle, so we see the stragglers clearly, without admixture of other ripples.

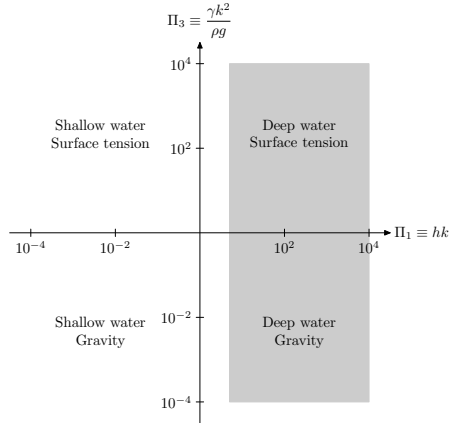


Figure 6.12. Deep-water portion of the world of waves (same as Figure 6.3).

6.2.3 Combining ripples and gravity waves on deep water

Having studied two corners of the puzzle—gravity waves and ripples in deep water—we connect the corners, and study the deep-water edge, as shown in Figure 6.12. The dispersion relations (6.14) and (6.39), for convenience restated here, are

$$\omega^2 = \begin{cases} gk, & \text{gravity waves;} \\ \gamma k^3 / \rho, & \text{ripples.} \end{cases} \quad (6.54)$$

We would like to combine the relations in these two extreme regimes, to produce a dispersion relation valid for gravity waves, for ripples, and for waves in between.

Both functional forms came from the same physical argument of balancing kinetic and potential energies. The difference was the source of the potential energy: gravity or surface tension. Gravity and surface tension are active in both regimes. On the top portion of the world of waves (Figure 6.19), surface tension dominates gravity; on the bottom portion (Figure 6.18), gravity dominates surface tension. You might therefore guess that, in the intermediate region, the two contributions to the potential energy simply add, and that the combination dispersion relation is just the sum of the two extremes:

$$\omega^2 = gk + \gamma k^3 / \rho. \quad (6.55)$$

Your guess would be correct (which is why we used an equality); when in doubt, try the simplest solution.

We can restate our reasoning using the effective gravitational field produced by surface tension, derived in (6.47). The two sources of gravity—real and effective—simply add, to make

$$g_{\text{total}} = g + g_{\text{surface tension}} = g + \frac{\gamma k^2}{\rho}. \quad (6.56)$$

When we replace g in $\omega^2 = gk$ with g_{total} , we rederive the deep-water dispersion relation (6.55):

$$\omega^2 = \left(g + \frac{\gamma k^2}{\rho} \right) k = gk + \gamma k^3 / \rho. \quad (6.57)$$

With this dispersion relation, we can calculate wave speeds for all wavelengths or wavenumbers. The phase velocity is

$$v_{\text{ph}} \equiv \frac{\omega}{k} = \sqrt{\frac{\gamma k}{\rho} + \frac{g}{k}}. \quad (6.58)$$

As a function of wavenumber, the two terms in the square root compete to increase the speed. The surface-tension term wins at high wavenumber; the gravity term wins at low wavenumber. So there is an intermediate, minimum-speed wavenumber, k_0 , which we can estimate by balancing the surface tension and gravity contributions:

$$\frac{\gamma k_0}{\rho} \sim \frac{g}{k_0}. \quad (6.59)$$

This computation is another example of order-of-magnitude minimization, which we introduced in Section 4.1.2 to calculate the Bohr radius. This minimum-speed wavenumber is

$$k_0 \sim \sqrt{\frac{\rho g}{\gamma}}. \quad (6.60)$$

At this wavenumber, $\Pi_3 = 1$: These waves lie just on the border between ripples and gravity waves. Their phase speed is

$$v_0 \sim \sqrt{2g/k_0} \sim \left(\frac{4\gamma g}{\rho} \right)^{1/4}. \quad (6.61)$$

In water, the critical wavenumber is $k_0 \sim 4 \text{ cm}^{-1}$, so the critical wavelength is $\lambda_0 \sim 1.5 \text{ cm}$; the speed is

$$v_0 \sim 23 \text{ cm s}^{-1}. \quad (6.62)$$

We derived the speed (6.61) dishonestly. Instead of using the maximum–minimum methods of calculus, we balanced the two contributions. A calculus calculation would agree with our minimum phase velocity. A tedious calculus calculation shows that the minimum group velocity is

$$v_g \approx 17.7 \text{ cm s}^{-1}. \quad (6.63)$$

Let's do the minimizations honestly. The calculation is not too messy if it's done properly, and the proper method is useful in many physical maximum–minimum problems. We

first find the minimum of the phase velocity (6.58). That equation contains constants— ρ , γ , and g —which we have to carry through all the differentiations. To simplify the manipulations, we instead choose a convenient set of units, in which

$$\rho = \gamma = g = 1. \quad (6.64)$$

Our treatment of waves uses three basic dimensions: mass, length, and time. Choosing three constants equal to unity is the most freedom that we have; it is equivalent to choosing canonical mass, length, and time, and thereby making all quantities dimensionless.

We restore the constants at the end of the minimization. In addition to constants, the phase velocity also contains a square root. So we minimize v_{ph}^2 , which in our new unit system is

$$v_{\text{ph}}^2 = k + \frac{1}{k}. \quad (6.65)$$

For this minimization, we do not need calculus. The two terms are both positive, so we may apply the arithmetic-mean–geometric-mean inequality (affectionately known as AM–GM) to k and $1/k$. The inequality states that, for positive a and b ,

$$\underbrace{(a+b)/2}_{\text{AM}} \geq \underbrace{\sqrt{ab}}_{\text{GM}}, \quad (6.66)$$

with equality when $a = b$. Figure 6.13 is a geometric proof of this inequality. You are invited to convince yourself, just for fun, that the figure is a proof. Using $a = k$ and $b = 1/k$, we find that the geometric mean is unity, so the arithmetic mean is ≥ 1 . Therefore,

$$k + \frac{1}{k} \geq 2, \quad (6.67)$$

with equality when $k = 1/k$, or when $k = 1$. At this wavenumber, the phase velocity is $\sqrt{2}$. In this unit system, the dispersion relation is

$$\omega = \sqrt{k^3 + k}, \quad (6.68)$$

and the group velocity is

$$v_g = \frac{\partial}{\partial k} \sqrt{k^3 + k}, \quad (6.69)$$

which is

$$v_g = \frac{1}{2} \frac{3k^2 + 1}{\sqrt{k^3 + k}}. \quad (6.70)$$

At $k = 1$, the group velocity is also $\sqrt{2}$: These borderline waves have equal phase and group velocity.

To convert $k = 1$ back to our usual units, we multiply it by 1, where we form the 1 using a product of ρ , γ , and g . How can we make a length from ρ , γ , and g ? The form of the result (6.60) tells us that $\sqrt{\rho g / \gamma}$ has units of $[\text{L}]^{-1}$. So $k = 1$ really means $k = 1 \times \sqrt{\rho g / \gamma}$, which is just what we found in (6.60), but now we know that the missing constant is 1.

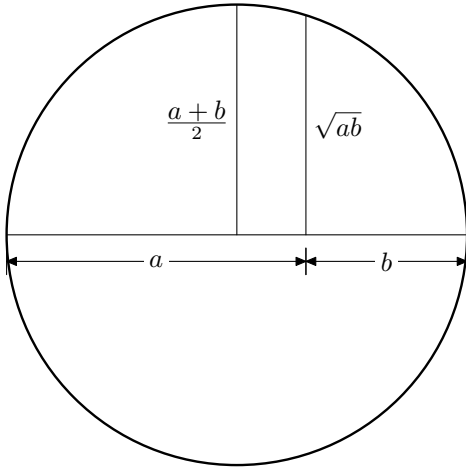


Figure 6.13. *Proof without words: A geometric proof of the arithmetic-mean–geometric-mean inequality. We leave to you the fun of translating the picture into a traditional proof. (Hint: Draw a triangle.)*

The minimum group velocity is more complicated. We minimize v_g^2 , to get rid of the square root. The derivative is

$$\frac{\partial}{\partial k} \underbrace{\frac{9k^4 + 6k^2 + 1}{k^3 + k}}_{v_g^2} = \frac{(3k^2 + 1)(3k^4 + 6k^2 - 1)}{(k^3 + k)^2}. \quad (6.71)$$

Equating this derivative to zero, we solve $3k^4 + 6k^2 - 1 = 0$, which is a quadratic in k^2 , and has positive solution

$$k_1 = \sqrt{-1 + \sqrt{4/3}} \sim 0.393. \quad (6.72)$$

At this k , the group velocity is

$$v_g(k_1) \approx 1.086. \quad (6.73)$$

In more usual units, this minimum velocity is

$$v_g \approx 1.086 \left(\frac{\gamma g}{\rho} \right)^{1/4}. \quad (6.74)$$

When we put in the density and surface tension of water, we get a minimum group velocity of 17.7 cm s^{-1} , as we claimed in (6.63).

After dropping a pebble in a pond, you see a still circle surrounding the drop point. Then, the circle expands at the minimum group velocity given in (6.63). If you do not have a pond handy, you can try the experiment in your kitchen sink: Fill it with water, and drop in a coin. The existence of a minimum phase velocity, given in (6.62), is useful for bugs that walk on water. If they move slower than 23 cm s^{-1} , they generate no waves, and thereby reduce the energy cost of walking.

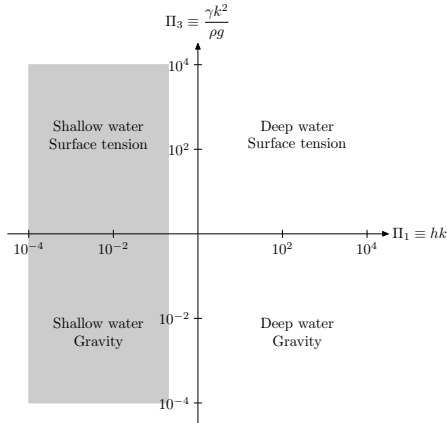


Figure 6.14. Shallow-water portion of the world of waves.

6.3 Shallow water

In shallow water, the height h , absent in the deep-water calculations, returns to complicate the set of relevant variables. We are now in the shaded region of Figure 6.14. With this extra length scale, we have too much freedom; using only dimensional analysis, we cannot deduce the shallow-water form of the magic function f in the dispersion relation (6.6). However, by modifying the slab argument, we can.

In deep water, the slab has depth $1/k$. In shallow water, however, where $h \ll 1/k$, the bottom of the ocean arrives before that depth. The shallow-water slab has depth h . Its length is still $1/k$, and its width is still w . However, the continuity argument changes. In deep water, where the slab has depth equal to length, the slab and surface move the same distance. In shallow water, with a slab thinner by hk , the surface moves more slowly than the slab, by the factor hk . With wave height ξ and frequency ω , the surface moves with velocity $\xi\omega$, so the slab moves (sideways) with velocity $v_{\text{slab}} \sim \xi\omega/hk$. The kinetic energy in the water is contained mostly in the slab, because the upward motion is much slower than the slab motion. This energy is

$$\text{KE}_{\text{shallow}} \sim \underbrace{\rho w h/k}_{\text{mass}} \times \underbrace{(\xi\omega/hk)^2}_{v^2} \sim \frac{\rho w \xi^2 \omega^2}{h k^3}. \quad (6.75)$$

We balance this energy against the potential energy, which we compute for the two limiting cases: ripples and gravity waves.

6.3.1 Gravity waves on shallow water

We first specialize to gravity waves—the shaded region in Figure 6.15, where water is shallow and wavelengths are long. These conditions include tidal waves and waves generated by undersea earthquakes. For gravity waves, we already computed the potential energy, in (6.21), and found that

$$\text{PE} \sim \rho g w \xi^2 / k. \quad (6.76)$$

This energy came from the distortion of the surface, and it is the same in shallow water (as long as the wave amplitude is small compared with the depth and wavelength). [The

dominant force (gravity or surface tension) determines the potential energy. As we see when we study shallow-water ripples, in Section 6.3.2, the water depth determines the kinetic energy.]

We balance this energy against the kinetic energy (6.75):

$$\underbrace{\frac{\rho w \xi^2 \omega^2}{h k^3}}_{KE} \sim \underbrace{\rho g w \xi^2 / k}_{PE}. \quad (6.77)$$

So

$$\omega^2 = g h k^2. \quad (6.78)$$

The equation has an equality, because, once again, the dimensionless constant is unity. So, for gravity waves on shallow water, the function f has the form

$$f_{\text{shallow}}(kh, \frac{\gamma k^2}{\rho g}) = kh. \quad (6.79)$$

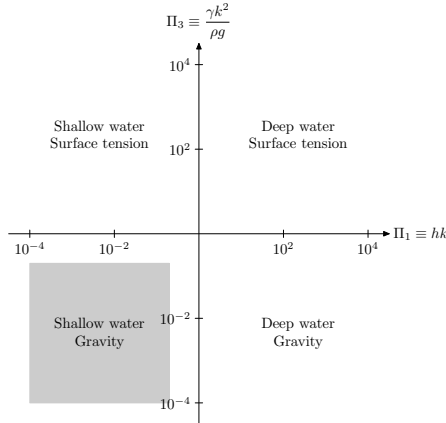


Figure 6.15. Shallow-water-gravity-wave portion of the world of waves.

The group and phase velocities are equal to each other, and are independent of frequency:

$$\begin{aligned} v_{\text{ph}} &= \frac{\omega}{k} = \sqrt{gh}, \\ v_{\text{g}} &= \frac{\partial \omega}{\partial k} = \sqrt{gh}. \end{aligned} \quad (6.80)$$

Shallow water is **nondispersive**: All frequencies move at the same velocity, so pulses composed of various frequencies propagate without smearing. Undersea earthquakes illustrate the danger in such unity. If there is an earthquake off the coast of Chile, and the ocean bottom drops, it generates a shallow-water wave, which travels without distortion to Japan. The wave speed is $v \sim \sqrt{3000 \text{ m} \times 10 \text{ m s}^{-2}} \sim 170 \text{ m s}^{-1}$: The wave can cross a 10^4 km ocean in 1 day. As it approaches shore, where the depth decreases, the wave slows, grows in amplitude, and becomes a large, destructive wave when it reaches land.

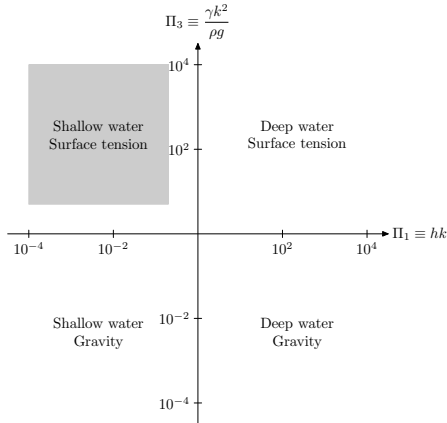


Figure 6.16. Shallow-wave-ripple portion of the world of waves.

6.3.2 Ripples on shallow water

Ripples on shallow water—the shaded region in Figure 6.16—are rare. They occur when raindrops land in a shallow rain puddle, one whose depth is less than 1 mm. Even then, only the longest-wavelength ripples, where $\lambda \sim 1$ cm, can feel the bottom of the puddle. What is the potential energy of the surface? We have already computed the potential energy, in (6.41). Although we derived that formula for deep water, the water depth does not affect the potential energy; the dominant force—here, surface tension—determines the potential energy. We equate the potential energy from (6.41) with the kinetic energy (6.75):

$$\underbrace{\frac{\rho w \xi^2 \omega^2}{h k^3}}_{KE} \sim \underbrace{\frac{w}{k} \gamma (k \xi)^2}_{PE}. \quad (6.81)$$

We then find that

$$\omega^2 \sim \frac{\gamma h k^4}{\rho}. \quad (6.82)$$

The phase velocity is

$$v_{\text{ph}} = \sqrt{\frac{\gamma h k^2}{\rho}}, \quad (6.83)$$

and the group velocity is $v_g = 2v_{\text{ph}}$ (the form of the dispersion relation is $\omega \propto k^2$). For $h \sim 1$ mm, this speed is

$$v \sim \left(\frac{72 \text{ erg cm}^{-2} \times 0.1 \text{ cm} \times 36 \text{ cm}^{-2}}{1 \text{ g cm}^{-3}} \right)^{1/2} \sim 16 \text{ cm s}^{-1}. \quad (6.84)$$

6.3.3 Combining ripples and gravity waves on shallow water

We have studied the two corners of the world of shallow-water waves: ripples and gravity waves. Now we connect the corners to make an edge: We study general shallow-water waves. This region of the world of waves is shaded in Figure 6.17. We can combine the dispersion

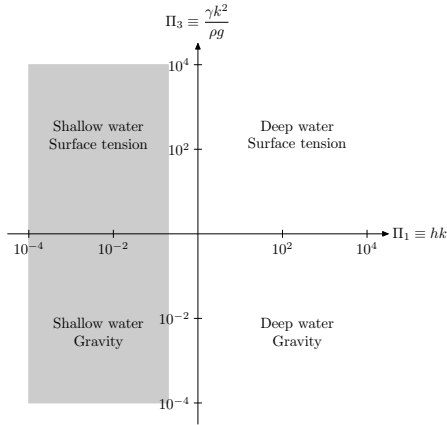


Figure 6.17. Shallow-water portion of the world of waves (same figure as Figure 6.14).

relations for ripples with that for gravity waves using two equivalent methods. We either can add the two extreme-case dispersion relations, (6.78) and (6.82); or, can use the effective gravitational field given in (6.56) in the gravity-wave dispersion relation (6.78). Either method produces

$$\omega^2 \sim k^2 \left(gh + \frac{\gamma h k^2}{\rho} \right). \quad (6.85)$$

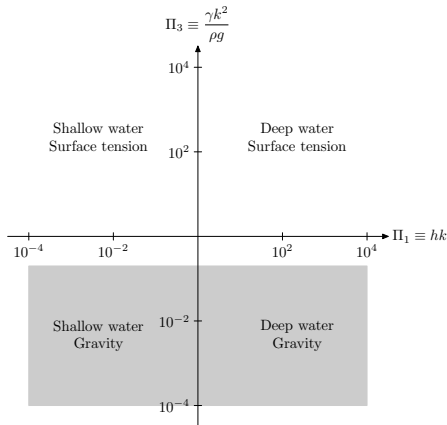


Figure 6.18. Gravity-wave portion of the world of waves.

6.4 Combining deep- and shallow-water gravity waves

We now examine the gravity-wave edge of the world, as shown in Figure 6.18. The deep- and shallow-water dispersion relations are, from (6.14) and (6.78):

$$\omega^2 = gk \times \begin{cases} 1, & \text{deep water;} \\ hk, & \text{shallow water.} \end{cases} \quad (6.86)$$

To interpolate between the two regimes, we want a function $f(hk)$ that asymptotes to 1 as $hk \rightarrow \infty$, and to hk as $hk \rightarrow 0$. One such simple function is $\tanh hk$, so we guess that the One True Gravity Wave Dispersion Relation is:

$$\omega^2 = gk \tanh hk. \quad (6.87)$$

Our guess is plausible, because $\tanh hk$ falls off exponentially as $h \rightarrow \infty$, in agreement with the argument based on Laplace's equation. In fact, our guess is correct.

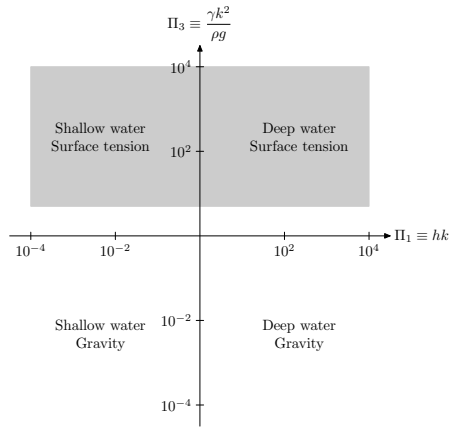


Figure 6.19. *Ripple portion of the world of waves.*

6.5 Combining deep- and shallow-water ripples

We now examine the final edge: ripples, in both shallow and deep water, as shown in Figure 6.19. In Section 6.4, we found that $\tanh kh$ did the yeoman work of interpolating between hk and 1, as hk went from 0 to ∞ (as the water went from shallow to deep). We guess that the same trick holds for ripples, because the Laplace-equation argument, which justified the $\tanh kh$, does not depend on the restoring force. The relevant dispersion relations are (6.82), for shallow water, and (6.39), for deep water:

$$\omega^2 = \begin{cases} \gamma k^3 / \rho, & \text{if } kh \gg 1; \\ \gamma h k^4 / \rho, & \text{if } kh \ll 1. \end{cases} \quad (6.88)$$

If we factor out $\gamma k^3 / \rho$, the necessary transformation becomes clear:

$$\omega^2 = \frac{\gamma k^3}{\rho} \times \begin{cases} 1, & \text{if } kh \gg 1; \\ hk, & \text{if } kh \ll 1. \end{cases} \quad (6.89)$$

This ripple result looks similar to the gravity-wave result (6.86), so we make the same replacement:

$$\begin{cases} 1, & \text{if } kh \gg 1, \\ hk, & \text{if } kh \ll 1, \end{cases} \quad \text{becomes} \quad \tanh kh. \quad (6.90)$$

Then we get the general ripple dispersion relation:

$$\omega^2 = \frac{\gamma k^3}{\rho} \tanh kh. \quad (6.91)$$

This dispersion relation does not have much practical interest, because, at the cost of greater complexity than the deep-water ripple dispersion relation (6.39), it adds coverage of only a rare case: ripples on ponds. We include it for completeness, so that we have treated all four edges of the world, in preparation for the grand combination coming up next.

6.6 Combining all the analyses

Now we can replace g with g_{total} from (6.56), to find the One True Dispersion Relation:

$$\omega^2 = (gk + \gamma k^3/\rho) \tanh kh. \quad (6.92)$$

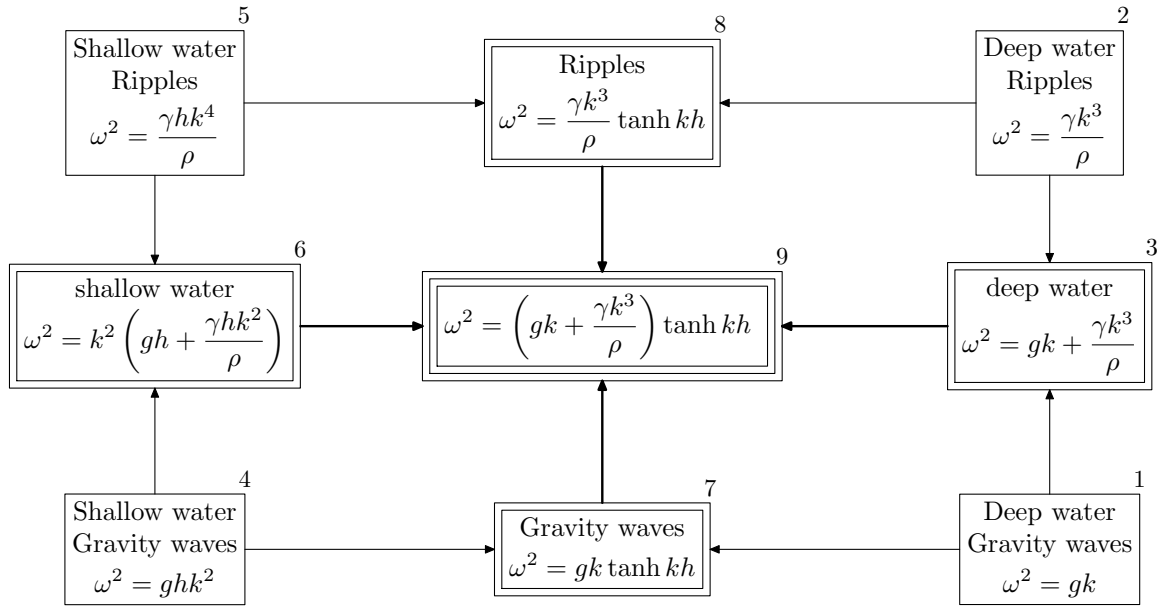


Figure 6.20. What we did. Each box represents a limit. The numbers next to the boxes mark the order in which we studied that limit. In the final step (9), we combined all the analyses into the superbox in the center, which contains the dispersion relation for all waves: gravity waves or ripples, shallow water or deep water. The arrows show how we combined smaller, more specialized corner boxes into the more general edge boxes (double ruled), and the edge regions into the universal center box (triple ruled).

6.7 What we did

We studied water waves by investigating dispersion relations. We mapped the world of waves, explored the corners, then the edges, and then assembled the results to form an understanding of the complex, complete solution. We can now draw the whole map, shown in Figure 6.20. Considering limiting cases, as we did, and then stitching them together, makes the analysis tractable, and, more important, comprehensible.

7 RETINAL ROD

7.1 Problem that the rod solves

Try the following experiment, which Hubel [23, p. 56–57] describes. Take a light meter and a newspaper to your favorite beach. Measure the light reflected from the big black letters. Go inside and measure the light reflected from the white space. You will find that the black letters outdoors reflect more light than the white space indoors. Yet to our eye, black is dark and white is light, no matter the background light.

The eye’s remarkable illumination independence arises because the eye computes contrast in the first steps of its image processing, as Shapley, Kaplan and Purpura [55] discuss. Reflectivity—the fraction of incident light that an object reflects—is an innate property of an object, determined by the energy levels of the electrons in the object and by the texture of the surface, rather than by the light level. The light reflected from an object is RI_b , where R is the reflectivity and I_b is the background-light intensity.

When you read indoors or at the beach, the light is bright enough that rods are useless and cones transduce the signal. The analysis in this paper is for the rod, because we know more detailed biochemistry for the rod than for the cone. For this newspaper example, pretend anyway that rods transduce the signal, and consider what one rod sees. As the eye sweeps (saccads) across the newspaper title, this rod receives light reflected from white space, from black print, and then from white space: The black print generates a (negative) flash. The size of the flash is $\propto I_b \Delta R$, where ΔR is the difference in reflectivity between black print and white space. If the rod’s flash sensitivity is $\propto I_b^{-1}$, then the signal that the rod generates as the black letter crosses it is $\propto \Delta R$ and is independent of I_b .

This I_b^{-1} flash sensitivity is the Weber–Fechner law, first proposed by Weber [60] and extended by Fechner [14]. It benefits the visual system in three ways. First, it uses the limited capacity of the optic nerve efficiently. By removing redundant information—the background light—the rod frees the optic nerve to carry a larger quantity of useful information: contrast. Second, it simplifies later stages of image processing: Those stages do not need to change their algorithms according to light level. Third, it increases the dynamic range of the rod. Suppose that the rod had constant sensitivity at all light levels. Then the same scene would produce much larger signals in bright than they would in low light. This needless variation would hinder the rod, which would have to encode both large and small signals with the same fractional accuracy. This task is difficult for small signals, for which is relatively large. The Weber–Fechner sensitivity avoids this problem, and allows the rod to accurately represent scenes in low and in bright light; it extends the dynamic range of the rod. An actual rod does not compute exact mathematical functions. However, to the extent that the rod’s flash response approximates I_b^{-1} , the rod gains these advantages. The experiments of

Matthews and coworkers [42] show that the rod's flash sensitivity approximates I_b^{-1} over a light intensity range of a factor of 100.

7.2 Scope of this analysis

This analysis is an approximate model for the flash sensitivity. With one caveat (described in Section 7.5), this model explains the Weber–Fechner law. Many workers have made detailed theoretical models of rod biochemistry; see, for example, the papers of Koutalos, Nakatani, and Yau [31], of Lamb and Pugh [37, 36], of Pugh and Lamb [48], Nikonov, Engheta, and Pugh [44], of Calvert and coworkers [5], of Tamura, Nakatani, and Yau [58], of Ichikawa [24], and of Ames [25]. (The paper by Ames describes a particularly interesting graphical method.) These models—whose predictions closely match experimental data—show that our understanding of rod biochemistry is nearly complete.

The accuracy comes at a price: numerical solution. The approximate model in this paper is simple enough that we can solve the equations analytically or, with feedback diagrams, pictorially. Furthermore, the approximate analysis, because it is simple, shows clearly what biochemical features are responsible for what properties of the response. The approximate analysis is therefore easier to generalize; these computational building blocks occur in other biochemical systems, such as protein synthesis regulation. The simplicity comes at the cost of accuracy. Both types of models—simple and inaccurate, and complicated and accurate—increase our understanding of the rod and of biochemical pathways.

The first part of this analysis (Section 7.3) estimates the response of a dark-adapted rod to a flash of light; it is a simplified presentation of the model of Lamb and Pugh [37]. The second part (Section 7.4) estimates the effect of steady background light. The third part (Section 7.5) combines the first two estimates to estimate the flash response as a function of background light level. Tamura, Nakatani, and Yau [58] describe a model that explains Weber–Fechner flash sensitivity in monkey rods; they evaluate the flash sensitivity by differentiating the steady-state response with respect to light intensity. That method is not always correct; the flash response so computed may be in error by a constant that depends on how rapidly the sensitivity varies with flicker rate. The model in this paper evaluates the flash sensitivity directly from the flash response. (Differentiating the steady-state response in this paper does give the correct scaling for the flash response. The result may be in error by a constant, but that error would be invisible to this model: For simplicity, this model ignores constants of proportionality.) This model uses data from salamander rods, the photoreceptors for which we have the most detailed biochemical information.

7.3 Flash response in darkness: The model of Lamb and Pugh

Lamb and Pugh's model, which they present in [37] and review in [36, 48], explains the first few hundred milliseconds of the response of a dark-adapted rod to a flash of light. It assumes that calcium concentration remains constant during the response, an assumption that fails after roughly 300 ms. This section reviews their model.

Figure 7.1 shows the reactions by which light lowers membrane current. The mathematical analysis follows the same sequence of reactions. At time $t = 0$, a light flash instantly activates S rhodopsin molecules:

$$R^* = S \text{ step}(t), \quad (7.1)$$

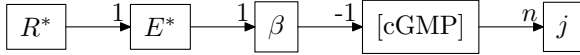


Figure 7.1. From light to current, without feedback. A flash of light activates rhodopsin molecules (R^*). Activated rhodopsin, via G-protein, activates phosphodiesterase molecules (increases E^*); the 1 on the arrow that connects R^* and E^* says that E^* is proportional to R^* . PDE* molecules hydrolyze cyclic GMP (cGMP) at a rate β ; the rate is proportional to E^* ; the linear proportionality is indicated on the figure with a 1 on the arrow between E^* and β . The -1 between β and [cGMP] means that (in the steady state) cyclic GMP concentration is inversely proportional to phosphodiesterase hydrolysis rate β . In the final stage of transduction, cyclic GMP molecules open channels in the membrane; the membrane current is proportional to the n th power of [cGMP].

where t is the time since the flash, and $\text{step}(t)$ is the step function. [Table 7.1 lists the symbols and parameters in this analysis.] The activated rhodopsin eventually inactivates, so the step function describes only the initial change in R^* level. The inactivation time constant is many seconds (as the experiments of Pepperberg and coworkers [46] show); by then, the constant-calcium approximation is already invalid. In the first few hundred milliseconds, when the calcium level is roughly constant, the step function is an accurate description.

Activated rhodopsin activates G-protein, which activates phosphodiesterase. The number of activated phosphodiesterase molecules increases linearly (it ramps):

$$E^* = S\nu_{\text{RE}}t, \quad (7.2)$$

where ν_{RE} is the rate at which one activated rhodopsin creates (via G-protein) activated phosphodiesterase. Activated phosphodiesterase hydrolyzes cGMP at a rate

$$\left(\frac{dc}{dt}\right)_{\text{hydrolysis}} = -\beta c = -(\beta_{\text{dark}} + \nu_{\text{EC}}E^*)c, \quad (7.3)$$

where c is the cGMP concentration, ν_{EC} is the rate at which one PDE* molecule hydrolyzes cGMP, and β_{dark} is the hydrolysis rate in darkness. Cyclase synthesizes cGMP at a rate

$$\left(\frac{dc}{dt}\right)_{\text{synthesis}} = \alpha. \quad (7.4)$$

The rate α depends on calcium concentration, but in the first few hundred milliseconds after the flash, the calcium concentration is nearly constant. So, in this section, we can assume that α is constant. Hydrolysis and synthesis of cGMP compete:

$$\frac{dc}{dt} = \left(\frac{dc}{dt}\right)_{\text{hydrolysis}} + \left(\frac{dc}{dt}\right)_{\text{synthesis}} = \alpha - (\beta_{\text{dark}} + \nu_{\text{EC}}E^*)c. \quad (7.5)$$

So

$$\frac{dc}{dt} = (\alpha - \beta_{\text{dark}}c) - \nu_{\text{EC}}E^*c. \quad (7.6)$$

Symbols		
Quantity	Units	Description
I	s^{-1}	Light level, as a rate of rhodopsin activation
R^*	—	Number of activated rhodopsin (Rh^*) molecules
E^*	—	Number of activated phosphodiesterase (PDE^*) molecules
c	M	Free cyclic GMP (cGMP) concentration
C	M	Free internal calcium concentration
α	$M s^{-1}$	Rate at which cyclase synthesizes cGMP
ν_{RE}	s^{-1}	Rate at which one activated rhodopsin molecule creates PDE^*
ν_{EC}	s^{-1}	Rate at which one PDE^* molecule hydrolyzes cGMP
β	s^{-1}	Rate at which all the PDE^* molecules hydrolyze cGMP
β_0	s^{-1}	Initial rate at which all the PDE^* molecules hydrolyze cGMP
α	s^{-1}	Rate at which cyclase synthesizes cGMP
$\text{step}(t)$	—	Step function: 0 for $t < 0$, 1 otherwise.
A	s^{-2}	Amplification constant, equal to $\nu_{RE}\nu_{EC}$

Parameters		
Quantity	Value	Description
c_{dark}	$< 4 \mu\text{M}$	cGMP concentration in the dark
β_{dark}	$0.5 s^{-1}$	PDE^* hydrolytic rate in the dark
n	2.0	Hill exponent for channel opening
m	2.0	Hill exponent for calcium inhibition of cyclase
w	1.0	Hill exponent for calcium activation of PDE^* hydrolytic activity
K_{Ca}	90 nM	Binding constant for calcium on cyclase
K	$20 \mu\text{M}$	Binding constant of channels for cyclic GMP
K_β	400 nM	Calcium concentration for half-maximal PDE^* hydrolytic activity
C_{dark}	300 nM	Internal free calcium concentration in the dark

Table 7.1. Symbols and parameters in this analysis. The symbols are mostly the same as Lamb and Pugh's [36]. One exception is that this A is their A/n . Another is that their β_E is ν_{EC} here; with this change, all quantities that contain β are rates, and ν_{EC} is a rate per activated phosphodiesterase molecule. The parameters are an amalgam of those used by Koutalos, Nakatani, and Yau [31], by Koutalos and Yau [32], and by Nikonov, Engheta, and Pugh [44].

Before the light flash, the rod was in the steady state, so creation and destruction of cGMP balanced: $\alpha - \beta_{\text{dark}}c_{\text{dark}} = 0$, where c_{dark} is the cGMP concentration in darkness. Relative to the initial cGMP concentration, $c = c_{\text{dark}} + \Delta c$, so

$$\frac{dc}{dt} = -\beta_{\text{dark}}\Delta c - \nu_{EC}E^*c. \quad (7.7)$$

(The α disappeared, because $\alpha = \beta_{\text{dark}}c_{\text{dark}}$.) Now substitute the ramp expression (7.2) for E^* into this equation:

$$\frac{dc}{dt} = -\beta_{\text{dark}}\Delta c - SAtc, \quad (7.8)$$

where $A \equiv \nu_{EC}\nu_{RE}$. This equation is simple if the $\beta_{\text{dark}}\Delta c$ term can be neglected. That term is negligible for small t , because $\Delta c = 0$ at $t = 0$; in that limit, (7.8) is

$$\frac{dc}{dt} = -(SAt)c. \quad (7.9)$$

For small t , the concentration c is roughly the initial concentration c_{dark} , so

$$\frac{dc}{dt} = -(SA)t c_{\text{dark}}, \quad (7.10)$$

whose solution is

$$c = c_{\text{dark}}(1 - SA t^2/2). \quad (7.11)$$

Using this solution, we can determine the relative contribution of the two terms in (7.8). The change in c is $\Delta c = SA t^2 c_{\text{dark}}/2$, so the two terms on the right side of (7.8) have ratio

$$\frac{\beta_{\text{dark}} \Delta c}{SA t c} \sim \frac{\beta_{\text{dark}} SA t^2 c_{\text{dark}}}{SA t c} \approx \beta_{\text{dark}} t, \quad (7.12)$$

as long as $c_{\text{dark}} \approx c$. (Here, the notation $a \sim b$ means that a/b is a dimensionless constant, generally close to 1.) In the dark, $\beta_{\text{dark}} \approx 0.5 \text{ s}^{-1}$, so the $\beta_{\text{dark}} \Delta c$ term is negligible for roughly the initial 1 s of the response. By then, calcium feedback (which this section neglects) has already quenched much of the response anyway.

The result (7.11) is valid for small t . Small t means until (1) calcium feedback starts to quench the response, or (2) $SA t^2 \sim 1$, whereupon the $c \approx c_{\text{dark}}$ approximation is no longer valid. Let the current change attain a maximum (and start to recover) at time t_p (on the order of 0.3 s), which is related to the time that it takes calcium concentration to change significantly. Then (7.11) is valid at early times for weak and moderate flashes—flashes for which $SA t_p^2 \ll 1$.

The next biochemical step is from cGMP to membrane current: cGMP opens channels in the membrane. Cyclic GMP concentration is highest in the dark, and even then, the level c_{dark} is well below the channel binding constant K . So the membrane current is given by

$$\frac{j(t)}{j_{\text{dark}}} = \left(\frac{c(t)}{c_{\text{dark}}} \right)^n, \quad (7.13)$$

where $j(t)$ is the membrane current, j_{dark} is the membrane current in the dark, and n is the Hill coefficient for channel opening. The membrane current is therefore

$$j(t) = j_{\text{dark}}(1 - SA n t^2/2), \quad (7.14)$$

and the peak (maximum) current decrease is

$$\Delta j(t_p) = j_{\text{dark}} - j(t_p) \approx j_{\text{dark}} SA n t_p^2/2. \quad (7.15)$$

This result is for a dark-adapted rod. The eventual goal of this analysis is to compute the flash sensitivity as a function of background light. What in (7.15) changes with background light? The current j_{dark} becomes j_b , the membrane current as a function of background light; the peak time t_p and the amplification constant A depend on I_b . With dependences on background light shown explicitly, the peak current change is

$$\Delta j(t_p) \approx j_b(I_b) SA(I_b) n t_p(I_b)^2/2. \quad (7.16).$$

The next section, by including the missing feedback paths, estimates the dependence of j_b and A on I_b . The more involved estimate of how t_p depends on I_b is in Section 7.5.

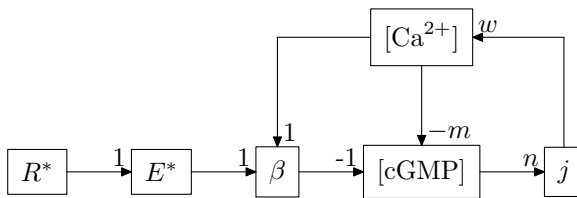


Figure 7.2. From light to membrane current, with feedback. The two new feedback paths depend on calcium concentration: Calcium increases phosphodiesterase hydrolytic rate β (by increasing A), and decreases cyclic GMP concentration (by decreasing cyclase velocity). Because of the kinetics of the calcium pump, the calcium concentration is proportional the membrane current; this linear proportionality is represented by the 1 on arrow from current to calcium. With the approximations in the model of this paper, the hydrolytic rate is roughly $\propto [\text{Ca}^{2+}]$. The cyclase velocity, and therefore the cGMP concentration, is $\propto [\text{Ca}^{2+}]^{-m}$.

7.4 Steady-state response

Figure 7.2 includes feedback paths, which Figure 7.1 lacks. How do these paths make membrane current and hydrolytic rate depend on steady background light level I_b ? The following steady-state model answers this question. It is similar to the steady-state model of Koutalos, Nakatani, and Yau [31], but is even simpler. Although less accurate, it can be solved analytically.

7.4.1 Equations of the model

The model for the steady-state response has four equations.

The first equation is from the calcium concentration. Calcium flows into the rod as part of the membrane current. It leaves the rod through the sodium-potassium-calcium pump. The experiments of Matthews [40] show that calcium current is roughly a fixed fraction (10 percent) of membrane current; the detailed model of Nikonov, Engheta, and Pugh [44] uses this value. So the inward calcium flux is $\propto j$. The experiments of Lagnado, Cervetto, and McNaughton [34] show that the pump is a first-order pump, with binding constant much greater than the free calcium concentration, even in the dark when calcium concentration is the highest. So the outward calcium flux is $\propto C$ (where C is the free internal concentration). The fraction of internal calcium that is free is controlled by a calcium buffer. This fraction is relatively independent of calcium level, as shown in the experiments of Lagnado, Cervetto, and McNaughton [34]. In the steady state, the inward and outward fluxes are equal, so $C \propto j$.

The second equation is from PDE* concentration. PDE* hydrolyzes cGMP at a rate βc (the same behavior as in the dynamics model). Cyclase synthesizes cGMP at a rate α . The models of Nikonov, Engheta, and Pugh [44] and of Koutalos and Yau [32] use

$$\alpha \propto \frac{1}{1 + (C/K_{\text{Ca}})^m}, \quad (7.17)$$

where K_{Ca} is the calcium concentration at which cyclase activity is half maximal, and m is the cooperativity. This equation (and slightly more complicated forms of it) has been

adopted by many other investigators as well. The form (7.17) is accurate enough for this approximate model. Actually, it is too accurate; in low and moderate light, $C \gg K_{\text{Ca}}$, and (7.17) simplifies to $\alpha \propto C^{-m}$. The first measurement of m , by Koch and Stryer [29] in bovine rods, gave $m \approx 4$. The current consensus—based, for example, on the experiments of Koutalos and coworkers [30], and on the experiments and model of Nikonov, Engheta, and Pugh [44]—is that, for salamander rods, $m \approx 2$. In the steady state, hydrolysis and synthesis balance, so $\beta c \propto C^{-m}$, or $c \propto C^{-m} \beta^{-1}$.

The third equation is from phosphodiesterase hydrolytic rate. Background light, by activating rhodopsin, creates PDE^* , which hydrolyzes cGMP. The experiments of Kawamura and Murakami [28], Lagnado and Baylor [33], Jones [27], Matthews [41] and others have shown that calcium increases the hydrolysis rate, by increasing either the production of PDE^* (increasing ν_{RE}) or the activity of PDE^* (increasing ν_{EC}). (Detwiler and Gray-Keller [10] review the studies on this effect of calcium.) For this model, it does not matter at which stage calcium acts, because the current change (7.16) contains ν_{EC} and ν_{RE} as the product $A = \nu_{\text{EC}} \nu_{\text{RE}}$. The experiments on calcium modulation of hydrolytic rate show that $A \propto C^w$ when $C \ll K_\beta$. (The Hill exponent w is approximately 1.0 and $K_\beta \approx 400 \text{ nM}$.) In low background light, when C is near its dark value of $\approx 250 \text{ nM}$, this approximation is not accurate, but it becomes increasingly accurate in brighter light. The feedback diagram places the feedback on ν_{EC} , by placing it on the β block; it could have placed it on the E^* block with equivalent effect. In either case, with $w = 1$, the hydrolytic rate is $\beta \propto C I_b$.

The fourth equation, familiar from the dynamics model, is from cyclic GMP opening channels: $j_b \propto c^n$.

7.4.2 Solution of the equations

The four equations are

$$\beta \propto I_b C, \quad (7.18)$$

$$c \propto \beta^{-1} C^{-m}, \quad (7.19)$$

$$j_b \propto c^n, \quad (7.20)$$

$$C \propto j_b. \quad (7.21)$$

The feedback diagram is equivalent to these equations: one block for one equation. For example, the [cGMP] block has one arrow from $[\text{Ca}^{2+}]$ labeled with $-m$; this value is exponent of C in (7.19). The [cGMP] block also has an arrow from β labeled with -1 ; this value is the exponent of β in (7.19).

A solution to these four equations in five variables is an equation between two variables. Which equations do we need to generalize the dynamics model of Section 7.3 to nonzero background light? The change in current (7.16) contains the amplification constant A , which depends calcium concentration; it also contains the steady-state membrane current, j_b , which depends on background light. So we solve for membrane current (j_b) and calcium concentration (C) as functions of I_b .

First, use (7.18) to eliminate β from the cGMP equation (7.19):

$$c \propto I_b^{-1} C^{-(m+1)}. \quad (7.22)$$

Using this result, eliminate c from (7.20):

$$j_b \propto I_b^{-n} C^{-n(m+1)}. \quad (7.23)$$

Then, eliminate C using (7.21), and combine powers of j_b :

$$j_b^{n(m+1)+1} \propto I_b^{-n}. \quad (7.24)$$

The membrane current is

$$j_b \propto I_b^{-\theta}, \quad (7.25)$$

where $\theta \equiv 1/(m + 1 + 1/n)$. Because $C \propto j_b$, the calcium concentration obeys the same scaling as the membrane current:

$$C \propto I_b^{-\theta}. \quad (7.26)$$

The calcium concentration and the membrane current are, because of calcium feedback, compressive functions of the background light. (A compressive function—for example, $\log x$ —is one that grows more slowly than x , or falls more slowly than $1/x$. Such functions are ubiquitous in psychophysics and sensory transduction.) The form of $1/\theta$ shows which feedback paths contribute the most compression. The major contributor to $1/\theta$ is m , the cooperativity of calcium binding on cyclase. The next largest contributor, 1, is from calcium feedback on the first stages of amplification. The smallest contributor, $1/n$, is from channel cooperativity. Control theory (see for example Siebert's text [56]) formalizes these intuitions. The rod system is a negative feedback system with one forward path. As long as the gain in the forward path is large enough, it becomes irrelevant, and the feedback paths dominate the closed-loop gain. The general principle is: If an ideal (infinite-forward-gain) feedback loop contains $f(x)$ in the feedback path, the full circuit computes $f^{-1}(x)$ (Figure 7.3). The rod places x^{-3} in the feedback path, and computes approximately $x^{-1/3}$. The response deviates from $x^{-1/3}$ because the forward gain is not infinite (it's n , which is merely 2).

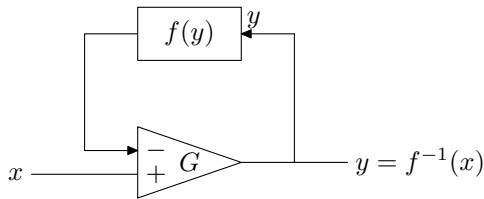


Figure 7.3. Computation of f^{-1} using a feedback loop. The triangle is a high-gain differential amplifier (for example, an operational amplifier): Its output is $G(v_+ - v_-)$, and G is large. The input voltage is x (on the noninverting terminal). The amplifier, because of its high gain (and the correct feedback sign!) maintains the inverting $(-)$ and noninverting $(+)$ terminals at the same voltage. So the input to the $f(y)$ box must be $f^{-1}(x)$, which is the output. In this representation of the rod system, the gains are the exponents on the arrows in Figures 7.1 and 7.2, rather than multiplicative constants, such as ν_{RE} or ν_{EC} , which this model ignores.

7.4.3 Similarity to the logarithm

Matthews and coworkers [42], Koutalos, Nakatani, and Yau [31, Figure 9], and others have shown that the rod's steady-state response is approximated by $\log I_b$ over an intensity range of a factor of 100.

The functional form of (7.25)—a high-order root—is an approximation to a logarithm, as the following argument shows. Consider the function a^n ; then $\int a^n da \propto a^{n+1}$. As $n \rightarrow -1$, the right side tends to x^0 , while the left side becomes $\int a^{-1} da$, whose integral is $\ln a$. Therefore, $\ln a \propto a^0$; a high-order root (represented by a zero exponent) approximates a logarithm. This heuristic argument neglected the constants of integration and of proportionality. Here is a more careful argument, which shows that $q(a^{1/q} - 1) \approx \log a$, for large n . Replace $a^{1/q}$ by $e^{(\log a)/q}$, and then approximate the exponential with the Taylor series $e^z = 1 + z + O(z^2)$:

$$q(a^{1/q} - 1) = q(e^{(\log a)/q} - 1) \approx q \left\{ \frac{\log a}{q} + O\left(\frac{\log a}{q}\right)^2 \right\}. \quad (7.27)$$

The last expression is roughly $\log a$, when $(\log a)^2 < q$. In this model, q is $1/\theta$, or roughly 3.5, which implies that $a < 6$ for the approximation to work.

Empirically, the rod's steady-state response approximates a logarithm over a larger range than what this model predicts. This model might lack pathways that would increase the order of the root at higher light levels. Also, these drastic approximations neglect effects that might increase the effect order of the root. Saturation is one such effect. The maximum change in membrane current is limited by the dark current. At high light levels, the current suppressed is less than proportional to the dark current (because Δj cannot exceed j_{dark}). This model neglects this kind of compression.

7.5 Flash sensitivity

The flash sensitivity is the peak current change, $\Delta j(t_p)$, divided by the flash size, S . Equation (7.16) gives the peak current change; with that result, the flash sensitivity becomes

$$S_F \equiv \frac{\Delta j(t_p)}{S} \sim j_b A t_p^2. \quad (7.28)$$

The \sim notation (meaning “correct except for a dimensionless constant”) gave us the freedom to discard the $n/2$ factor. The steady-state model shows that $A \propto I_b^{-\theta}$ and $j_b \propto I_b^{-\theta}$. One ingredient is missing: How does t_p depend on I_b ?

What determines t_p ? After a flash of light, the PDE* hydrolytic rate increases; this increase reduces c (the concentration of cGMP), and therefore reduces the membrane current. As the membrane current falls, less calcium enters the rod, so the internal calcium concentration falls. As the calcium concentration falls, the cyclase rate increases, which increases the rate of cGMP synthesis. Eventually, the increase in synthesis balances the increase in hydrolysis. At that moment—the peak time— c is constant (more precisely, $dc/dt = 0$), so the membrane current is constant (the membrane channels respond within milliseconds—that is, almost instantly—to changes in cGMP concentration). Let's translate this argument into mathematics.

The first step is to compute the rate of decrease in cGMP concentration due to the increase in hydrolysis. The increase in E^* is given by (7.2): $\Delta E^* = S\nu_{\text{Ret}}$. The contribution

to dc/dt is $\nu_{EC}(\Delta E^*)c$ (in absolute value), or $SAct$. For weak and moderate flashes, the initial cGMP concentration c_0 is an adequate approximation for c . Then the contribution to dc/dt is

$$\left(\frac{dc}{dt}\right)_{\text{hydrolysis}} \sim SAc_0t. \quad (7.29)$$

The initial cGMP concentration is $\propto j_b^{1/n}$. The result (7.25) turns this proportionality into $c_0 \propto I_b^{-\theta/n}$; then

$$\left(\frac{dc}{dt}\right)_{\text{hydrolysis}} \propto -SA t I_b^{-\theta/n}. \quad (7.30)$$

The next step is to compute rate of increase in cGMP synthesis due to the increase in cyclase velocity. The cyclase velocity increases because calcium concentration decreases. So let's determine the change in calcium concentration. The calcium concentration changes because the membrane current decreases. Equation (7.14) gives the change in membrane current: $\Delta j(t) = -SA t^2 j_b$ (we have to replace j_{dark} by j_b , and we explicitly include the sign, because it is crucial). The calcium influx is proportional to the membrane current, so the change in inward calcium current is

$$\Delta j_{Ca} = -SA t^2 \gamma j_b, \quad (7.31)$$

where γ is a constant that converts membrane current into calcium current (the constant will soon disappear). This current change causes a calcium-concentration change. The calcium pump has first-order kinetics. A more intuitive version of this fact is that the calcium pump integrates the calcium current to produce calcium concentration. For example, a spike in calcium current produces a step in calcium concentration; a step in calcium current produces a ramp in calcium concentration. In general, the integrator turns a current change $\propto t^x$ into a concentration change $\propto t^{x+1}$. In particular, a t^2 change in calcium current—which is what the light flash causes—produces a t^3 change in calcium concentration. So the fractional change in calcium concentration is

$$\frac{\Delta C}{C} \sim -SA t^3 / \tau, \quad (7.32)$$

where τ is the time constant of the calcium pump (τ is roughly 0.7 s). [This expression is valid for $t \ll \tau$; when $t \sim \tau$, the pump ceases to be an ideal integrator.] The cyclase rate is $\propto C^{-m}$, so the fractional change in cyclase rate is

$$\frac{\Delta \alpha}{\alpha} \sim mSA t^3 / \tau \sim SA t^3 / \tau. \quad (7.33)$$

(The \sim notation allows us to discard the constant m .) What we need is the change in cyclase rate itself:

$$\Delta \alpha \sim SA t^3 \alpha / \tau. \quad (7.34)$$

The calcium concentration determines the cyclase rate: $\alpha \propto C^{-m}$. From (7.26), $C \propto I_b^{-\theta}$, so $\alpha \propto I_b^{m\theta}$. Then the increased-cyclase-velocity contribution to dc/dt is

$$\left(\frac{dc}{dt}\right)_{\text{synthesis}} \propto SA t^3 I_b^{m\theta} / \tau. \quad (7.35)$$

To find the peak time, we find when the two contributions to dc/dt , given in (7.30) and (7.35), cancel:

$$SA t_p I_b^{-\theta/n} \propto SA t_p^3 I_b^{m\theta} / \tau. \quad (7.36)$$

The peak time is therefore

$$t_p \propto I_b^{-(m+1/n)\theta/2} = I_b^{-(1-\theta)/2}. \quad (7.37)$$

The experiments of Donner and coworkers [11] show that, for amphibian rods, $t_p \propto I_b^l$, where $l \approx -0.17$. With the cooperativities $n = 2$ and $m = 2$, as given in Table 7.1, $-(m+1/n)\theta/2 = -0.37$. This model, in its drastic approximations, must neglect compressive effects that would make the estimated exponent closer to the experimentally observed value.

With the dependence in (7.37), the flash sensitivity (7.28) is

$$S_F \propto \underbrace{I_b^{-\theta}}_{j_b} \times \underbrace{I_b^{-\theta}}_A \times \underbrace{I_b^{-(1-\theta)}}_{t_p^2} \propto I_b^{-(1+\theta)}. \quad (7.38)$$

Note that if we differentiate the steady-state response, $j_b \propto I_b^{-\theta}$, with respect to I_b , then we also get this expression for the flash sensitivity. With the values $n \approx 2$ and $m \approx 2$, the prediction is that

$$S_F \propto I_b^{-1.3}. \quad (7.39)$$

This exponent is a reasonable approximation to the Weber–Fechner exponent of -1 . Over a range of light intensity range of 100, it deviates from Weber–Fechner behavior by a factor of 4.

Although this model explains the exponent, it has a flaw: the assumption that $\beta_0 t \ll 1$. In the dark, where $\beta_0 t_p \ll 1$, the assumption is accurate. However, background light increases β_0 . In background light, $\beta_0 \propto I_b^{1-\theta}$. More precisely,

$$\frac{\beta_0(I_b)}{\beta_0(\text{dark})} \approx \left(\frac{I_b}{I_{\text{dark}}} \right)^{1-\theta}, \quad (7.40)$$

where I_{dark} is the dark light (typically $20 \text{ Rh}^* \text{ s}^{-1}$ for a salamander rod). At the moderate light intensity of $200 \text{ Rh}^* \text{ s}^{-1}$, the basal hydrolysis rate becomes 3.5 s^{-1} , whereupon $\beta_0 t_p$ is no longer $\ll 1$. I do not yet see how to repair this flaw. It hints that either a reaction is missing from this model; or that the approximations are too drastic, and neglect important features of the kinetics.

7.6 Conclusion

I made an inexact, but analytically solvable, model of flash sensitivity that explains the Weber–Fechner behavior in a salamander rod. The model combines the dark-adapted flash-response analysis of Lamb and Pugh [37] with an analytic model of steady-state sensitivity, to derive the flash response—and hence the flash sensitivity—as a function of background light. The simplicity of this model comes at a price: It predicts a response that roughly, but not exactly, matches the experiment. However, because the model is simple, it shows clearly how each biochemical pathway contributes to the flash response.

8 DENSITY OF PRIMES

8.1 Approximate primality by stages

The prime-number theorem states that \hat{p}_n , the density of primes around n , is $\sim (\log n)^{-1}$. The proof is difficult, and when it is complete, we may still wonder, “But *why* is it true?” We want to understand what features of primeness cause what features of prime behavior. To tease apart these behaviors and their causes, we make a model of primality, where *model* is used the physicists’ sense, as a simplification that still explains interesting features of the original phenomenon.

The first behavior to explain is why $\hat{p}_n \sim (\log n)^{-1}$. First, let’s consider a simpler question: Why do primes thin? A number n is prime if and only if no prime $\leq \sqrt{n}$ divides n . As n grows, so does the number of primes that might divide n ; therefore, n is less likely to be prime. Already, by invoking likelihood, we have sinned. Hardy and Littlewood [19] write that “*Probability* is not a notion of pure mathematics, but of physics or philosophy.” As a physicist, I am bathed in sin; a few more transgressions in this chapter will not harm me (you must decide for yourself how safe you feel). So let’s continue by formalizing the likelihood argument. Suppose that divisibility is random and independent of other divisibilities: Let $m|n$ with probability $1/m$. For \hat{p}_n , this model predicts

$$p_n = \prod_{p \text{ prime}}^{\lfloor n^{1/2} \rfloor} \left(1 - \frac{1}{p}\right). \quad (8.1)$$

In this chapter, p_n refers to the density given by a model of primality, and \hat{p}_n refers to the density of actual primes.

In the preceding paragraph, I did not carefully define the meaning of random divisibility, because when I created these models, I had no clear notion of its definition. In Section 8.1.2, I define it carefully, because the careful definitions allow us to prove statements about the behavior of p_n .

I expected $p_n \sim (\log n)^{-1}$ to be the asymptotic solution of the product (8.1). When I computed p_n for large n (Figure 8.1), I was pleased to see $p_n \log n$ approaching a constant, but I was not pleased that the constant was ≈ 1.12 (instead of 1). So this model explains the crudest feature in the density of primes—the presence of $(\log n)^{-1}$ —but its usefulness stops there. As Hardy and Littlewood [19] and Pólya [47] have pointed out, the discrepancy is a consequence of Mertens’ theorem (Theorem 429 of Hardy and Wright [20]):

$$\prod_{p \text{ prime}}^{\lfloor n^{1/2} \rfloor} \left(1 - \frac{1}{p}\right) \sim \frac{2e^{-\gamma}}{\log n}, \quad (8.2)$$

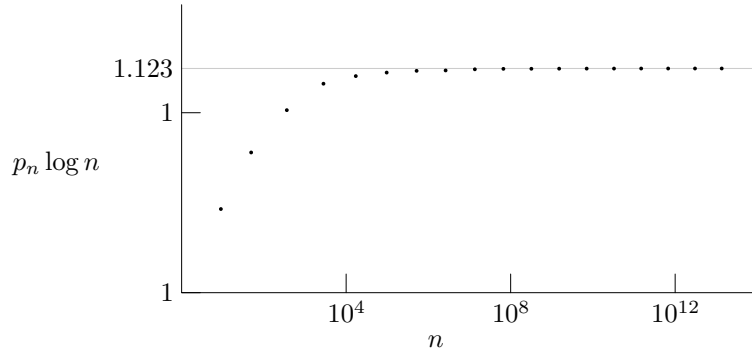


Figure 8.1. Predicted prime probability p_n compared against the prime-number theorem value of $(\log n)^{-1}$. The probability p_n looks like it approaches $(\log n)^{-1}$ times a constant. The constant, unfortunately, is $2e^{-\gamma} \approx 1.123$, not 1.

where $\gamma \approx 0.577$ is Euler's constant, and $\lfloor a \rfloor$ is the greatest integer $\leq a$. The factor $2e^{-\gamma}$ is ≈ 1.123 . In other words, actual primes strike out candidates more efficiently than independence would predict.

In the prediction (8.1), $\lfloor n^{1/2} \rfloor$ is the upper limit, because the square-root is an important feature of actual primes. For example, in actual divisibility, if $3 \mid 81$, then $27 \mid 81$. Such statements are not valid when divisibility is random; in such a world, 3 may divide 81, yet 27 may not divide 81. The $\lfloor n^{1/2} \rfloor$ limit is an attempt to capture, in a random-divisibility model, the correlations of actual divisibility. This attempt is worthwhile. If we ignore these correlations and use $n-1$ as the upper limit, instead of $\lfloor n^{1/2} \rfloor$, then Mertens' theorem gives $p_n \sim e^{-\gamma}/\log n$. This result is too small by a factor of 0.56, whereas the preceding method, though not exact, was too large by a factor of only 1.12. The $\lfloor n^{1/2} \rfloor$ limit captures yet more correlations.

We could fix the remaining error by improving the independence assumption. However, I do not know how to improve it in a convincing way. One method is to increase the upper limit in the product (8.1), but I have no principled way of choosing the new limit, short of using Mertens' theorem (varying the limit is the topic of Section 8.6).

An alternative fix is to be more consistent—to drink more deeply from the well of probabilistic sin. We computed p_n , which implicitly assumes that primality is random, using actual primes, and got the incorrect $1 - 1/p$ factor in (8.1). We can fix this inconsistency by making primality random. How does random primality change the $1 - 1/p$ factor in the computation of p_n ? Because any number could be prime (in this random world), the product runs over all $k \leq \lfloor n^{1/2} \rfloor$, instead of over only primes. If k is prime (which happens with probability p_k), then the algorithm checks whether $k \mid n$, giving a factor of $1 - 1/k$. If k is not prime, then the factor is 1. The expected factor is $p_k \times (1 - 1/k) + (1 - p_k) \times 1 = 1 - p_k/k$. So, in this model,

$$p_n = \prod_1^{\lfloor n^{1/2} \rfloor} \left(1 - \frac{p_k}{k}\right). \quad (8.3)$$

This recurrence was my second model, which I call the square-root model. I later

learned¹ that it is similar to the random sieve proposed by Hawkins' [21, 22].

8.1.1 Hawkins' model

The Hawkins' random sieve (Algorithm 8.1) generates a random sequence of H-primes H_1, H_2, \dots . Hawkins calls them *sieving numbers*. I have changed Hawkins' terminology and notation slightly, to make them consistent with the notation of this chapter.

Algorithm 8.1 Hawkins' model

1. Define $H_1 = 2$.
2. Strike out each integer with probability $1/H_1$ (that is, with probability $1/2$).
3. Define H_2 to be the first integer $> H_1$ that survived. Strike out each integer with probability $1/H_2$.
4. Define H_3 to be the first integer $> H_2$ that survived. Strike out each integer with probability $1/H_3$.
5. and so on

The sequence of H-primes is different each time that the algorithm is run. Hawkins' studied h_n , the probability that n is an H-prime. The probability satisfies

$$h_{n+1} = h_n \left(1 - \frac{h_n}{n}\right). \quad (8.4)$$

To solve this recurrence, define $q_n = 1/h_n$:

$$q_{n+1} = q_n \left/ \left(1 - \frac{1}{nq_n}\right)\right. = q_n + \frac{1}{n} + O\left(\frac{1}{n^2q_n}\right). \quad (8.5)$$

The error term, which is the tail of a geometric series with positive ratio, is always positive. Therefore, $q_n \geq 1$, and the recurrence has bounds

$$\frac{1}{n} \leq q_{n+1} - q_n \leq \frac{1}{n} + \frac{1}{n^2} \leq \frac{1}{n-1}. \quad (8.6)$$

The solution is $q_n \geq \log n$ (by an easy induction argument). So h_n , which is $1/q_n$, approaches $(\log n)^{-1}$. Though this argument does not show it, p_n monotonically approaches $(\log n)^{-1}$.

8.1.2 Square-root model

The heuristic argument that led to the recurrence (8.3) does not specify the process for which p_n is a probability, and thereby leaves vague the notions of random primality and random divisibility. I now describe such a process, as a modification of Hawkins' algorithm. Algorithm 8.2 generates a sequence of random primes (R-primes).

Algorithm 8.2 Square-root model

1. Define $R_1 = 2$.
2. Strike out each integer ≥ 4 with probability $1/R_1$ (that is, with probability $1/2$).

1. I thank Warren Smith of NEC for pointing out Hawkins' work to me.

3. Define R_2 to be the first integer $> R_1$ that survived. Strike out each integer $\geq R_2^2$ with probability $1/R_2$.
4. Define R_3 to be the first integer $> R_2$ that survived. Strike out each integer $\geq R_3^2$ with probability $1/R_3$.
5. and so on

The quantity p_n is the probability that n appears on the list of R-primes. This algorithm is the same as Hawkins' (Algorithm 8.1), except that it uses $\geq R_n^2$ as the least number to consider striking out. Hawkins [22] briefly mentioned this modification as a natural extension of his model, but he did not analyze its consequences.

Algorithm 8.3 is an alternative, equivalent version of Algorithm 8.2; it also generates a sequence of R-primes. It is the form that I use to deduce the recurrence (8.7) for p_n .

Algorithm 8.3 Equivalent square-root model

1. Define $R_1 = 2$.
2. For $n > 2$, check whether any R-prime $\leq \sqrt{n}$ divides n . Here, *divide* is random. To decide whether $m|n$, roll an m -sided die, with faces labeled $1 \dots m$. If the die shows 1, then $m|n$ and n is not R-prime. If no die roll shows 1 (namely, if no R-prime divides n), then n is R-prime.

Section 8.1.3 shows that p_n satisfies the product (8.3).

The square-root model more closely models actual primes. Whereas h_n tends monotonically to $(\log n)^{-1}$, the probability p_n oscillates around $(\log n)^{-1}$, as Section 8.3.3 shows. A heuristic argument shows why the probability oscillates. Consider a perturbation around the nonoscillating part of the density. For example, suppose that the die in Algorithm 8.3 rolls too many 1s while testing primality in a region around some number A . In that region, too few numbers are declared R-prime. This perturbation makes itself felt when the algorithm starts to test numbers around A^2 ; in that region, too many numbers are declared R-prime, because there are too few candidate divisors (R-primes) around A . The density displays negative feedback. The excess around A^2 makes itself felt when the algorithm starts to test numbers around A^4 ; in that region, too few numbers are declared R-prime, because there are too many candidate divisors around A . And so on. The feedback points grow by successive squaring: A to A^2 to A^4 , and so on. The oscillation period, which includes two feedback points, therefore grows by successive fourth powers: A to A^4 to A^{16} . As we see in Section 8.4.1, this heuristic argument is mostly correct: The period does grow superexponentially, but the growth is closer to A to A^{64} to A^{4096} . The density has feedback with a delay; an electrical engineer will confirm that oscillations are a natural consequence of this combination.

This oscillation models the oscillations in the density of actual primes. I find an asymptotic expressions for the oscillatory contribution, and estimate its magnitude. This estimate leads to a conjecture for where $\pi(n)$ —the number of primes $\leq n$ —first crosses $\text{Li}(n)$. (I use the usual definition: $\text{Li}(n) = \int_2^n (\log x)^{-1} dx$.) No n is known for which $\pi(n) > \text{Li}(n)$. However, Littlewood [39] showed that $\pi(n)$ crosses $\text{Li}(n)$ infinitely often; Littlewood's result shows that the density of actual primes oscillates around $(\log n)^{-1}$. His argument was nonconstructive—it did not provide a explicit n for any crossing. The most recent, related

result is partially constructive: Te Riele [51] showed that the first crossing occurs for some $n < 7 \times 10^{370}$. The oscillations in the density of R-primes leads me to conjecture that $\pi(n)$ crosses $\text{Li}(n)$ near 10^{27} (Section 8.4).

The remainder of the chapter analyzes the recurrence and discusses its consequences.

8.1.3 Recurrence for p_n

To solve for p_n , I need an equation for it, which is given by Theorem 8.1. I already have an equation for p_n , namely Equation (8.3); however, I arrived at that equation using heuristic arguments about random divisibility. Theorem 8.1 shows that the heuristically generated recurrence correctly describes the p_n that results from Algorithms 8.2 or 8.3.

Theorem 8.1 The probability p_n satisfies

$$p_n = p_{n-1} \times \begin{cases} 1, & \text{if } n \text{ is not a perfect square;} \\ 1 - p_{\sqrt{n}}/\sqrt{n}, & \text{if } n \text{ is a perfect square.} \end{cases} \quad (8.7)$$

Proof. If n is not a perfect square, then $n-1$ and n have the same set of R-primes that could divide them, so $p_n = p_{n-1}$, as claimed,

If n is a perfect square, say k^2 , then there are two cases: k is or is not R-prime. I compute the conditional probability that n is R-prime in both cases, and combine these probabilities to verify the recurrence.

Case 1: k is not R-prime. This case occurs with probability $1 - p_k$. Going from $k^2 - 1$ to k^2 could introduce k as a candidate divisor. However, k is not R-prime, so $p_n = p_{n-1}$.

Case 2: k is R-prime. This case occurs with probability p_k . Now k is a new candidate divisor, and it eliminates k^2 with probability $1/k$. So $p_n = p_{n-1}(1 - 1/k)$.

The conditional probabilities combine to give

$$p_n = p_{n-1} \left\{ (1 - p_k) \times 1 + p_k \times \left(1 - \frac{1}{k} \right) \right\}. \quad (8.8)$$

So $p_n = p_{n-1}(1 - p_k/k)$, which verifies the recurrence (8.7). ■

Unfolded into a product, the recurrence (8.7) becomes the recursive product (8.3). To solve the recurrence, I convert it into a delay-differential equation, then solve that equation.

8.2 Convert the recurrence into a delay-differential equation

The square-root model has delayed negative feedback. A delay-differential equation is therefore a natural form for the recurrence.

The conversion to a delay-differential equation has four steps:

1. Change variables to $l_n = \log p_n$. This change converts multiplication into addition, which converts the recurrence (8.7) into a difference equation.
2. Change to a continuous version of l_n , which I call c_n (for *continuous*). A differential equation requires functions without jumps, because jumps imply infinite derivatives. The function l_n jumps when n is a perfect square, and is flat in between; c_n distributes the change continuously.
3. Change to $y_n = c_n + \log \log n$. If $p_n \sim (\log n)^{-1}$, then $y_n \rightarrow 0$. Because $y \rightarrow 0$ (Theorem 8.4), this change of variable allows us to linearize the eventual delay-differential equation.

4. Change independent variable to $t = \log_2 \log n$. Feedback means, roughly, that y_n affects y_{n^2} , which affects y_{n^4} , and so on. The definition of t changes the superexponential growth in the feedback points to the more tranquil sequence: $y(1)$ affects $y(2)$, which affects $y(3)$, and so on. This change of variable is much more convenient in a differential equation than in a difference equation (which is why the transformations above are designed to manufacture a differential equation).

8.2.1 Change multiplication to addition

A differential equation needs summation, not multiplication. So define $l_n = \log p_n$. Creeping toward continuous variables, further define $l_a = l_{[a]}$ for nonintegral a . See Figure 8.2 for a picture of l . It changes at only perfect squares, and is flat in between. So (8.7) becomes $l_{n^2} - l_{n^2-x} = \log(1 - e^{l_n}/n)$ for $0 < x < 2n - 1$. Now get rid of the logarithm:

$$l_{n^2} - l_{n^2-x} = -\frac{e^{l_n}}{n} + O(n^{-2}), \quad \text{for } 0 < x < 2n - 1, \quad (8.9)$$

because $e^{l_n} < 1$.

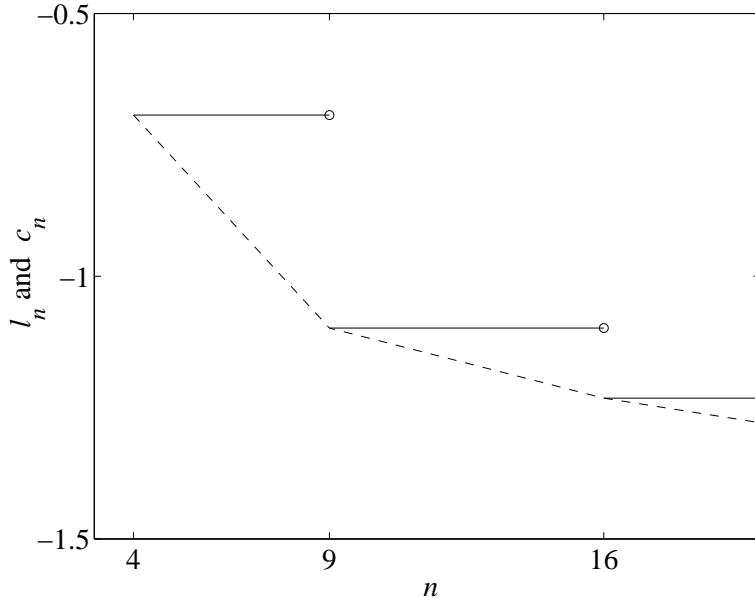


Figure 8.2. The functions l_n and c_n . The function l_n (the solid lines) is flat, with jumps at perfect squares, as indicated by the open circles. The discontinuity in l_n at a perfect square is, from (8.9), $O(n^{-1/2})$, so $l_n - l_{n-1} = O(n^{-1/2})$. The function c_n (the dotted line) is the continuous version of l_n . When n is a perfect square, $c_n = l_n$, and c_n interpolates linearly between the values at perfect squares. The distance between perfect squares in the neighborhood of n is $\geq \sqrt{n}$, so $|c_n - c_{n-1}| \leq |l_n - l_{n-1}|/\sqrt{n} = O(n^{-1})$.

8.2.2 Make the dependent variable continuous

The function l jumps at perfect squares, so it is unsuitable for a differential equation, because discontinuous functions have infinite derivatives. So define c_n , a continuous version of l_n (Figure 8.2). [The more common way to make l_n continuous is to round the corners of the cliffs in l_n , and to leave it flat in between the cliffs. That method produces a function with huge derivatives, which would make the error estimates difficult.]

From (8.9), compute $dc/dn = c'$ (except when n is a perfect square):

$$c'_{n^2-x} = \frac{l_{n^2} - l_{n^2-1}}{2n-1} = -\frac{e^{l_n}}{2n^2} + O(n^{-3}), \quad \text{for } 0 \leq x < 2n-1. \quad (8.10)$$

The conversion is not complete, because the right side still contains l_n , and because the $x = 0$ case (perfect squares), where c' is discontinuous, still requires discussion. To eliminate l_n , note that $l_n = c_n + O(n^{-1/2})$, so $e^{l_n} = e^{c_n} (1 + O(n^{-1/2}))$, and

$$c'_{n^2-x} = -\frac{e^{c_n}}{2n^2} + O(n^{-5/2}), \quad \text{for } 0 \leq x < 2n-1. \quad (8.11)$$

This equation is valid for $x = 0$, as I show now, even though its ancestor (8.9) was not. The left- and right-sided derivatives at n^2 are not equal. Define c'_{n^2} as their average; then c' is defined everywhere (although it is not continuous). The right-sided derivative is, for $x = 0^+$,

$$c'_{n^2+x} = c'_{(n+1)^2-x} = -\frac{e^{c_{n+1}}}{2(n+1)^2} + O(n^{-5/2}), \quad (8.12)$$

where I used (8.11) with n replaced by $n+1$ to rewrite $c'_{(n+1)^2-x}$. Because $c_{n+1} - c_n = O(n^{-1})$ (Figure 8.2), $e^{c_{n+1}} = e^{c_n} (1 + O(n^{-1}))$. Therefore, $c'_{n^2+x} = -e^{c_n}/2n^2 + O(n^{-5/2})$ for $x = 0^+$. This expression for the right-sided derivative is identical to the expression (8.11) for the left-sided derivative. So the average of the left- and right-sided derivatives also satisfies (8.11), which is therefore valid for $x = 0$.

Now get rid of the x in (8.11), which I do by defining $q^2 = n^2 - x$, so $n = q + \xi$, where $0 \leq \xi < 1$, because $x < 2n-1$. And $c_{q+\xi} = c_q + O(q^{-1})$, so (8.11) becomes

$$c'_{q^2} = -\frac{e^{c_q}}{2q^2} + O(q^{-5/2}). \quad (8.13)$$

8.2.3 Convert to a variable with zero limit

If $p_n \sim (\log n)^{-1}$, then $c_q \sim -\log \log q$, a value that does not limit to zero. To make neglecting terms easier, transform to a variable with a (hopefully) small limit by defining $y_q = c_q + \log \log q$, and $y'_q = c'_q + 1/q \log q$. The definition of y'_q is valid everywhere, because c'_q is defined everywhere (although y'_q , like c'_q , is discontinuous). Then (8.13) becomes

$$y'_{q^2} = -\left(\frac{e^{y_q} - 1}{2q^2 \log q}\right) + O(q^{-5/2}). \quad (8.14)$$

The equation for q is

$$y'_q = -\left(\frac{e^{y\sqrt{q}} - 1}{q \log q}\right) + O(q^{-5/4}). \quad (8.15)$$

This equation is the delay-differential equation.

8.2.4 Taming the growth of n

The mixed differential–recurrence relation (8.15) has delayed negative feedback, but the feedback points grow by successive squaring. To tame this wild growth, change variables from n to t , where t is defined implicitly by $y(t) = y_{e^{2^t}}$, or $t = \log_2 \log q$. So t counts the number of successive squarings, and I expect $y(t)$ to oscillate with period roughly 2. With this change of variable, the delay–differential equation becomes

$$\dot{y}(t) = -\alpha \left(e^{y(t-1)} - 1 \right) + O \left(2^t e^{-2^{t-2}} \right), \quad (8.16)$$

where $\alpha = \log 2$ and $\dot{y}(t) = dy/dt$. We can tell that $t = \log_2 \log q$ is a useful definition for t : It makes the ugly $q \log q$ factor in (8.15) vanish (because $dq = q \log_2 \log q \, dt$), and makes the delay constant (because $y_{\sqrt{q}} = y(t-1)$).

If $\alpha < 0$, the feedback would reinforce deviations from the correct density, and we could not expect that $p_n \sim (\log n)^{-1}$. So, without further statement, let's assume that $\alpha > 0$. Let's not restrict ourselves to models with $\alpha = \log 2$, because varying α produces closely related, and interesting models (the topic of Section 8.6). Note that, by construction, $y(t)$ is continuous; its derivative $\dot{y}(t)$ has discontinuities, but those discontinuities are swallowed in the error term. Because (8.16) contains a delay, the initial conditions have to specify $y(t)$ on an interval of length ≥ 1 (because of the $y(t-1)$ term), instead of at a point, as they would for a typical differential equation.

8.2.5 Related equations

Wright [65] discusses

$$w'(t) = -\alpha w(t-1) \{1 + w(t)\}, \quad (8.17)$$

where $w(t) = p_n \log n - 1$ and $\alpha = \log 2$. He defines $t = \log_2 \log n$, which is nearly identical to our definition, $t = \log_2 \log q$, because $q = n + O(1)$. For small y , $w(t) \approx y(t)$. The linearized forms of (8.17) and (8.16) are identical, as are the asymptotic solutions. Wright was led to (8.17) first by the (unpublished) work of Lord Cherwell on the distribution of primes, and later by de Visme's [9] heuristic equation for the density of primes. Wright's studies of nonlinear delay–differential equations [61, 62, 63, 64] were inspired by his study of (8.17). To solve the closely related delay–differential equation (8.16), I use many of these methods that Wright developed, most particularly those in [64].

8.3 Solve the delay–differential equation

Delay–differential equations are difficult to solve, because their initial conditions are functions, not merely numbers; making statements about functions is more difficult than is making statements about numbers. We can solve (8.16) by successive refinement, using a method that works with minimum restriction on the initial conditions:

1. Show that $y \rightarrow 0$ as $t \rightarrow \infty$, independent of initial condition.
2. Use $y \rightarrow 0$ to show that $y = O(e^{-Ct})$, where $C > 0$.
3. Use $y = O(e^{-Ct})$ to find a more detailed asymptotic expression for y .

8.3.1 Show that $y \rightarrow 0$

To show that $y \rightarrow 0$, consider two cases: The zeros of y are either bounded or unbounded. Here, *bounded zeros* means that $y(t)$ is never zero for all t greater than some t_0 ; *unbounded zeros* means there is no such t_0 . First tackle the easier, bounded-zero case.

Theorem 8.2 If the zeros of y are bounded, then $y \rightarrow 0$.

Proof. Suppose that y has a finite limit—say, y^* . Take the limit $t \rightarrow \infty$ in the delay-differential equation (8.16), to get $\lim_{t \rightarrow \infty} \dot{y}(t) = -\alpha(e^{y^*} - 1)$, which is some constant. That constant must be zero for y to have a finite limit. If $\dot{y} \rightarrow 0$, then (8.16) requires that $y \rightarrow 0$ as well. So we need to show only that y has a finite limit, which we now do.

Let t_0 be the final zero of y , so $y(t)$ has one sign for $t > t_0$. Suppose that y is positive. If we could show that $\dot{y} < 0$, then we could conclude that y has a finite limit, because y would be bounded from below and would be monotonic decreasing. Therefore, compute the upper bound on \dot{y} . For $t > t_0 + 1$, use $e^{y(t-1)} - 1 > 0$ to get the bound $\dot{y}(t) < O(t^{-2})$ (using a much weaker error statement than what is available). For large enough t —say, for $t > t_1$ —the error term is $\leq A/t^2$ (in absolute value), for some $A > 0$. Then, $\dot{y}(t) < A/t^2$ for $t > \max(t_1, t_0 + 1)$. Unfortunately, $\dot{y}(t)$ can be slightly positive, because of the error term in (8.16). The method almost works, but does not quite succeed.

Therefore, define the related function $q(t) = y(t) + A/t$ (for $t > 0$). Then, $q(t) > 0$ and $q'(t) = \dot{y}(t) - A/t^2 < 0$. So $q(t)$ has a limit—say, Q —and y has the same limit. If y is negative, then the same method works, except that I define $q(t) = y(t) - A/t$, and conclude that $y(t) \rightarrow Q$ again. In either case, y has a finite limit, so $y \rightarrow 0$. ■

The unbounded-zero case is more subtle. I first explain the proof intuitively, forgetting about the error term. The most negative y limits the maximum positive \dot{y} , which limits how fast y can rise. Furthermore, y cannot rise for longer than one time unit, because if $y(t) = 0$, then $\dot{y}(t+1) = 0$, so y reaches a maximum at $t+1$. This maximum in turn determines the most negative \dot{y} , and, therefore, the new minimum. Thus, we get sequences of upper and of lower bounds, and it turns out that both sequences converge to zero (to take the limit, we need to know that the zeros are unbounded). The following proof formalizes this procedure, and is careful about the error term.

Theorem 8.3 If $\alpha \leq 1$ and the zeros of y are unbounded, then $y \rightarrow 0$.

Proof. First, formalize maximum and minimum, by defining $u(t) = \sup_{(t, \infty)} y(z)$ and $v(t) = \inf_{(t, \infty)} y(z)$. Because y never runs out of zeros, $u(t) \geq 0$ and $v(t) \leq 0$. We wish to show that $u^* = \lim_{t \rightarrow \infty} u(t) = 0$ and that $v^* = \lim_{t \rightarrow \infty} v(t) = 0$. We allow the functions $u(t)$ and $v(t)$ and the quantities u^* and v^* to be $\pm\infty$ (but they will not be, as subsequent results show).

Find how $v(t)$ determines $u(t)$ by computing an upper bound on y . Let T be any maximum of y . Integrate (8.16) to get

$$y(T) = y(T-1) - \alpha \int_{T-2}^{T-1} (e^{y(z)} - 1) dz + O(T^{-2}), \quad (8.18)$$

again using the weaker, yet sufficiently strong, error form $O(t^{-2})$. To get an upper bound for $y(T)$, we need an upper bound for $y(T-1)$. We already have an equation that contains $y(T-1)$: the frequently cited (8.16), which relates $y(T-1)$ and $\dot{y}(T)$. We must be slightly careful, however. Even though T is a maximum, we cannot use $\dot{y}(T) = 0$, because \dot{y} is

discontinuous. Fortunately, all the discontinuity lives in the error term, so $\dot{y}(T) = O(T^{-2})$. From (8.16), $\dot{y}(T) = -\alpha(e^{y(T-1)} - 1) + O(T^{-2})$. With $\dot{y}(T) = O(T^{-2})$, this equation for $\dot{y}(T)$ simplifies to

$$y(T-1) = \log(1 + O(T^{-2})) = O(T^{-2}). \quad (8.19)$$

Furthermore, $-\alpha \int_{T-2}^{T-1} (e^{y(z)} - 1) dz \leq 1 - e^{v(T-2)}$, because $y \geq v(T-2)$ in the integration range and $\alpha \leq 1$. Therefore, $y(T) \leq 1 - e^{v(T-2)} + O(T^{-2})$. We can apply this same formula at a subsequent maximum—say T_1 —to find that $y(T_1) \leq 1 - e^{v(T_1-2)} + O(T_1^{-2})$. Because $v(t)$ is nondecreasing, and $O(T_1^{-2}) \prec O(T^{-2})$, we have also $y(T_1) \leq 1 - e^{v(T-2)} + O(T^{-2})$. The right side is therefore a bound on u :

$$u(T) \leq 1 - e^{v(T-2)} + O(T^{-2}). \quad (8.20)$$

Note that, even if $v(t) = -\infty$, then $e^{v(T-2)} \leq 0$, and $u(T) \leq 1 + O(T^{-2})$, which is bounded. Because u is nonincreasing, this bound implies that $u(t) < \infty$ for $t > T$ as well.

As $T \rightarrow \infty$, the bound (8.20) becomes $u^* \leq 1 - e^{v^*}$. This limit step is valid as long as y never runs out of maxima. Because y is continuous, between zeros it attains either a maximum or a minimum. However, y might attain only minima beyond a certain point—say t_3 —in which case, we cannot take the limit $T \rightarrow \infty$. In that case, however, we do not need to take the limit, because $y(t) \leq 0$ for $t \geq t_3$, so $u^* = 0$. Thus, either

$$u^* \leq 1 - e^{v^*} \quad \text{or} \quad u^* = 0. \quad (8.21)$$

If y has a minimum at T' , we find, by similar reasoning, that $v(T') \geq 1 - e^{u(T'-2)} + O(T'^{-2})$. Because $u(T'-2)$ is bounded, so is $v(T')$. The relation for $v(T')$ has limit $v^* \geq 1 - e^{u^*}$, as long as y never runs out of minima; if it does run out, then $v^* = 0$. So, either

$$v^* \geq 1 - e^{u^*} \quad \text{or} \quad v^* = 0. \quad (8.22)$$

The inequalities (8.21) and (8.22) imply that, if either u^* or v^* is zero, so is the other, in which case the proof is done. Therefore, in the remainder of the proof, we can neglect the $u^* = 0$ or $v^* = 0$ part of each inequality.

The inequalities then combine into

$$u^* \leq 1 - \exp(1 - e^{u^*}). \quad (8.23)$$

One solution is $u^* = 0$, when the inequality becomes an equality. That solution is unique, because the left side grows faster than the right for any $u^* > 0$, as we can show by taking the derivative with respect to u^* . The right side has derivative $\exp(u^* + 1 - e^{u^*})$. However, $e^{u^*} > 1 + u^*$ for $u^* > 0$, so the derivative is < 1 . Therefore, $u^* = 0$ is the only solution (because $u^* \geq 0$), whereupon $v^* \geq 1 - e^{u^*}$ and $v^* \leq 0$ imply that $v^* = 0$ as well. Thus, $y \rightarrow 0$. ■

This theorem, combined with the one on bounded zeros (Theorem 8.2), results in Theorem 8.4.

Theorem 8.4 If $\alpha \leq 1$, then $y \rightarrow 0$.

Inverting the myriad variable changes from p_n to $y(t)$, we first $y_n = o(1)$; then, $c_n = -\log \log n + o(1)$; then, $l_n = -\log \log n + o(1)$; and, finally, $p_n = e^{l_n} \sim (\log n)^{-1}$. The only subtle point is the use of $l_n - c_n = O(n^{-1/2}) = o(1)$.

I suspect that Theorem 8.4 holds for $\alpha < \pi/2$, rather than for only $\alpha \leq 1$. It is certainly possible to show that y is always bounded, for all $\alpha > 0$. However, in bounding the integrals, we assumed the worst—that $y(t)$ is a step function. However, $y(t)$ can be bounded by a triangle; with this more careful bound, the range of validity of Theorem 8.4 might increase to, say, $\alpha \leq 1.3$. The analysis is messier, and is not necessary for this model of prime numbers. (Compare the proof lengths for $\alpha \leq 1$ and for $\alpha \leq 3/2$ in Theorem 3 of [64].)

8.3.2 Show that y is exponentially small

To refine the estimate for p_n , show that y is exponentially small. If the delay-differential equation (8.16) had no error term, and were linear, then the job would be done, because $z'(t) = -\alpha z(t-1)$ is a standard equation [1, 63], which can be solved with the Laplace transform, and it has only exponentially decaying solutions (for $\alpha < \pi/2$). If we could only linearize (8.16)! We can, roughly because $y \rightarrow 0$, so the nonlinear terms eventually become negligible.

We rewrite (8.16) as

$$\dot{y}(t) = -\alpha \left(\frac{e^{y(t-1)} - 1}{y(t-1)} \right) y(t-1) + O\left(e^{-2^{t-2}}\right), \quad (8.24)$$

As $t \rightarrow \infty$, $y \rightarrow 0$, so the factor in parentheses limits to unity; also, as $t \rightarrow \infty$, the error term limits to zero. Delay-differential equations with asymptotically constant coefficients have been treated extensively by Wright [62], whose results I quote.

Wright's results about such equations will make more sense if we first discuss the solution of the linearized equation [1]:

$$z'(t) = -\alpha z(t-1). \quad (8.25)$$

Try the substitution $z(t) = e^{st}$, which gives

$$se^s + \alpha = 0. \quad (8.26)$$

Making this substitution is equivalent to solving (8.25) with the Laplace transform. Let $\{s_k\}_{k \geq 1}$ be the solutions of (8.26), arranged in order of nonincreasing real part (nonincreasing instead of decreasing, because there can be complex-conjugate solutions). Then,

$$z(t) = \sum A_k(t) e^{s_k t}, \quad (8.27)$$

where the coefficients $A_k(t)$ are polynomials and are determined by the initial conditions. If s_k is an m th-order root of (8.26), then $A_k(t)$ is a polynomial of degree $m-1$, as discussed in [1]; this result is a consequence of the Laplace inversion formula and Cauchy's residue theorem.

The most common case is a first-order root, whereupon $A_k(t)$ is a constant. At a higher-order root, $(se^s + \alpha)' = 0$, so $e^s(s+1) = 0$, which implies that $s = -1$. But $s = -1$ is a root

iff $\alpha = 1/e$ as well. In that case, $m = 2$, so the corresponding coefficient would be linear. Even for $\alpha = 1/e$, only the $s = -1$ root has $m > 1$; all other roots are simple ($m = 1$).

For large enough t , the term corresponding to s_1 (and to s_2 , if s_1 and s_2 are complex conjugates) dominates. If the root is simple, we can write $z(t) \sim O(e^{\sigma_1 t})$, where $s_k = \sigma_k + i\omega_k$; if $\alpha = 1/e$ and $s_1 = -1$, then $z(t) \sim O(e^{(\sigma_1 + \epsilon)t})$, for any $\epsilon > 0$.

For $\alpha < \pi/2$, Wright's [64] Theorem 5 shows that $\sigma_k < 0$ for all k . For the particular case of $\alpha = \log 2$,

$$s_{1,2} \approx -0.571624 \pm 1.08646i, \quad (8.28)$$

which I found using Newton's method. A minute with a calculator will convince you that s_1 satisfies (8.26); however, it's more tricky to show that no other solution has greater real part. Because Wright [64] proves the behavior of the roots for all α , here I give only a graphical argument (for $\alpha = \log 2$), which illustrates the principles.

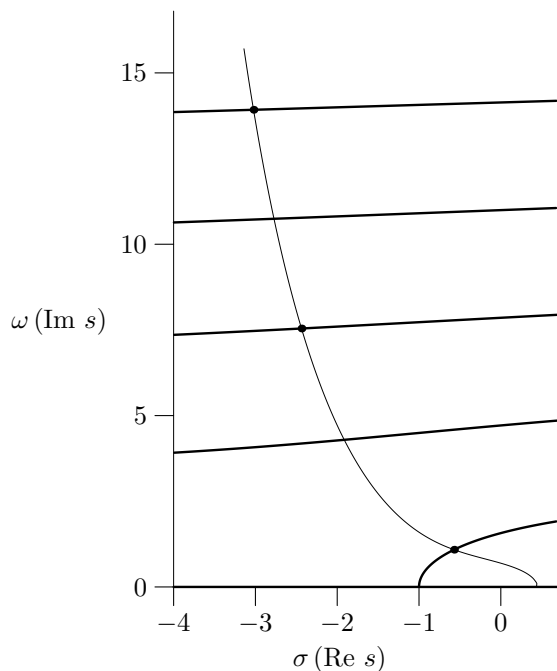


Figure 8.3. The roots of $se^s + \alpha = 0$, shown graphically, for $\alpha = \log 2$. Here $s = \sigma + i\omega$. I have drawn only the $\omega \geq 0$ region, because the $\omega < 0$ region adds no new information. The dotted line is the magnitude constraint (8.29). The set of solid lines (including the σ axis) is the phase constraint (8.30). The magnitude constraint introduces spurious roots—those of $se^s = \alpha$. I have marked the actual roots with a circle. The actual and spurious roots alternate along the dotted line. Note that all roots have $\sigma < 0$ and $\omega \neq 0$, so all solutions of the linear equation (8.25) oscillate and decay exponentially.

If s is a solution, so is \bar{s} , so consider only the case $\omega \geq 0$. The magnitude equation is $(\sigma^2 + \omega^2)e^{2\sigma} = \alpha^2$, or

$$\omega = \sqrt{\alpha^2 e^{-2\sigma} - \sigma^2}. \quad (8.29)$$

The phase equation is $\arg(s) + \arg(e^s) = (2n+1)\pi$, or $\{m\pi + \arctan(\omega/\sigma)\} + \omega = (2n+1)\pi$, where m and n can be any integer. This equation is equivalent to

$$\sigma = -\omega / \tan \omega \quad \text{or} \quad \omega = 0. \quad (8.30)$$

Figure 8.3 plots the magnitude and phase constraints, and shows their simultaneous solution. In taking the magnitude, a calculation that involves squaring, we introduce spurious roots: the roots of $se^s = \alpha$. Get rid of them with a restriction on ω . If $\sigma < 0$, then $\arg(s) \in (\pi/2, \pi]$. Because $\omega = \arg(e^s)$, I find that $\omega = 2n\pi + \theta$, where $\theta \in [0, \pi/2)$. Similarly, if $\sigma \geq 0$, then $\arg(s) \in [0, \pi/2$ and $\omega = 2n\pi + \theta$, where $\theta \in [\pi/2, \pi]$. These restrictions eliminate alternate intersections along the dotted line. Note that all roots have negative real part, so all terms in (8.27) decay exponentially to zero, and that there are roots with arbitrarily negative real part.

Following Wright [64], define the set \mathcal{M} to be the real parts of s_k , with their limit points; and define the characteristic number $\omega_1(y)$ of a solution as the greatest lower bound of the real σ such that

$$\int_1^\infty \{|y(t)|^2 + |\dot{y}(t)|^2\} e^{-2\sigma t} dt \quad (8.31)$$

converges. By Theorem 1 of Wright [62], either $\omega_1(y) \in \mathcal{M}$ or $\omega_1(y) = \omega(\text{error term})$, where $\omega(f)$ is defined by the greatest lower bound of the real σ such that $\int_1^\infty |f(t)|^2 e^{-2\sigma t} dt$ converges. The error term is superexponentially small, so $\omega(\text{error term}) = -\infty$. So, $\omega_1(y) \in \mathcal{M}$, because \mathcal{M} already contains $-\infty$ as one of its limit points. Therefore, $\omega_1(y) \leq \sigma_1$. From (2.1) of [62],

$$y = O(e^{\sigma_1 t}). \quad (8.32)$$

Wright's results justify the natural procedure of replacing the coefficients in (8.24) by their limits, and solving the resulting constant-coefficient equation, to get (8.32).

This more refined result for y in (8.32) becomes

$$p_n = \frac{1}{\log n} \left(1 + O((\log n)^{-0.824679\dots}) \right), \quad (8.33)$$

because $\sigma_1 / \log 2 \approx 0.824679$.

8.3.3 Find an asymptotic expression for y

With y exponentially small, as given by (8.32), the delay-differential equation (8.16) becomes $\dot{y}(t) = -\alpha y(t-1) + O(e^{2\sigma_1 t})$. So we need to solve

$$\dot{y}(t) = -\alpha y(t-1) + f(t), \quad (8.34)$$

where $f(t) = O(e^{2\sigma_1 t})$. We can write any solution of (8.34) as a sum of the particular solution, which is any solution of (8.34) with initial conditions $y = 0$ for $0 \leq t \leq 1$; and of the general solution, which is any solution of $\dot{y}(t) = -\alpha y(t-1)$. The general solution is given by (8.27). It is a series of exponentials, all decaying (and possibly oscillating, depending on α).

We can obtain the asymptotics of the particular solution with the Laplace transform. Multiply (8.34) by e^{-st} , and integrate from $t = 1$ to $t = \infty$ (which is the range of validity of (8.34)):

$$\int_1^\infty \dot{y}(t)e^{-st} dt = -\alpha \int_1^\infty y(t-1)e^{-st} dt + f(s), \quad (8.35)$$

where $f(s) = \int_1^\infty f(t)e^{-st} dt$. The left side becomes (after an integration by parts) $-y(1)e^{-s} + sy(s) = sy(s)$, because $y(1) = 0$ for the particular solution. The integral on the right is

$$e^{-s} \int_0^\infty y(t)e^{-st} dt = e^{-s}y(s), \quad (8.36)$$

because $y(t) = 0$ for $0 \leq t \leq 1$. So (8.35) becomes $sy(s) = -\alpha e^{-s}y(s) + f(s)$, or

$$y(s) = \frac{f(s)}{s + \alpha e^{-s}}. \quad (8.37)$$

Because $y(t) = O(e^{\sigma_1 t})$, the transformed function $y(s)$ is analytic for $\Re s > \sigma_1$. The Laplace inversion formula says that

$$y(t) = \int_{c-i\infty}^{c+i\infty} \frac{f(s)}{s + \alpha e^{-s}} e^{st} ds, \quad (8.38)$$

for any $c > \sigma_1$. Close the contour in the left half-plane (equivalent to assuming positive t), and find that

$$y(t) = \sum B_k(t)e^{q_k t}, \quad (8.39)$$

where q_k are the poles of $y(s)$, arranged in order of nonincreasing real part, and $B_k(t)$ are the residues (which will depend on t for only nonsimple poles). I have used q_k instead of s_k , because the poles of $y(s)$ are not identical to the roots of $se^{-s} + \alpha = 0$ (which are the s_k). Because $f(t) = O(e^{2\sigma_1 t})$, the transformed function $f(s)$ is analytic for $\Re s > 2\sigma_1$; in that region, $q_k = s_k$. For $\alpha = \log 2$, the first two zeros are complex conjugates $\sigma_1 \pm i\omega_1$ given by (8.28). Because y is real,

$$y(t) = A \cos(\omega_1 t + \phi) e^{\sigma_1 t} + O(e^{\sigma_2 t}) + O(e^{2\sigma_1 t}), \quad (8.40)$$

where A and ϕ are determined by the initial conditions (in a complicated way, because the delay-differential equation is nonlinear). This result is the particular solution, but it is also the general solution, as a quick inspection of (8.27) will convince you. For $\alpha = \log 2$, $\sigma_2 \approx -2.43658$, which is $< 2\sigma_1$. So

$$y(t) = A \cos(\omega_1 t + \phi) e^{\sigma_1 t} + O(e^{2\sigma_1 t}). \quad (8.41)$$

Because the error term in (8.16) is superexponentially small, we can get an asymptotic series with superexponentially small error term by modifying the methods of [61]. The result (8.41) suffices for prime numbers. We have now shown Theorem 8.5.

Theorem 8.5 For $\alpha = \log 2$,

$$p_n = \frac{1}{\log n} \left\{ 1 + A(\log n)^{-\beta} \cos(\omega \log \log n + \phi) + O(\log n)^{-2\beta} \right\}, \quad (8.42)$$

where $\beta = -\sigma_1/\log 2 \approx 0.824679$ and $\omega = \omega_1/\log 2 \approx 1.56743$, and where A and ϕ are determined by the initial conditions.

8.4 Analyze the oscillations

The most interesting feature of this model is the oscillations in p_n around $(\log n)^{-1}$. Let's take a closer look at these oscillations, and speculate about their implications for real primes.

8.4.1 Bogus initial conditions

I computed p_n numerically using (8.7), and used several of the data to fit the prediction (8.41) without the error term. Figure 8.4 shows the excellent agreement between the fit and the remainder of the numerical data. Notice that $y(t)$ oscillates around 0 (so p_n oscillates around $(\log n)^{-1}$), with period approximately 6. I had expected the period to be 2, because of the feedback argument in Section 8.1.2. The argument is basically correct, but it was too sloppy, because p_{n^2} depends not just on p_n , but also on p_k for $k < n$.

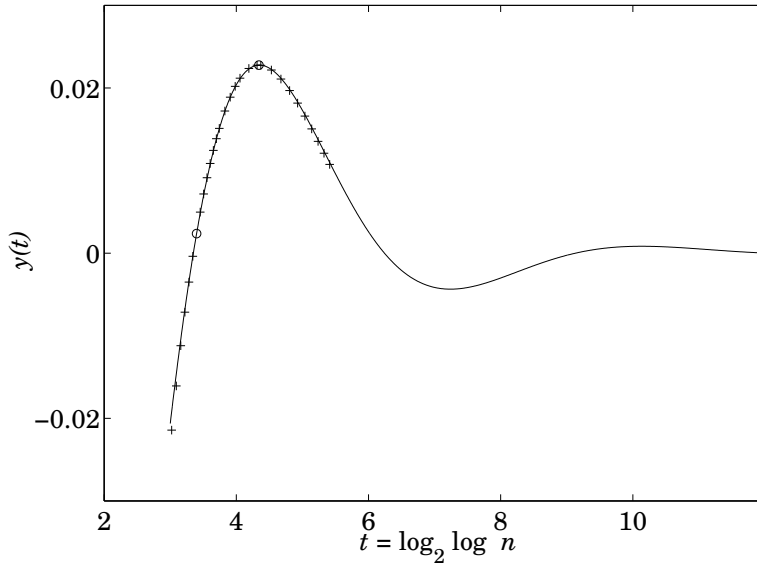


Figure 8.4. Theory and experiment, starting with $p_1 = 0$ and $p_2 = 1$. With those initial conditions, I computed p_n numerically using (8.7), for various n up to 1.15×10^{19} , and thereby computed $y(t)$. I fit the dominant term in the solution, $Ae^{\sigma_1 t} \cos(\omega_1 t + \phi)$, using the two data points marked with circles. The only free parameters in the fit are A and ϕ , whose fitted values are $A = 0.3091$ and $\phi = 1.0733$. Note that the fitted curve goes through the remainder of the numerical data, indicating that, by $t = 3$, the transients have died out.

8.4.2 More realistic initial conditions

The $y(t)$ in Figure 8.4 does not accurately model the density of actual primes, because the initial conditions ($p_1 = 0$ and $p_2 = 1$) are lousy estimates of the density \hat{p}_n ; there are not enough data at low n to define \hat{p}_n accurately. So I estimated \hat{p}_n for actual primes, for various $n \in [10^{4.5}, 10^9]$. I used those estimates as the initial $y(t)$ and extrapolated y using (8.16), without the error term. The resulting curve is shown in Figure 8.5. I estimated $\hat{p}(P_m)$ as $k/(P_{m+k} - P_m)$, where P_m is the m th prime, and I chose $k = \max(1000, \lfloor m/100 \rfloor)$. (I

have temporarily used the notation $p(f(n))$ instead of $p_{f(n)}$, because $p(P_n)$ written with subscripts is hard to distinguish from pP_n .) A reasonable alternative would be the estimate $\hat{p}_n \approx k/(\pi(n+k) - \pi(n))$.

The scatter in \hat{p}_n is small when n is large, because k becomes huge, and the estimate becomes reasonably reliable. I could not use such a huge k for the smaller n , because that would introduce too much error in determining \hat{p}_n : I would not know whether I was determining \hat{p}_{n+k} or \hat{p}_n .

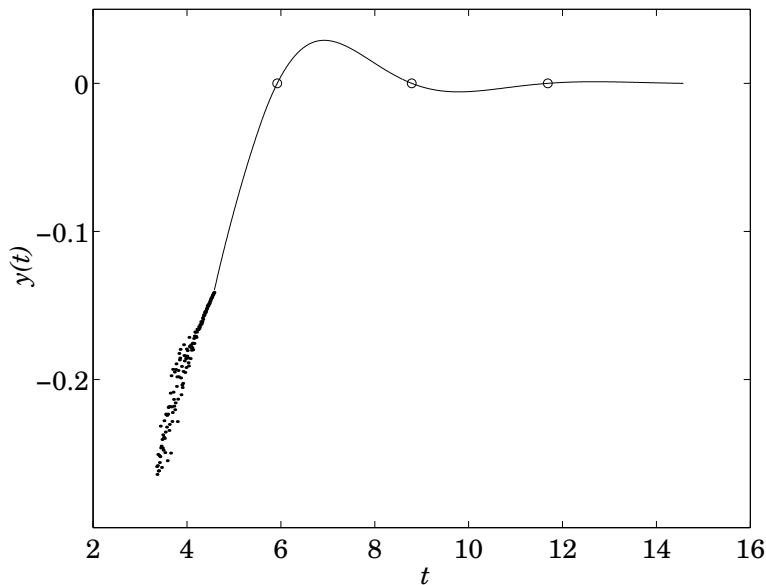


Figure 8.5. Predicted density, initialized using actual density. I estimated \hat{p}_n for actual primes, using the method described in the text. This computation gave $y(t)$ for $3.35 \leq t \leq 4.57$ (the points). I computed the curve for $y(t)$ using (8.16) without the error term, and marked the zero crossings with circles. The first zero crossing in $y(t)$ is located at $t \approx 5.92$, or $n \approx 2 \times 10^{26}$. The cumulative distribution, $\int p_n dn$, crosses $\text{Li}(n)$ around 6×10^{26} .

8.4.3 Predictions

We can use the square-root model to make predictions about the density of actual primes. According to Ribenboim [50], Gauss and Riemann believed that $\pi(n) < \text{Li}(n)$ for all $n > 3$. Their conjecture is reasonable: For all $\pi(n)$ thus far computed, $\pi(n) < \text{Li}(n)$ (also according to Ribenboim's fascinating book [50]). However, Littlewood [39] showed that $\pi(n)$ crosses $\text{Li}(n)$ infinitely often, but he did not show where the first crossing was. By computing a huge number of zeros of the zeta function, de Riele [51] showed that $\pi(n) > \text{Li}(n)$ for some $n < 7 \times 10^{370}$.

In the square-root model, p_n oscillates around $(\log n)^{-1}$, and $\bar{\pi}(n) = \int_2^n p_n dn$ oscillates around $\text{Li}(n)$, where $\bar{\pi}(n)$ is the estimate of $\pi(n)$ given by the square-root model. From Figure 8.5, $\bar{\pi}(n) < \text{Li}(n)$ until n gets huge (remember that $n = e^{2^t}$). How huge? The

function $y(t)$ crosses zero for the first time at $t \approx 5.92$, or $n \approx 2 \times 10^{26}$; $\bar{\pi}(x)$ crosses $\text{Li}(x)$ a bit later, at $n \approx 6 \times 10^{26}$ (I found this crossing using numerical integration), because it takes a while for the excess density to overcome the accumulated deficit. I suggest looking near 6×10^{26} for sign changes in $\pi(n) - \text{Li}(n)$. This estimate is sensitive to the initial conditions, so “near 10^{27} ” is probably all the accuracy that I can claim. The current record [50] for computing $\pi(n)$ is $\pi(10^{18})$, so an exact calculation of $\pi(10^{27})$ may be unattainable for another decade. However, an accurate approximation to $\pi(n)$ may suffice to show that $\pi(n) - \text{Li}(n) > 0$, because $\text{Li}(n)$ can be approximated accurately with an asymptotic series.

One argument against this prediction (and against the square-root model) is the kink in $y(t)$ where the estimated densities end and the theoretical extrapolation begins. I have not found an explanation for this behavior.

8.5 Speculate on the Riemann hypothesis

Section 8.4.3 described a prediction of the square-root model for properties of $\pi(n)$. The model also suggests, heuristically, other properties of actual primes. For sequences generated by the Hawkins’ sieve, a random analogue of the Riemann hypothesis holds [43]. Suppose that this analogue also holds for these R-prime sequences, with their oscillations. That result would not be so interesting. If the analogue did not hold, however, we would have a reason to doubt, although not a disproof of the actual Riemann hypothesis. The failure would suggest that oscillations in the density of actual primes—an important feature of the square-root model—perturb the density enough to invalidate the Riemann hypothesis.

8.6 Go beyond the square root

The original reason for the square-root model was the too-high prediction for p_n in (8.1), because actual primes are more efficient at sieving than random divisibility would predict. We fixed the excess by making primality random. An alternative fix is to extend the range of the product, as Pólya [47] suggested:

$$p_n = \prod_{p \text{ prime}}^{\lfloor n^{1/r} \rfloor} \left(1 - \frac{1}{p}\right), \quad (8.43)$$

where $r = e^\gamma \approx 1.781$. Now, by Mertens’ theorem, $p_n \sim (\log n)^{-1}$.

The success of this alteration makes me wonder what would happen if we combine it with random primality. Then,

$$p_n = \prod_1^{\lfloor n^{1/r} \rfloor} \left(1 - \frac{p_k}{k}\right). \quad (8.44)$$

The square-root model has $r = 2$ and $t = \log_2 \log n$. With the definition $t = \log_r \log n$, the product (8.44) results in the same delay-differential equation (8.16), but with $\alpha = \log r$. For $r = e^\gamma$, the dominant solution of the characteristic equation (8.26) is $s_{1,2} = -0.696647 \pm i 0.925611$. The density approaches $(\log n)^{-1}$, and oscillates as for the $r = 2$ case. Perhaps a generalized random-primality model, for some $r \neq 2$, models actual primes more accurately than the $r = 2$ model does.

Wild behavior occurs if r is too large. Then, one solution of the characteristic equation $se^s + \alpha = 0$ has *positive* real part, and the oscillations around $(\log n)^{-1}$ grow without bound. (Actually, they would grow if the delay–differential equation (8.16) were linear; the nonlinear terms keep y bounded.) This qualitative change happens when the solution s_1 has zero real part, at $s = i\omega_1$. So $i\omega_1 e^{i\omega_1} + \alpha = 0$, or $\omega_1 = (n + 1/2)\pi$, for some integer n . The smallest α allowed under this restriction is $\alpha = \pi/2$, or $r \approx 4.81$. At this value of α , the system is critically damped ($r = 2$ is in the underdamped regime; your car’s shock absorbers are tuned for critical damping). In a world where numbers are either prime or split into at least five factors, we would need to check R-primes up to only $n^{1/5}$. In such a world, the R-prime-number theorem (that $p_n \sim (\log n)^{-1}$) would no longer be true; perhaps the actual prime-number theorem would not be valid either.

The same instability occurs in negative-feedback electric circuits. Such a circuit becomes unstable if the loop gain is too large or the delay too long (Nyquist’s theorem [57]); either condition turns negative into positive feedback. These primality models also have negative feedback and a delay. With t as the independent variable, the delay is always unity. The parameter analogous to gain is α , the coefficient in front of the integral in (8.16), and it increases as r increases. When $\alpha > \pi/2$, the loop becomes unstable. Equivalently, with n as the independent variable, the delay—from $n^{1/r}$ to n —increases as r increases. For large enough r , the loop is unstable. In a linear system, the oscillations would grow without bound. This model is nonlinear; when y becomes sufficiently large, nonlinear terms take effect. In this model, the nonlinear terms probably prevent y from growing without bound (however, I have not proved this statement).

The Hawkins’ model, $p_{n+1} = p_n(1 - p_n/n)$, is the overdamped extreme. The delay is almost exactly zero, because $n + 1$ and n hardly differ for large n . This low delay leaves no time for oscillations to arise, and the density converges monotonically to $(\log n)^{-1}$, as shown in [21]. Equivalently, we can take $\alpha \rightarrow 0$ in the characteristic equation (8.26). In that regime (actually, for $\alpha < e^{1/e}$), the zeros are purely real and negative: y decays without oscillating.

8.7 Lower the temperature

R-primes could be generated by a noisy computer. Take a computer that computes remainders properly, and use the sieve to find primes. The computer generates just the sequence of actual primes. Now pour acid on the CPU, in a way that damages only the circuitry that computes $a \bmod b$, and makes the result random. Then, the computer generates R-prime sequences, as described in Section 8.1.2.

What if the acid damages the CPU only partly (dump, say, vinegar instead of nitric acid)? Then the computer generates sequences that partly resemble the sequence of actual primes, and partly resemble sequences of R-primes. One way to parameterize some of these intermediate models is to borrow the idea of temperature from statistical mechanics [49]. Suppose that the computer has to compute $a \bmod b$, with M (a random variable) as the result. Let M be correct with relative probability 1, and let each wrong answer have relative probability $e^{-1/T}$, where T is the temperature. Then $p_{\text{right}} = 1/Z$, and the probability of each of the wrong answers is $e^{-1/T}/Z$, where $Z = 1 + (b - 1)e^{-1/T}$. I call this normalizer “ Z ”, in analogy with the partition function from statistical mechanics. I do not know how

well that part of the temperature analogy holds. At $T = 0$ (the frozen state), the computer produces the sequence of actual primes. At $T = \infty$ (the completely melted state), the computer produces a sequence of R-primes (a different sequence on each run). What happens at intermediate temperatures? We could study $p_n(T)$, even if only for high T , and obtain a power series in $1/T$, similar to the high-temperature expansions in statistical mechanics [66]. Perhaps the first few terms of $p_n(T)$ would suggest properties of actual primes. For example, let

$$p_n(T) = f_0(n) + f_1(n)T^{-1} + \cdots. \quad (8.45)$$

If $f_1(n)$ is positive, then probably n is prime. I doubt that it's possible to determine f_1 with such resolution, but less accurate bounds on f_1 may still result in rigorous statements about the distribution of actual primes.

8.8 Finish

We studied the square-root model, a probabilistic model of primality that is similar to the Hawkins' random sieve. The model results in a density $p_n \sim (\log n)^{-1}$, as expected by the prime-number theorem, but it also exhibits oscillations around that density, which I believe are related to the crossings of $\pi(n)$ and $\text{Li}(n)$ discovered by Littlewood. The analysis of the square-root model leads to the conjecture that $\pi(n)$ first crosses $\text{Li}(n)$ near $n \sim 10^{27}$.

BIBLIOGRAPHY

- [1] Richard Bellman and Kenneth L. Cooke. *Differential–Difference Equations*. Academic Press, New York, 1963.
- [2] Howard C. Berg. *Random Walks in Biology*. Princeton University Press, Princeton, NJ, 1993.
- [3] Carl Bernstein and Bob Woodward. *All the President’s Men*. Simon and Schuster, New York, 1974.
- [4] P. W. Bridgman. *Dimensional Analysis*. Yale University Press, New Haven, Conn., 1931.
- [5] Peter D. Calvert, Theresa W. Ho, Yvette M. LeFebvre, and Vadim Y. Arshavsky. Onset of feedback reactions underlying vertebrate rod photoreceptor light adaptation. *Journal of General Physiology*, 111:39–51, 1998.
- [6] S. Chandrasekhar. *An Introduction to the Study of Stellar Structure*. University of Chicago Press, Chicago, Ill., 1939.
- [7] S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability*. Oxford University Press, Oxford, 1961.
- [8] Frank S. Crawford. *Waves*, volume 3 of *Berkeley Physics Course*. McGraw-Hill, New York, 1968.
- [9] G. Hoffman de Visme. The density of prime numbers. *Mathematical Gazette*, 45:13–14, 1961.
- [10] Peter B. Detwiler and Mark P. Gray-Keller. The mechanisms of vertebrate light adaptation: Speeded recovery versus slowed activation. *Current Opinion in Neurobiology*, 6:440–444, 1996.
- [11] Kristian Donner, Ari Koskelainen, Kaj Djupsund, and Simo Hemilä. Changes in retinal time scale under background light: Observations on rods and ganglion cells in the frog retina. *Vision Research*, 35:2255–2266, 1995.
- [12] Lyn Dupré. *BUGS in Writing: A Guide to Debugging Your Prose*. Addison–Wesley, Reading, Mass., revised edition, 1998.
- [13] T. E. Faber. *Fluid Dynamics for Physicists*. Cambridge University Press, Cambridge, England, 1995.
- [14] G. T. Fechner. *Elemente der Psychophysik*. Breitkopf and Härtl, Leipzig, Germany, 1860.
- [15] Richard Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics*, volume I. Addison-Wesley, Reading, MA, 1963.

- [16] William Elmer Forsythe. *Smithsonian Physical Tables*, volume 120 of *Smithsonian Miscellaneous Collections*. Smithsonian Institution, Washington, DC, ninth revised edition, 1964.
- [17] H. W. Fowler and F. G. Fowler. *The King's English*. Oxford University Press, Oxford, 3rd edition, 1931. Reprinted 1962.
- [18] Stephen Gasiorowicz. *Quantum Physics*. Wiley, New York, second edition, 1996.
- [19] G. H. Hardy and J. E. Littlewood. Some problems of “partitio numerorum:” III: on the expression of a number as a sum of primes. *Acta Mathematica*, 44:1–70, 1922.
- [20] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, Oxford, third edition, 1954.
- [21] D. Hawkins. The random sieve. *Mathematics Magazine*, 31:1–3, 1958.
- [22] D. Hawkins. Random sieves, II. *Journal of Number Theory*, 6:192–200, 1974.
- [23] David Hubel. *Eye, Brain, and Vision*. Scientific American Library, New York, 1988.
- [24] Kazuhisa Ichikawa. Critical processes which characterize the photocurrent of retinal rod outer segments to flash stimuli. *Neuroscience Research*, 19:201–212, 1994.
- [25] Adelbert Ames III. Steady state feedback in mammalian phototransduction illustrated by a nomogram. *Vision Research*, 6:821–827, 1994.
- [26] D. C. Ipsen. *Units, Dimensions, and Dimensionless Numbers*. McGraw-Hill, New York, 1960.
- [27] Gregor J. Jones. Light adaptation and the rising phase of the flash photocurrent of salamander retinal rods. *Journal of Physiology*, 487:441–451, 1995.
- [28] Satoru Kawamura and Motohiko Murakami. Calcium-dependent regulation of cyclic GMP phosphodiesterase by a protein from frog retinal rods. *Nature*, 349:420–423, 1991.
- [29] K. W. Koch and L. Stryer. Highly cooperative feedback control of retinal rod guanylate cyclase by calcium ions. *Nature*, 334:64–66, 1988.
- [30] Y. Koutalos, K. Nakatani, T. Tamura, and K.-W. Yau. Characterization of guanylate cyclase activity in single retinal rod outer segments. *Journal of General Physiology*, 106:863–890, 1995.
- [31] Y. Koutalos, K. Nakatani, and K.-W. Yau. The cGMP-phosphodiesterase and its contribution to the sensitivity regulation in retinal rods. *Journal of General Physiology*, 106:891–921, 1995.
- [32] Yiannis Koutalos and King-Wai Yau. Regulation of sensitivity in vertebrate rod photoreceptors by calcium. *Trends in Neurosciences*, 19:73–81, 1996.
- [33] Leon Lagnado and Denis Baylor. Calcium controls light-triggered formation of catalytically active rhodopsin. *Nature*, 367:273–277, 1994.

- [34] Leon Lagnado, L. Cervetto, and P. A. McNaughton. Calcium homeostasis in the outer segments of retinal rods from the tiger salamander. *Journal of Physiology*, 455:111–142, 1992.
- [35] Horace Lamb. *Hydrodynamics*. Dover, New York, sixth edition, 1945.
- [36] T. D. Lamb and E. N. Pugh Jr. G-protein cascades: Gain and kinetics. *Trends in Neurosciences*, 15:291–298, 1992.
- [37] T. D. Lamb and E. N. Pugh Jr. A quantitative account of the activation steps involved in phototransduction in amphibian photoreceptors. *Journal of Physiology*, 449:719–758, 1992.
- [38] David R. Lide, editor. *CRC Handbook of Chemistry and Physics: A Ready-Reference Book of Chemical and Physical Data*. CRC Press, Boca Raton, FL, 74th edition, 1993.
- [39] J. E. Littlewood. Sur la distribution des nombres premiers. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, 158:1869–1872, 1914.
- [40] H. R. Matthews. The role of cytoplasmic calcium in photoreceptor light adaptation. *Journal of Physiology* (Paris), 86:146–155, 1992.
- [41] H. R. Matthews. Static and dynamic actions of cytoplasmic Ca^{2+} in the adaptation of responses to saturating flashes in salamander rods. *Journal of Physiology*, 490:1–15, 1996.
- [42] H. R. Matthews, R. L. W. Murphy, G. L. Fain, and T. D. Lamb. Photoreceptor light adaptation is mediated by cytoplasmic calcium concentration. *Nature*, 334:67–69, 1988.
- [43] Werner Neudecker and David Williams. The ‘Riemann hypothesis’ for the Hawkins random sieve. *Compositio Mathematica*, 29:197–200, 1974.
- [44] S. Nikonov, N. Engheta, and E. N. Pugh Jr. Kinetics of recovery of the dark-adapted salamander rod photoresponse. *Journal of General Physiology*, 111:7–37, 1998.
- [45] R. C. Pankhurst. *Dimensional Analysis and Scale Factors*. Chapman and Hall, London, 1964.
- [46] D. R. Pepperberg, M. C. Cornwall, M. Kahlert, K. P. Hofmann, J. Jin, G. J. Jones, and H. Ripps. Light-dependent delay in the falling phase of the retinal rod photoresponse. *Visual Neuroscience*, 8:9–18, 1992.
- [47] G. Pólya. Heuristic reasoning in the theory of numbers. *American Mathematical Monthly*, 66:375–384, 1959.
- [48] E. N. Pugh and T. D. Lamb. Amplification and kinetics of the activation steps in phototransduction. *Biochimica et Biophysica Acta*, 1141:111–149, 1993.
- [49] F. Reif. *Fundamentals of Statistical and Thermal Physics*. McGraw-Hill, New York, 1965.
- [50] Paulo Ribenboim. *The New Book of Prime Number Records*. Springer-Verlag, New York, 3rd edition, 1996.

- [51] H. J. J. te Riele. On the sign of the difference $\pi(x) - \text{li}(x)$. *Mathematics of Computation*, 48:323–328, 1987.
- [52] William L. Riordan. *Plunkitt of Tammany Hall*. E. P. Dutton, New York, 1963.
- [53] Bertrand Russell. *A History of Western Philosophy, and Its Connection with Political and Social Circumstances from the Earliest Times to the Present Day*. Simon and Schuster, New York, 1945.
- [54] Knut Schmid-Nielsen. *Scaling: Why Animal Size is So Important*. Cambridge University Press, Cambridge, England, 1984.
- [55] Robert Shapley, Ehud Kaplan, and Keith Purpura. Contrast sensitivity and light adaptation in photoreceptors or in the retinal network. In Robert Shapley and Dominic Man-Kit Lam, editors, *Contrast Sensitivity*, volume 5 of *Proceedings of the Retina Research Foundation Symposia*, chapter 7, pages 103–116. MIT Press, Cambridge, Mass., 1993.
- [56] William M. Siebert. *Circuits, Signals, and Systems*. MIT Press, Cambridge, Mass., 1985.
- [57] William M. Siebert. *Circuits, Signals, and Systems*. McGraw-Hill, New York, 1986.
- [58] T. Tamura, K. Nakatani, and K.-W. Yau. Calcium feedback and sensitivity regulation in primate rods. *Journal of General Physiology*, 98:95–130, 1991.
- [59] Steven Vogel. *Life in Moving Fluids: The Physical Biology of Flow*. Princeton University Press, Princeton, NJ, 1994.
- [60] E. H. Weber. De pulsu, resorptione, auditu et tactu (on stimulation, response, hearing and touch). *Annotationes Anatomicae et Physiologicae*, 1:1–75, 1834.
- [61] E. M. Wright. The non-linear difference–differential equation. *Quarterly Journal of Mathematics*, 17:245–252, 1946.
- [62] E. M. Wright. The linear difference–differential equation with asymptotically constant coefficients. *American Journal of Mathematics*, 70:221–238, 1948.
- [63] E. M. Wright. The linear difference–differential equation with constant coefficients. *Proceedings of the Royal Society of Edinburgh*, 62:387–393, 1949.
- [64] E. M. Wright. A non-linear difference–differential equation. *Journal für die Reine und Angewandete Mathematik*, 194:66–87, 1955.
- [65] E. M. Wright. A functional equation in the heuristic theory of primes. *Mathematical Gazette*, 45:15–16, 1961.
- [66] J. M. Yeomans. *Statistical Mechanics of Phase Transitions*. Oxford University Press, Oxford, 1992.