Notes & Problems for GEOL 0350: Mathematical Methods of Fluid and Solid Geophysics and Geology

Baylor Fox-Kemper
Brown University Dept. of Earth, Environmental, and Planetary Sciences

September 9, 2019

Contents

Syllabus

0.1 Course Descriptions

0.1.1 GEOL0350 Mathematics of Fluid and Solid Geophysics and Geology

Intended for undergraduates conalphcentrating in geological and physical sciences or engineering, especially those interested in the quantitative study of Earth. Problem sets will cover common approaches to quantify the dynamics and chemistry of solids and fluids in nature. Mathematical topics to be introduced include linear algebra, vectors and tensors, differential equations, dynamical systems, eigenvalues and eigenvectors, empirical orthogonal functions, fractals, chaos, and statistics. Applications include waves in the oceans, atmosphere, and solid earth, convective and conductive heat flow, reaction rates, gravitational potential energy, Newtons laws on a rotating planet, measuring coastlines and ranges, and dating errors in stratigraphy. Prerequisites: GEOL 0220 and (MATH 0100, 0170, 0190, 0180, 0350 or 0200).

0.2 Contacts

Portions of the website are password-protected to ensure that fair use and copyrights are correctly obeyed as I share images from books, etc. You can access these by using:

```
username:  io
password:  ocean
```

0.3 Getting Help!

We are usually available by email. Baylor's office hours will be Monday 1:30-2:30 and Thursday 2-3 or by appointment (see my schedule at http://fox-kemper.com/contact). You can also drop into the Math Resource Center (MRC, http://www.math.brown.edu/mrc/) or sign up or drop in to a tutoring session (http://www.brown.edu/academics/college/support/tutor).

0.4 Meetings and Places

We will meet Monday, Wednesday, and Fridays from 10:00 to 10:50AM in GeoChem 0390. Baylor's office hours will be Monday 1:30-2:30 and Thursday 2-3 or by appointment (see my schedule at http://fox-kemper.com/contact) in his office (GeoChem 133) or lab (GeoChem 134). You should also have signed up for a lab/practicum/practice session for one of MTWR 3:00-3:50.

0.5 Website and Canvas

The primary resource for this class is the webpage: http://fox-kemper.com/0350. The class webpage is where all of your assignments will be announced, solution sets posted, links to additional reading will be posted, etc. Assignments should be turned in as pdfs using canvas. The copiers in GeoChem and elsewhere can be used to scan handwritten assignments (for free).

You will want to familiarize yourself with Wolfram Alpha (http://www.wolframalpha.com), it is a great resource for looking up math definitions. Wikipedia is also handy in a pinch (due to the armies of math & physics grad students who apparently have so very few social commitments that they punch in all the details of their dissertation appendices).

0.6 Required Course Activities, Expected Times, and Structure of Classtime

The regular class time will be presentation of new materials and discussion. This format requires buy-in from you, the student, however. You *must do the reading of the notes before class*, and preferably also *at least skim the associated chapters in the book* before class. If an individual student fails to do this, it will negatively influence her or his ability to follow, and if the class fails to do this I will have to expand the lecture mode–decreasing the problem-solving mode–which is not good for learning. In addition, you will visit the lab/practicum/practice sessions once a week (you signed up for one day from MTWR at 3-3:50), where you can work on homework problems in small groups, with input from the TA. The TA will be scheduling students from GEOL2300 to come to discuss your understanding and help with homework problems during this time. You should provide feedback to the TA or the professors about these interactions, because the GEOL2300 students are being assessed on how well they interact with you.

Magdalene Lampert, a researcher in math education, has shown that learning and retention in mathematical methods is improved by inverting the common classroom presentation order. Lecture, followed by discussion, followed by individual homework is not as effective as individual effort, group effort, full discussion. We will use the discussions and the practicum sessions to adhere to the latter format as best as possible.

Your individual effort begins with reading the notes and skimming the chapter before lectures. Then you will be challenged with questions throughout the class and practicum. In the practicum sessions, you will work individually for a few minutes, and then discuss in a group for a few minutes, and ask the leader for guidance or clarification. Finally, the whole class will discuss approaches to problems and the correct solutions. You will then review these problems again as you review the chapter reading and finalize your (related) homework problems, and study for exams.

0.6.1 Assignments, Exams, and Expected Time for Activities

- Scheduled class meetings, which will be suspended in the Reading Period (3 hours/week; 38 hours) and practicum meetings, which will continue in the Reading Period (1 hours/week; 13 hours) [Grading: 10% Attendance and participation.]
- Reading and reviewing class work (2 hours/week; 26 hours)
- Weekly assignments (6 hours/week for 12 weeks; 72 hours) [Grading: 50% Weekly homework]
- Weekly peer reviews (1 hours/week for 12 weeks; 12 hours) [Grading: 10% Reviews of other students' homework assignments.]
- Preparation for Midterm and Final (16 hours) [Grading: 20% Final, 10% Midterm]
- Final Exam (3 hours)
- Total: 180 hours [Grading: 100%]

What can I do to get a good grade? Turn in all of the assignments on time. For the format of the course to work, ON TIME matters, so that we can get to the reviewing. Also, *BONUS POINTS* are available on homework and exams for spotting typos in the notes, homework assignments, and exam problems. The more promptly you point them out (by email), and the more important they are, the more points you get!

The scheduling of the assignments are listed on the webpage, and other than the exceptional weeks around holidays will be as follows.

- Weekly assignment due by class time on Monday.
- Solution sets distributed by midnight Monday (assignments not accepted afterward).
- Peer reviewing and grading due by following Friday.
- *Iterum usque ad finem*

All of this will be charted out on the calendar on the website and in canvas.

Peer review

In addition to doing the problem sets, you will each be performing reviews of each others work. We will be using a rubric based on the AGU guidelines for review. A-F for presentation quality and 1-5 for science/math. Such a guide is useful to go by, and when you do reviews of your fellow students, I'll expect to get a A1 or B2 or B1 score, etc. An A1 will count for 100%, and presentation and accuracy will be equally weighted (an F5 will be 20%). There are a few lessons to be learned here, that will help you write your own papers and will help you provide effective and useful reviews in your career.

- Learning to spot unfounded claims
- Learning how to properly support claims
- Learning to distinguish poor writing/presentation from poor thinking
- Learning to label equations, graphs, and numerical information understandably

- Revisiting problems from a different perspective

You will have each of your homework assignments peer-reviewed by more than one person, and inconsistent results will be rechecked. The assignments for reviewers will rotate (ensuring fairness in grading by randomization). You should feel free to contact me with any concerns about the process or specific issues.

0.6.2 Calendar

The main webpage for the class http://fox-kemper.com/0350 will have the calendar with all assignment deadlines, readings, etc. set up by the first class session. There will be weekly problem sets, one midterm, and a final exam.

0.7 Goals

In this class you will:

- Learn how to quantify some of the physical processes of the earth system.
- Learn how observations and budgets are quantified, evaluated and quality-controlled, and compared.
- Get practice solving diverse geophysical and geological problems using new mathematical techniques.
- Gain a broader perspective and more practice by peer reviewing and collaborating.

This class cannot possibly provide a complete understanding of all of the mathematical topics presented, instead the goal is to introduce the most basic ideas and give geophysical and geological examples where the mathematical tools are useful. A key goal is to introduce the mathematical language, so that students can better choose later mathematics classes and look up mathematical concepts on their own (e.g., using Wolfram Alpha).

0.8 Textbooks and Software

We will work from the course notes. There is not a required textbook, although for reference you should familiarize yourself with a copy of something like Boas (2006) or Arfken et al. (2013). You might also check out Wilks (2011), and Snieder (2004), which are in the library. Arfken et al. (2013), which is similar to Boas, is available electronically through the Brown Library.

We will solve problems drawn from many geophysics and geology textbooks (LeBlond and Mysak, 1978; Turcotte, 1997; Schubert et al., 2001; Turcotte and Schubert, 2002; Aki and Richards, 2002; Drazin and Reid, 2004; Holton, 2004; Snieder, 2004; McWilliams, 2006; Vallis, 2006; Marshall and Plumb, 2008; Cushman-Roisin and Beckers, 2010; Fowler, 2011; Kaper and Engler, 2013; Bourguignon et al., 2015), but these books are not required for the course. If electronic copies of them are available at Brown, I have added an url to the bibliography here and on the course website. Sufficient background will be provided along with each problem so that no further reading will be required. You may want to use software, which is allowed for homework (although not required and

you must still be able to explain your work without the program). I strongly recommend Matlab and Mathematica, but there are lots of others.

0.8.1 Applications

Geophysical and geological applications touched on in this class are:

- Global Energy Balance
  - Ice Ages
  - Energy Balance Models
- Data constrained models and maps
  - Climate Variability Patterns
  - Stochastic versus Deterministic Variability
- Waves and Oscillations
  - Ocean Waves, Tides, and Tsunamis
  - Earthquakes and Seismic Waves
  - Diurnal, Seasonal, and Orbital Variation Cycles
  - Dispersive Wave Kinematics: Phase & Group velocity
- Transport Budgets
  - Diffusion and Advection
  - Heat transfer
  - Tracers in Fluids
  - Rheology
- Boundary Layers
- Landscape Evolution
- Flows
  - Oceanic
  - Atmospheric
  - Groundwater
  - Mantle Convection
- Chemical Reactions, Rates, and Equilibria
- Mechanics
  - of Solids
  - of Fluids

- Gravity

  Potential and Conservative Forces

- Stratigraphy

  Dating and errors

  Mapping

0.8.2 Math Tools & Critical Concepts

A list of the mathematical topics to be touched on in this class, and associated critical concepts:

- Review of Mathematical Preliminaries (1.5 Weeks)

  Series and Sequences

  Real, Imaginary, Complex

  Trigonometry

  Exponentials and Logarithms

  Units and Dimensions

  Derivatives and Integrals

- Linear Algebra (2 Weeks)

  Vector Spaces

  Matrices and Linear Equations

  Bases and Orthogonality; Rank; Null Space and Span

  Inverse Methods

  Eigenvalues and Eigenvectors; Singular Value Decomposition/Empirical Orthogonal Functions/Principal Component Analysis

- Multivariate Calculus and Differential Geometry (2 Weeks)

  Vectors

    Coordinate Transformations

    Rotation & Reflection, Angular Momentum, and Vorticity

    Vector Differentiation and Integration: Div, Grad, Curl; Gauss, Green & Stokes

  Tensors

    Cartesian Tensors

    Inner and Outer Products versus Matrix Multiplication

    Symmetries: Principle of Tensor Covariance, Tensor Invariants, & Anisotropy

Curvilinear Coordinates and Transforms, especially Spherical Coordinates

Deriving Calculus Identities from Tensor Symmetries

- Differential Equations

  Linear Ordinary Differential Equations and Dynamical Systems (2 Weeks)

  Rate equations

  First and Second Order Equations

  Homogeneous and Inhomogeneous Equations

  Linear and Nonlinear Equations

  Series Solutions: Perturbation Analysis, Asymptotics, and Linearization

  Sturm-Liouville Problems: Free Modes of Oscillation, Superposition

  Time Series and Fourier Analysis

  Linear Partial Differential Equations and Dynamical Systems (2 Weeks)

  Boundary and Initial Value Problems

  Separation of Variables

  Laplace and Poisson Equations (applications of Elliptic PDEs)

  Heat Flow and Wave Equations (applications of Elliptic and Hyberbolic PDEs)

  Separation of variables in linear wave problems: Cramer's Rule and Oscillation Modes

  Decompositions: Helmholtz Streamfunction and Potential, Toroidal and Polloidal, Polarization

- Chaos and Nonlinear Dynamics (1 Week)
- Probability and Statistics (1.5 Weeks)

0.9 Policies

0.9.1 Deadlines

Because of the reviewing process, the scheduling of assignments is tight. Thus, I will have to insist that all problem sets be turned in on time. If they are late, they will drop a letter grade. If they are really late (so that they mess up the next step in the reviewing process) they will be counted as missed and can not be made up. If you foresee that there are big problems coming up (medical, family, etc.) let me know *before* an assignment is due and we can figure something out.

0.9.2 Collaboration

I encourage you to work together, and I do not mind at all if you have similar problem sets or share figures or computer code. However, in this case, I want you to list all of your study group on

each homework assignment (so I can avoid you peer-reviewing your group). You are all required to submit a version of each assignment as first author (that is, one that you wrote yourself), so don't submit identical versions of a problem. You need to be careful to cite your colleagues or the textbooks, websites, or papers you might be working from.

0.9.3 Miscellany

- Attendance is expected. If you will miss a class, please let me know when and why so I can be sure you'll get any announcements, etc.
- Clothing and behavior (e.g., cell & laptop use) should be appropriate for a learning environment.
- Discrimination and harassment will not be tolerated.
- Please contact me if you have any disabilities that require accommodation.

Chapter 1

Series Expansions

1.1 Why Start Here?

There are no exact relationships, exact measurements, or exact theories in science. Instead, each generation of scientists examine the relationships they can find and quantify them with working definitions and units to suit the best techniques of their era. Through observation or by experiment, this study leads them to hypothesize, theorize and construct models that share or approximate similar relationships among variables. Field science is never exact, but it is always quantitative.

Mathematics, on the other hand, is often expressed in exact relationships, precise definitions and proofs, and exact solutions. This work is challenging, and often just proving that a solution exists is the goal. Even probability and statistics, where variables are "random", is often built using equalities among random variables and definitions that only become true after an infinite number of experiments.

One way to build an increasingly accurate theory is to begin with a first guess, then add dependence on more variables and larger and larger excursions of each variable from their initially guessed value. Mathematically, this approach is comfortably set in the language of series, where each term in the series is a correction to the terms that came before. Technically, a series has an infinite number of small corrections, but in regular scientific practice a finite set of corrections will capture the accuracy possible today.

A classic example from atmospheric sciences illustrates. Jule Charney, while working on his Ph.D. on the origin of storms in the late 1940s, found a simple equation that he believed described the evolving waves. It could be related to a known equation, whose solutions were called the confluent hypergeometric functions. Great! All he needed to do was look up the values of these functions he wanted and plot the solutions to his theory. After some research in the library, Jule found that the confluent hypergeometric functions cannot be written in closed form! And, in the 1940s, computers were rare. There wasn't a table of sufficient accuracy for Charney's purpose–no one had ever calculated these functions at sufficient precision.

So Jule did what any mathematically-minded geoscientist[1] would do when the values of a special function were needed, Charney got a calculator, found a series that converged to the function, and started crunching away. In fact, even though his thesis results turned out to be really important

[1] Like you will be after reading this book!

(the equations of motion he derived made the first numerical weather predictions possible), many people looked up his thesis mostly to get at the table of hypergeometric function values he had tabulated in the appendix.

Let's suppose we wanted to determine the value of a particular hypergeometric function, $F[z]$ at the point $z = 1.0$.[2] Let's first assume that we can make accurate observations of the function (maybe even by observing the weather!), but not exactly at the point $z = 1.0$. We might make a series of figures like those in Fig. 1.1.

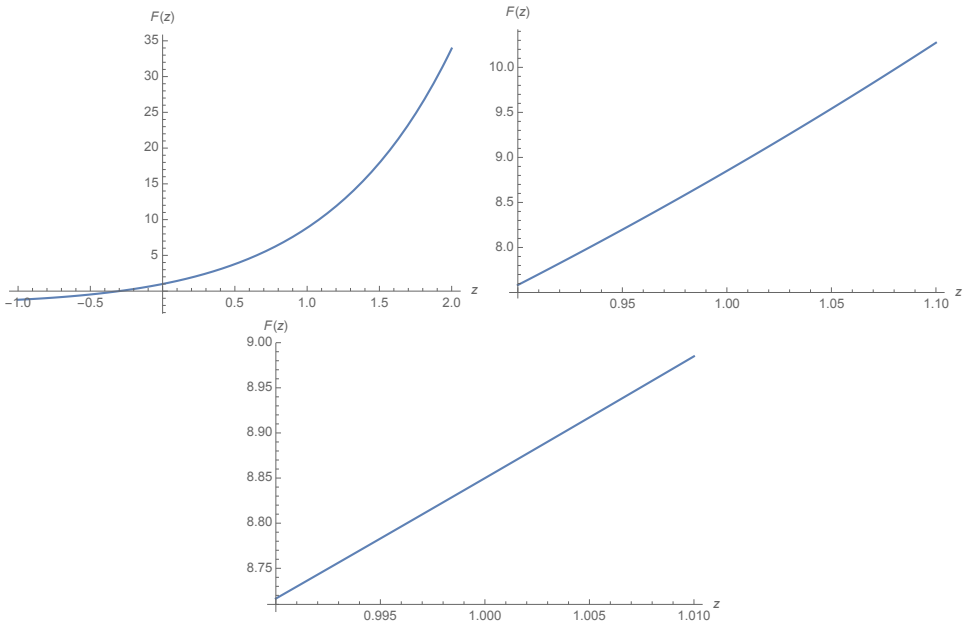


Figure 1.1: Zooming in on the hypergeometric function $F[z]$ near $z = 1.0$.

We can see that as we make the measurements more and more precise, we can better and better determine the values of the function: $10, 8.9, 8.85, \ldots$ This list is a *sequence* of numbers which converge to the true value, which these days a computer quickly reveals is very close to $8.849906238720033\ldots$ Here the trailing dots indicate that additional digits were not printed.

Definition 1.1 (Sequence) *A set of quantities ordered by association of each quantity to a natural number, which may be finite or infinite in length.*

Alternatively, perhaps we can't make such figures directly, but we can begin with an initial guess, and then improve it by adding or subtracting increasingly small bits at a time. The *Taylor series* method, described later in this chapter and closely related to Charney's method, is just such a procedure. The hypergeometric function is exactly 1 at z=0, so the Taylor series allows us to approximate the function in the neighborhood of z=1 by adding correction after correction to it's value nearby at z=0. Here are the first few terms in the series, rounded off to three digits:

$$1 + 4 + 2.54 + 0.969 + 0.266 + 0.057 + 0.010 = 8.847 \approx 8.849906238720033\ldots \tag{1.1}$$

Definition 1.2 (Series) *The sum of a sequence, which may be finite or infinite in length and may converge, diverge, or oscillate.*

[2]The confluent hypergeometic function actually has three inputs, but two of them have been chosen to have arbitrary values here.

When Charney went to tabulate his values, he had to use an approximation method–a sequence or series–for *each* datum he wanted to plot or put in his table. A research sequence often begins toward a scientific goal, then turns to mathematics, then returns to science with new mathematical tools, then returns to mathematics to develop new tools, and so on. Will your sequence oscillate, converge, or diverge?

1.1.1 Rigor, Difficulty, and Intuition

Precise definitions and statements, rigorous rules for manipulation and computation, and thorough proofs are the basis of mathematical rigor. The difficulty of a mathematics course or textbook or proof often stems from the rigor required. Mathematical intuition stems from having studied enough rigorous systems as to be able to rapidly appreciate what are the defining features of a new system when first exposed to it.

Geoscience also importantly works from definitions and statements, but they tend to be deliberately imprecise or inclusive to begin with–"working definitions"–until the world reveals itself through experiment to make additional precision valuable. The difficulty of a scientific course or textbook or experimental design usually stems from the complexity, subtlety, and unfamiliarity of the topic, the challenge of mastering measurement, field, experiment, and modeling techniques, and finding the dedication required to access remote locations or spend long hours in the lab acquiring this mastery. Scientific intuition stems from familiarity with how measurements and experiments tend to go wrong, a judicious sense of how many results can be squeezed out of an experiment accurately, and deep familiarity with the workings of a particular geophysical system.

Often, geoscientists find the rigor of mathematics unappealing and constrictive, while mathematicians find geosciences lacking in precision and lacking in the comforts of staying near home. Both disciplines are challenging and progress in both is is triumph of human achievement, but what makes each difficult is quite different. Mathematical geosciences–the topic of this book–strives to bring modest levels of rigor to geosciences and modest levels of geoscientific knowledge to mathematics, but the key idea is to build the kind of intuition that helps advance both fields. Thus, the level of rigor is not as high as in a mathematics course covering similar topics, nor is the level of geological information as broad or as deep as in a geosciences course.

1.2 Numbers as sequences and series

Sequences and series are so much a part of mathematics that even the numbers we use can be thought of–or derived or proven to exist–by considering sequences and series and the numbers they approach as the sequence gets longer and longer. Here are some examples, can you see some ways how they relate sequences and series?

Definition 1.3 (Natural Numbers) *The "counting numbers": 1, 2, 3, . . .*

Definition 1.4 (Integer Numbers) *The set of all natural numbers and their sums and differences.*

Definition 1.5 (Rational Numbers) *The set of all integer numbers and their ratios.*

Definition 1.6 (Real Numbers) *The set of all rational numbers, and all the numbers to which a converging*[3] *sequence of natural numbers can converge.*

Definition 1.7 (Decimal Representation) *A series approximation to the value of a number, based on the number of digits on most human hands. The leading digit gives the nearest value at that order, and then each successive digit is the remainder fraction in units of* $1/10$ *toward the next increment. Most rational and real numbers are represented by an infinite decimal series.*

Definition 1.8 (Binary Representation) *A series approximation to the value of a number, based on a series of ones and zeroes. The leading digit gives the nearest value at that order, and then each successive digit is the remainder fraction in units of* $1/2$ *toward the next increment. Most rational and real numbers are represented by an infinite binary series. In computers, each term in a binary series requires one bit of storage.*

Definition 1.9 (Floating Point Numbers) *The way that real and rational numbers are approximately represented in computers. Floats are rational numbers, but not all rational numbers can be represented as floats due to limited precision (i.e., the number of computer bits used to distinguish the number from similar numbers). Like scientific notation (e.g.,* $6.23 \cdot 10^{23}$*) there are computer memory bits used for the leading digits and bits used for the exponent and a bit used for positive and negative. It is important to remember that computer data types cannot directly represent most real numbers, just rational approximations to them.*

Bonus question: Which of the preceding sets of numbers has the most numbers in it?

1.3 Common Sense, Common Series: The Geometric Series

The geometric series is a useful standard series, because it is easy to find out what the sum of the terms is. A finite number (n) terms in a geometric series is

$$S_n = a + ar + ar^2 + ar^3 + \cdots + ar^n. \tag{1.2}$$

The series is called *geometric*, because each term is proportional to the preceding term. In this sense, *geometric* means "involving products or powers" much like the line, area, volume, etc., spanned by a unit length r involve successive powers of r.

Consider what happens to the sum as the number of terms increases. If $|r| \geq 1$, the terms get bigger with each increase in n. While the partial sum (or finite series) S_n for any n can be calculated, as n goes to infinity the sums don't head anywhere in particular. Each is large and the next is larger. In fact, for any big number N you choose, it is fairly easy to find an n such that $|S_n| > N$. For such a series, where each term gets successively larger, we can be sure that the series is divergent (i.e., it diverges away from the value it has for any finite number of terms).

If $r = -1$, then we have an oscillating series, since $S_n = a - a + a - a \cdots + a(-1)^n$. Each new term brings a change of magnitude $|a|$ in the sum, so the series never settles down! An oscillating series is also not convergent.

If, on the other hand, if each term gets smaller then we can imagine the case where the corrections become negligibly small, and we can identify a *limit* of the series. If $|r| < 1$, is this enough to say that the series converges? Well, in this case each term in the series is smaller than the one that

[3] Technically, converging in the Cauchy sense.

precedes it, or $|r^{n-1}| > |r||r^{n-1}| = |r^n|$. For example, with $r = 0.1$, then $r^2 = 0.01, r^3 = 0.001, \ldots$. For large n, the additional terms are very small compared to the preceding ones, so as the sequence continues the partial sum stops changing much with each new term, and we can consider the *limit as n goes to infinity* of the sequence of finite-length sums, which we take as a definition of the sum of the infinite series S.

$$S = \lim_{n\to\infty} S_n. \tag{1.3}$$

We will show that for all geometric series, if $|r| < 1$ then

$$S = \frac{a}{1-r}, \text{ where } S = \lim_{n\to\infty} S_n \text{ and } S_n \equiv \sum_{j=0}^{n} ar^n. \tag{1.4}$$

Demonstrating this result is part of the next example problem.

Example 1.1 (Prove the Formula) *Find a simple formula for a finite geometric series and an infinite geometric series when* $|r| < 1$.

We can compare a partial sum with r times itself,

$$S_n = a + ar + ar^2 + ar^3 + \cdots + ar^{n-1}, \tag{1.5}$$
$$rS_n = \quad ar + ar^2 + ar^3 + \cdots + ar^{n-1} + ar^n. \tag{1.6}$$

Subtracting the two equations, we find the value of any partial sum in a convenient form.

$$S_n - rS_n = a - ar^n,$$
$$S_n = \frac{a(1-r^n)}{1-r}. \tag{1.7}$$

Now, we can examine the limit of the partial sums as n gets large,

$$S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{a(1-r^n)}{1-r} = \frac{a}{1-r}\left[1 - \lim_{n\to\infty} r^n\right].$$

If $|r| < 1$, then the limit in the last term is zero, and (1.4) is proven.

$$S = \frac{a}{1-r}.$$

Example 1.2 (Purity) *In a water purification process, one-mth of the impurity is removed in the first stage. In each succeeding stage, the removal gets harder, so the amount of impurity removed is only one-mth of that removed in the preceding stage. If* $m = 2$*, how pure can the water be made? If* $m = 3$*?*

The purification removes a factor of $1/m$ more each stage. The impurity remaining after the n^{th} refining process, I_n, is therefore

$$I_n = 1 - 1/m - 1/m^2 - 1/m^3 \cdots - 1/m^n = 1 - (1/m + 1/m^2 + 1/m^3 + \ldots). \tag{1.8}$$

We use (1.7) to find a convenient form of the impurity at the n^{th} stage,

$$I_n = 1 - \frac{a(1-r^n)}{1-r} = 1 - \frac{\frac{1}{m}\left(1 - \left[\frac{1}{m}\right]^n\right)}{1-1/m}, \tag{1.9}$$

$$\lim_{n\to\infty} I_n = 1 - \frac{\frac{1}{m}}{1-1/m} = \frac{1-\frac{2}{m}}{1-\frac{1}{m}} = \frac{m-2}{m-1} \qquad \text{if} \qquad m > 1. \tag{1.10}$$

So for $m = 2$, $\lim_{n\to\infty} I_n = 0$, and for $m = 3$, $\lim_{n\to\infty} I_n = \frac{1}{2}$.

1.4 Convergence

Geometric series can be applied to some scientific problems (e.g., Example 1.2), but they are more commonly valuable as a benchmark for other more complicated series. The problem is that for many series, it is not clear whether they converge or not. By comparing them term-by-term to a carefully selected series, their rate of increase can be proven to be larger than that of a diverging series or smaller than that of a converging series.

Obviously, if the terms in an infinite series do not tend to zero then the series diverges. If they do tend to zero, does it imply convergence? Take, for example, the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots \tag{1.11}$$

It seems like this series converges, but does it? Each term is smaller than the preceding one, but is there a related series that doesn't converge?

Example 1.3 (Harmonic) *By comparison to a divergent series, show that the harmonic series does not converge.*

The harmonic series is bigger, term by term, than a geometric series with $r = \frac{1}{2}$.

$$\begin{aligned} 1 \quad &+ \frac{1}{2} + \left[\frac{1}{3} + \frac{1}{4}\right] + \left[\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right] + \ldots \\ > 1 &+ \frac{1}{2} + \left[\frac{1}{4} + \frac{1}{4}\right] + \left[\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right] + \ldots \\ = 1 &+ \frac{1}{2} + \left[\frac{1}{2}\right] \quad + \left[\frac{1}{2}\right] \quad + \ldots \end{aligned}$$

The last series is clearly divergent. The key to this comparison is noticing that

$$\frac{1}{2} > \frac{1}{4}, \frac{1}{3} > \frac{1}{4}, \frac{1}{5} > \frac{1}{8}, \frac{1}{6} > \frac{1}{8}, \frac{1}{7} > \frac{1}{8}, \tag{1.12}$$

$$\text{or in general } \frac{1}{2^n - k} \geq \frac{1}{2^n} \qquad \text{if } 2^{n-1} > k \geq 0 \tag{1.13}$$

This series (1.11) is called *harmonic*, because the sequence $1/n$ is related to the harmonies–i.e., the overtones–in musical instruments. A guitar string, for example, might vibrate with the whole

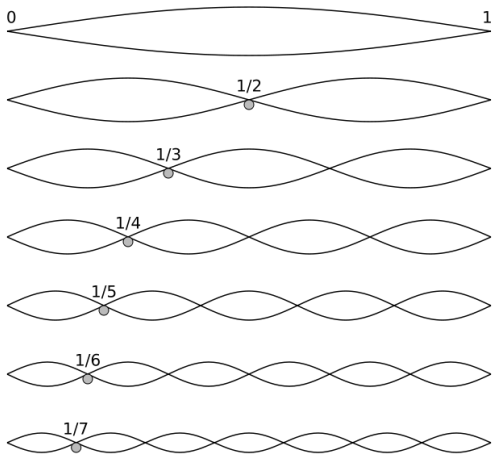


Figure 1.2: The harmonic sequence relates to vibrations of a string fixed on each end (source: Wikimedia commons).

string moving in the same direction (1/1), half of the string moving up and half moving down (1/2), thirds of the strings moving in opposite directions (1/3), and so on (Fig. 1.2). We will see that this sequence is important in describing many "wavy" phenomena, such as waves on the ocean, waves in the atmosphere, tides, and seismic waves in the Earth.

A digital readout is an interesting example. Definition 1.7 reveals that a decimal representation is closely related to a series. Obviously, any finite number of decimals is a meaningful number, but is it possible that the series might diverge as we take on more and more digits?

Example 1.4 (Decimal) *By comparison to a geometric series, show that the a decimal representation of a finite number must converge.*

First, divide whatever number we are considering by a convenient neighboring round number of the same sign, so we need only consider numbers N between 0 and 1. Its decimal representation up to the n^{th} digit is therefore a finite series like

$$N = 0 + \frac{a_1}{10} + a_2\left(\frac{1}{10}\right)^2 + a_3\left(\frac{1}{10}\right)^3 + \ldots,$$
$$N_n = 0 + \frac{a_1}{10} + a_2\left(\frac{1}{10}\right)^2 + a_3\left(\frac{1}{10}\right)^3 + \cdots + a_n\left(\frac{1}{10}\right)^n.$$

Each of the coefficients a_i is a number between 0 and 9. Subtracting, and comparing to these limiting values, we see that

$$N - N_n = a_{n+1}\left(\frac{1}{10}\right)^{n+1} + a_{n+2}\left(\frac{1}{10}\right)^{n+2} + \ldots,$$
$$0 < \left(\frac{1}{10}\right)^{n+1}\left[a_{n+1} + a_{n+2}\left(\frac{1}{10}\right)^1 + \ldots\right] \leq \left(\frac{1}{10}\right)^{n+1}\left[9 + 9\left(\frac{1}{10}\right)^1 + 9\left(\frac{1}{10}\right)^2 + \ldots\right],$$

because each term in the coefficient is limited by either 0 or 9. If the inequality is true for each

term, it is true for the sum. Using (1.4), we have

$$0 < N - N_n \leq \left(\frac{1}{10}\right)^{n+1} \frac{9}{1-\frac{1}{10}} = \left(\frac{1}{10}\right)^{n} \frac{9}{10-1} = \left(\frac{1}{10}\right)^{n}. \tag{1.14}$$

Thus, for any n, we have limited how far N_n can be away from N, and with enough digits, we can get N_n arbitrarily close to N.

1.5 Power

A *power series* is a series like that just examined in Example 1.4. It involves increasing powers of a variable, but unlike the geometric series the coefficients c_i may vary.

$$S = c_0 + c_1 r + c_2 r^2 + c_3 r^3 + \ldots \tag{1.15}$$

Many examples of power series are scientifically useful. As in Example 1.4, it is often useful to compare the power series to a related geometric series if there are doubts about convergence. Something must be known about the coefficients c_i to make the comparison meaningful.

1.6 Taylor

The most important scientific series from calculus is the Taylor series expansion,

$$f(x) = \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!} f^{(n)}(a). \tag{1.16}$$

This formula approximates any differentiable function as a power series (which may or may not converge). Estimating a few terms in the Taylor series is often the basis for building a model to fit data–and for estimating how much more would be known if one more term was measured. I used a Taylor series to make (1.1).

Let's see how the Taylor series can be generated, just by trying to fit a power series to a function $f(x)$ step by step. Consider the power series in powers of $(x-a)$,

$$S = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \ldots. \tag{1.17}$$

We can set $x = a$, which zeroes out every term but the first. Setting this equal to the value of the function $f(x)$ at $x = a$,

$$f(a) = c_0. \tag{1.18}$$

Now, let's take the derivative of every term in the series with respect to x and compare to the derivative of $f(x)$. Now, if we compare these two results when $x = a$,

$$\frac{dS}{dx} = S' = S^{(1)} = 0 + c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \ldots \tag{1.19}$$

$$\frac{df(x)}{dx} = f'(x) = f^{(1)}(x), f^{(1)}(a) = c_1. \tag{1.20}$$

Note that the number in parentheses in a superscript indicates the number of times the derivative has been taken (or the number of primes). Taking the derivative again, and again setting $x = a$, we find

$$f^{(2)}(a) = 2 \cdot 1 \cdot c_2, \tag{1.21}$$

$$f^{(3)}(a) = 3 \cdot 2 \cdot 1 \cdot c_3, \tag{1.22}$$

and so on. Solving each of these equations for c_i results in (1.16), where the *factorial* (!) notation has been used. For example, $n! = n \cdot (n-1) \cdot (n-2) \cdots \cdot 1$.

Here are some Taylor series expansions to common functions (actually, these are Maclaurin series, which are just Taylor series where $a = 0$).

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \tag{1.23}$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \tag{1.24}$$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \tag{1.25}$$

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \tag{1.26}$$

$$\tan^{-1}(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^{2n-1}}{2n-1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \dots \tag{1.27}$$

$$(1+x)^p = \sum_{n=0}^{\infty} \binom{p}{n} x^n = 1 + \frac{p!}{(p-1)!} x + \frac{p!}{2!(p-2)!} x^2 + \dots \tag{1.28}$$

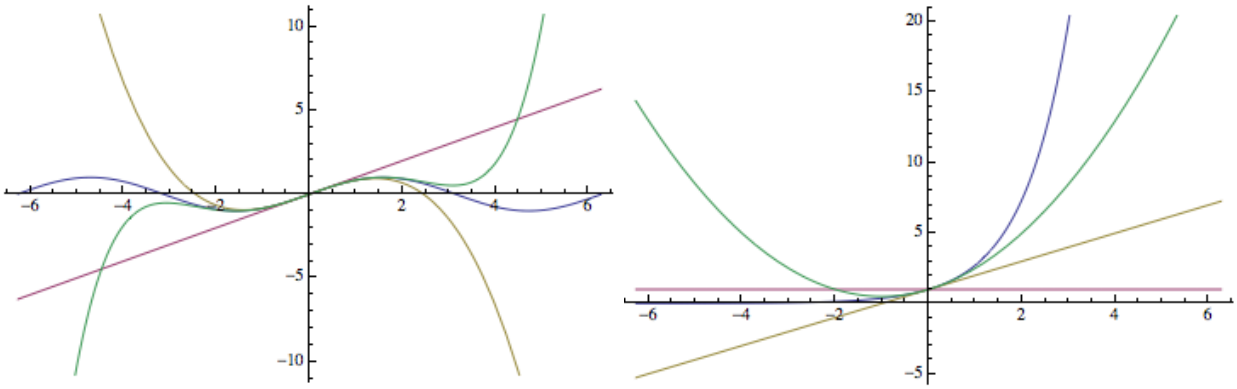
$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \dots \tag{1.29}$$

A shorthand notation for the *binomial coefficient* is used in (1.28), which is equivalent to the following factorial combination

$$\binom{p}{n} \equiv \frac{p!}{n!(p-n)!}. \tag{1.30}$$

By examining these formulae, the conventions must be $0! = 1$ and that all binomial coefficients are taken to be zero when $n > p$.

Example 1.5 (Taylor Realize) *Plot the* $\sin(x)$ *and* e^x *functions in* (1.23) *and* (7.1)*. Then overlay a) a plot of the first term, b) the sum of the first two terms, and c) the sum of the first three terms in the series. You may use MATLAB, mathematica, excel, or other plotting software if you like. Explain what is explained and what is missed by the truncated series approximations.*



The more terms that are kept, the more the fit resembles the function (blue line) near the y-axis. Furthermore, the higher the polynomial order, the more maxima and minima can be represented, Each wiggle in the function requires another polynomial order, as is clear in the sine approximation. The leading term is a constant, the first two terms is a straight line tangent, and the higher order approximations keep closer and closer to the function. The number of derivatives of the function matched increases with the order of the approximation.

Example 1.6 (Climate Sensitivity) *Climate sensitivity is the response of the climate system (as measured in terms of a change in global mean surface temperature ($\delta GMST$) or global surface air temperature ($\delta GSAT$)) to a change in atmospheric carbon dioxide. There are many ways to quantify sensitivity, but three are in common scientific use. Equilibrium climate sensitivity (ECS) is the change in temperature after a doubling of CO_2 as measured after decades or centuries, allowing the energy to disperse into the deep ocean. ECS is, by definition, the temperature after all of these changes have occurred, but it is often estimated before the eventual value, because the changes in temperature get smaller and smaller over time (i.e., a convergent sequence!). Transient climate response (TCR) is the amount of warming that occurs just after CO_2 doubles, in an experiment where it is increased by 1% each year, roughly mimicking the historical record. As TCR is measured instantaneously, rather than after equilibration, if the CO_2 was held fixed at the doubling level changes in temperature would continue afterward. Finally, earth system sensitivity (ESS), includes very long-term earth system feedbacks, such as changes in ice sheets or changes in the distribution of vegetative cover, and is often measured similarly to ECS except in models featuring more aspects of the earth system than just the atmosphere, oceans, and sea ice. The Intergovernmental Panel on Climate Change (IPCC) fifth assessment report (), assesses the ECS to likely range between 1.5C to 4.5C while TCR is assessed to be within 1C to 2.5C.*
We'll the Taylor Series to think a bit more deeply about ECS versus TCR. Let us imagine that there are a variety of time series from experimental simulations–of global mean temperature (T) and of atmospheric carbon dioxide concentration (C)–in experiments. The key difference between ECS and TCR is at what time *are the two quantities ECS and TCR estimated in these simulations (Gregory et al., 2004). So, let's think about this as choosing the a in the Taylor series.*
If we take a simulation when the CO_2 is increasing by 1% per year from 280ppm (a pre-industrial estimate) to 560ppm 100 years later (doubling), and set $a = t_{2\times}$ to be the time when this doubling occurs, then we can estimate the TCR by Taylor expanding the $T(t)$ record. The first term in the Taylor series $T(t_{2\times})$ minus the pre-industrial temperature T_0 is the TCR: $TCR = T(t_{t_{2\times}}) - T_0$. However, at that moment, the temperature is still adjusting, so we have to wait until later, when equilibration occurs, i.e., $T^{(1)}(t) = 0$. If we set a at that time, $a = t_{T^{(1)}=0}$, then the ECS is the first term in that Taylor series minus the pre-industrial temperature, $ECS = T(t_{T^{(1)}=0}) - T_0$.

1.7 Newton-Raphson Iteration

It might have occurred to you that in Exercise 1.5, some of the lower order approximations to the function might be very useful since they are simpler to work with than $\sin(x)$, for example. Suppose we wanted to use this simplicity to find an approximate solution to, e.g., $\sin(x) = 0.1$. Perhaps your calculator doesn't have the function for inverse of sine, so solving this equation would be quite hard. Perhaps you might try the repeated graphing and zoom approach from Fig. 1.1 to find an approximate guess. Or, maybe you think $\sin(x) \approx x$, based on the Taylor series (1.23), so maybe the solution is approximately $x \approx 0.1$?

Newton's method (or the Newton-Raphson method) is a method for improving such guesses to the solution of a equation that you can take the derivatives of (i.e., a smooth, continuous, differentiable function), which can be iterated (repeated) until a desired accuracy is reached (or no solution is found at all). Typically, the method results in a convergent sequence toward a solution of the equation. The steps are as follows:

1. Arrange the equation so that it reads: $f(x^*) = 0$. The * indicates a solution to the equation. There may be more than one...
2. Make a guess x_i of x^*.
3. Find an improved version by evaluating the Taylor series of the function at x_i, and choosing the next x_{i+1} where the Taylor series estimate is equal to zero,

$$f(x_{i+1}) \approx f(x_i) + f^{(1)}(x_i)\,(x_{i+1} - x_i) = 0, \tag{1.31}$$

$$\therefore \qquad x_{i+1} = x_i - \frac{f(x_i)}{f^{(1)}(x_i)}. \tag{1.32}$$

Step 3 can be repeated indefinitely. When you are near a solution, the steps will become increasingly small and you will converge. If the solution converged upon is not the desired one, or the iteration method diverges (which is rare), repeat the whole process from step 2 with a different initial guess.

Let's try with $\sin(x) = 0.1$. Step 1 is $f(x^*) = \sin(x^*) - 0.1 = 0$. For step 2, we'll use $x = 0.1$, as mentioned above. For step 3, we find $x_2 = 0.1 - \frac{\sin(0.1)-0.1}{\cos(0.1)} = 0.1 - \frac{0.0998334\ldots\ -0.1}{0.995004\ldots} = 0.100167\ldots$. We could repeat the iteration again, but $x = 0.1$ was only wrong by 0.1%, and $x = 0.100167$ is wrong by only 0.0004%.

Let's work through a geophysical example.

Example 1.7 (Wave Solution) *The equation that relates the speed of a wave on the ocean's surface to it's wavelength (L) and the depth of the ocean (H) is*

$$c = \sqrt{\frac{gL}{2\pi}} \tanh\left[\frac{2\pi H}{L}\right]. \tag{1.33}$$

This formula is tricky to solve for L, being a product of a square root of L times a hyperbolic tangent! What if we want to know the wavelength of a wave whose speed (c) is $10\,\mathrm{m\,s^{-1}}$ in $100\,\mathrm{m}$ deep water?

Approximating the acceleration due to gravity g as $10\,\mathrm{m\,s^{-2}}$, we have

$$10 = \sqrt{\frac{10L}{2\pi}}\tanh\left[\frac{2\pi 100}{L}\right], \tag{1.34}$$

$$\tanh\left[\frac{2\pi 100}{L}\right] - \sqrt{\frac{20\pi}{L}} = 0. \tag{1.35}$$

The last equation can't be solved in closed form, since there is an L both inside and outside of the hyperbolic tangent. We could solve graphically, by looking for where the function on the left of 1.35 crosses equals zero (Fig. 1.3). But, we couldn't easily draw this function. We can take the

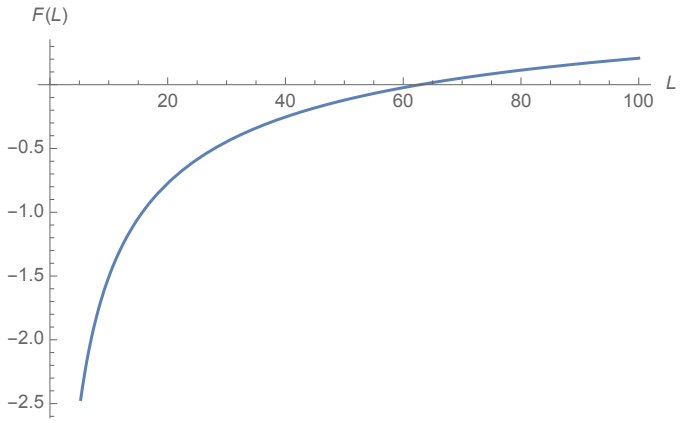


Figure 1.3: The value of the left hand side of (1.35) versus L.

derivative of it, though.

$$\frac{d}{dL}\left(\tanh\left[\frac{2\pi 100}{L}\right] - \sqrt{\frac{20\pi}{L}}\right) = L^{-3/2}\sqrt{5\pi} - \frac{200\pi\,\mathrm{sech}^2[(200\pi)/L]}{L^2} \tag{1.36}$$

Since we can do that, we have all of the ingredients to use Newton's method (1.32)!

$$L_{i+1} = L_i - \frac{\tanh\left[\frac{2\pi 100}{L_i}\right] - \sqrt{\frac{20\pi}{L_i}}}{L_i^{-3/2}\sqrt{5\pi} - \frac{200\pi\,\mathrm{sech}^2[(200\pi)/L_i]}{L_i^2}} \tag{1.37}$$

This is a bit scary to calculate, but using Matlab or Mathematica it is easily programmed. The sequence of results, if we guess L_1 =50 m to begin with is:

$$L_i = \underbrace{50\,\mathrm{m}}_{i=1}, \underbrace{60.8\,\mathrm{m}}_{i=2}, \underbrace{62.8\,\mathrm{m}}_{i=3}, \underbrace{62.83\,\mathrm{m}}_{i=4}, \underbrace{62.832\,\mathrm{m}}_{i=5}, \ldots \tag{1.38}$$

So, you can see that this series for L_i quickly converges with iteration number i. The corresponding values of the left hand side of (1.35) are

$$\tanh\left[\frac{2\pi 100}{L_i}\right] - \sqrt{\frac{20\pi}{L_i}} = \underbrace{-0.12}_{i=1}, \underbrace{-0.016}_{i=2}, \underbrace{-0.00025}_{i=3}, \underbrace{-0.000015}_{i=4}, \underbrace{1.1\cdot 10^{-6}}_{i=5}, \ldots \tag{1.39}$$

Thus, they quickly approach zero as intended.

1.8 Introducing Fourier Series

Another common series is the Fourier series. Like the polynomial series that results from Taylor expansions, the Fourier series is a sum of terms that are higher powered in each term in the series when written in (complex) exponentials,

$$f(x) = c_0 + c_1 e^{ix} + c_2 e^{2ix} + c_3 e^{3ix} + \dots \tag{1.40}$$

The c coefficients are constants that can be chosen to select one particular Fourier series from the general form written here. In the next chapter, we will see that these complex exponentials are exactly equivalent to sines and cosines, which makes the Fourier series equivalent to

$$f(x) = \frac{1}{2}a_0 + a_1 \cos x + a_2 \cos 2x + a_3 \cos 3x + \dots + b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \dots \tag{1.41}$$

At this stage, it is only important to appreciate that a Fourier series is another example of a series, and that it is constructed from a set of oscillating functions. Some of you who are used to thinking about the coefficients of the argument of sines and cosines as frequencies may note that the frequencies here obey the harmonic progression (like Fig. 1.2)! After a few more chapters, we will be in a position to see how a Fourier series can be fit to any periodic function, which is a powerful idea.

1.9 More Wiggly with More Derivatives

Let's examine the Taylor and Fourier series to think about whether integrals or derivatives of a function tend to be smoother or wigglier. This will be an important concept in handling real data.

First, let's examine the degree of wiggles in the Fourier series. Each subsequent term oscillates more quickly: $\frac{1}{2}a_0$ does not oscillate, $a_1 \cos x$ oscillates, and $a_2 \cos 2x$ oscillates with twice the frequency of $a_1 \cos x$. However, it's not just the frequency of oscillation that contributes to wiggliness in the overall series, it's also the amplitude. We don't know much about the initial amplitudes, as that depends on the specific function being approximated with the Fourier series, but we *do know* about the derivatives and integrals of the Fourier series in comparison to the original function. Since we can distribute the derivatives over the sums constituting each series, then

$$f(x) = \frac{1}{2}a_0 + a_1 \cos x + a_2 \cos 2x + a_3 \cos 3x + \dots + b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \dots \tag{1.42}$$

$$\frac{df(x)}{dx} = 0 - a_1 \sin x - 2a_2 \sin 2x - 3a_3 \sin 3x + \dots + b_1 \cos x + 2b_2 \cos 2x + 3b_3 \cos 3x + \dots \tag{1.43}$$

$$\int f(x)\mathrm{d}x = c_0 + \frac{1}{2}a_0 x + a_1 \sin x + \frac{a_2}{2}\sin 2x + \frac{a_3}{3}\sin 3x + \dots - \frac{b_1}{1}\cos x - \frac{b_2}{2}\cos 2x - \frac{b_3}{3}\cos 3x + \dots \tag{1.44}$$

While we can't say how quickly the a_n and b_n coefficients decrease in the original function, note that taking the derivative tends to increase the coefficients more based on the frequency of the sine or cosine, while taking the integral tends to decrease each terms contribution. Thus, the derivative

is expected to be more wiggly, while the integral is less so. You might be concerned about the sines being exchanged for cosines, and while thinking that through is a good exercise, it is easy enough to take another derivative and another integral to illustrate.

$$f(x) = \frac{1}{2}a_0 + a_1\cos x + a_2\cos 2x + a_3\cos 3x + \cdots + b_1\sin x + b_2\sin 2x + b_3\sin 3x + \ldots \tag{1.45}$$

$$\frac{d^2f(x)}{dx^2} = 0 - a_1\cos x - 4a_2\cos 2x - 9a_3\cos 3x - \cdots - b_1\sin x - 4b_2\sin 2x - 9b_3\sin 3x - \ldots \tag{1.46}$$

$$\int\int f(x)\mathrm{d}x\mathrm{d}x = c_1 + c_0x + a_0x^2 - a_1\cos x - \frac{a_2}{4}\cos 2x - \frac{a_3}{9}\cos 3x + \cdots - \frac{b_1}{1}\sin x - \frac{b_2}{4}\sin 2x - \frac{b_3}{9}\sin 3x + \ldots \tag{1.47}$$

Now it is clear that the second derivative function will be very similar to the original function, but with amplification of the higher frequency terms, leading to a more wiggly function. Inversely, the integral will be smoother.

How does this look in the Taylor series? Well, each term in the Taylor series is likewise easy to integrate and differentiate. Let's use the particularly easy case of $a = 0$.

$$f(x;0) = f(0) + xf^{(1)}(0) + \frac{x^2}{2!}f^{(2)}(0) + \frac{x^3}{3!}f^{(3)}(0) + \ldots, \tag{1.48}$$

$$\frac{df(x)}{dx} = 0 + f^{(1)}(0) + xf^{(2)}(0) + \frac{x^2}{2!}f^{(3)}(0) + \ldots, \tag{1.49}$$

$$\int f(x)\mathrm{d}x = c_0 + xf(0) + \frac{x^2}{2}f^{(1)}(0) + \frac{x^3}{3!}f^{(2)}(0) + \ldots \tag{1.50}$$

So, since the higher powers are wigglier, the differentiation tends to reduce the power

What if the original function has a kink in it, where it is non-differentiable? Well, then any Taylor or Fourier finite series approximations to it will not have the kink, but will have smooth transitions over that part of the function. Indeed, since the Taylor and Fourier methods actually require integration or differentiation, such a kink may make it difficult or impossible to calculate the Fourier or Taylor series representations unless a smoothed approximation is first formed.

1.10 Small Corrections

1.10.1 Asymptotics

The convergence and accuracy of finite truncations of the geometric series and Taylor series depends critically on the terms growing small with n. Most of the time, this is because some part piece of the makeup of each term is already small ($(x-a)^n$) or big ($n!$), and increasing n makes it smaller still. Having a small or big parameter is called having an *asymptotic parameter*, because it can be useful in finding a limit to which the series asymptotes as n goes to infinity. Asymptotics is the mathematical formulation of considering such limiting behavior in a wide range of problems with small or large parameters. Chapter 14 presents some of the most common techniques.

Science is always in the business of eliminating parameters with negligible contributions to the theory. Asymptotic parameters take this one step farther. Suppose we are doing a Taylor series

at a point x nearby a, so $(x-a)$ is our small parameter. It may not be negligible on its own, but $(x-a)^2$ will be smaller so it may be negligible, or $(x-a)^3$ which is smaller still.

1.10.2 Accuracy

Sometimes, we want to find the limit of the infinite sum S, and try to do so by summing only a finite number of terms S_n. This approximation is called a finite truncation of the series. This may be because we don't know how to calculate S directly, or because (e.g., due to limited observations or the cost of evaluating lots of terms) we only can evaluate a limited number of terms. We know that ultimately the *residual* or *truncation error* R_n between the two will become small, since $R_n = S - S_n$ and $\lim_{n\to\infty} R_n = 0$, but without further study we don't know how quickly.

Even when a series converges, it may do so excruciatingly slowly. For example, we know that $1 = \cos(0) = \sin(\pi/2)$, since sine and cosine are nearly the same function except sine begins at $x = 0$ with the climbing part of the function whereas cosine starts at a maximum. However, evaluating the first few terms of the Taylor series expanding $\sin(x)$ near $x = 0$ in (1.23) at $x = \pi/2$, it takes up to the term derived from the *fifth* derivative of sine before the answer is within 1% of 1. It takes up to the *thirteenth* derivative term to get the series to approximate the answer at $x = 3\pi/2$ with similar accuracy. Taking derivatives of sine is easy (if you remember the formula), but taking the derivative of some functions is costly, so a more rapidly converging series would be valuable. Considering truncation and building better algorithms and series that converge faster are important parts of computation (Chapter 15).

1.10.3 Small Compared to What? Introducing Units

Applying the preceding techniques to the real world is missing one major ingredient: units. For example, if $(x-a)$ is 0.1 meters, is it small? 0.1 is small, but 10 is big and that's the same distance when measured in centimeters. Which interpretation is right? Is it small in meters or big in centimeters? Also, how do we compare $(x-a)$ with units of meters to $(x-a)^2$ with units of meters squared (area?). The answer lies in noting that we can really only compare meters to meters, or apples to apples. Thus, smallness needs to be measured in a way that is independent of arbitrary choices about units, and in a way that allows us to take small parameters to arbitrary powers and still compare meaningfully. This idea forms the basis of *dimensional analysis* (Chapter 4), which is a crucial part of mathematical science which we will revisit again and again.

For now, the critical precursor to doing a good job with units and dimensional analysis is being careful to give units their due in problem sets. Don't write bare numbers when a unit should be attached!

1.11 Homework Problems

1.11.1 Manipulation

Exercise 1.1

$$1 + 0.1 + 0.01 + 0.001 + \cdots = \tag{1.51}$$

Exercise 1.2

$$1 + 1/2 + 1/4 + 1/8 + \cdots = \tag{1.52}$$

Exercise 1.3

$$1 + 1 + 1 + 1 + \cdots = \tag{1.53}$$

Exercise 1.4 (Taylor Realize) *Confirm the first two nonzero terms of each of the series in (1.24), (1.26), and (1.28) by plugging into those formulae for $n = 0, 1, 2, \ldots$. For some n, the terms will vanish, so keep going until you get two nonzero ones. Then, use (1.16) and derive the first two nonzero terms in the series by taking the derivatives of each function. Again, some terms will vanish (and they may not match up one-to-one with the formula in (1.23-1.28) until you have enough terms calculated).*

1.11.2 Application

Exercise 1.5 (30N) *An important parameter in the consideration of the physics of the rotating Earth is the Coriolis parameter: $2\Omega\sin(\varphi)$, where Ω is the angular rate of rotation of the earth in radians (2π in a day, or $2\pi/(24hr/s \cdot 3600s)$) and φ is the latitude. Taylor expand the first terms in this parameter around 30 degrees North, or $30^\circ \cdot (2\pi radians/360^\circ) = \pi/6$ radians. Carry out the expansion to the term including $(\varphi - \pi/6)^3$.*

Exercise 1.6 (Rocking with Taylor Swiftly) *Consider pushing on the rock depicted in Fig. 3.2. If you push or pull gently, the rock will push back. If you push or pull hard, not so much. We will use this example to consider how nonlinear functions are sensitive to amplitude in a way that linear functions are not.*
a, Equal and Opposite: *Make a graph depicting the force applied, F_A (positive=push in horizontal direction, negative=pull in horizontal direction) versus the force back from the rock F_R (after equilibration). You can suppose that all forces go in the same direction. (Hint: prevent acceleration up to a point, and $F_A + F_R = ma$.*
b, Constant Approximation: *If the function $F_R(F_A)$ is fit with a Taylor series around $F_A = 0$ and truncated at the first (constant) term ($F_R(F_A) \approx c_0$). Describe this system's response to applied forces.*
c, Linear Approximation: *If the Taylor series for $F_R(F_A)$ around $F_A = 0$ is truncated after two terms ($F_R \approx c_0 + c_1 F_A$), it predicts how much F_R for $F_A = 1$ Newton? 2 Newtons? 10^5 Newtons? Can fracture happen?*
d, Breaking Point: *What is the minimum number of terms in the Taylor series that must be*

Figure 1.4: **A situation that's sensitive to perturbation amplitude.**

retained to make F_R do something other than double when F_A is doubled?
e, Extremes: *Consider the response force F_R at positive and negative F_A. Is it approximately an even or odd function? If this were exactly true (i.e., symmetry or antisymmetry between pushing and pulling) what is the first term in the Taylor series that could make F_R sensitive to amplitude of F_A?*

Exercise 1.7 (Taylor rhymes with Baylor) *Suppose the function $h(x)$ plotted in the figure is found by measuring topography along the x direction, and in particular consider fitting it with a series expansion near the point marked A. Sea level is $h = 0$, so we are particularly interested in $h(x) = 0$, which indicates the location of coastlines. The function is not known, but we consider the possibility of approximating a Taylor series expansion to it.*

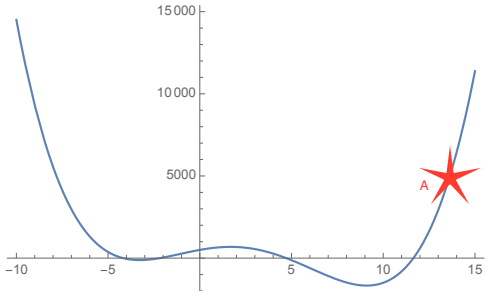


Figure 1.5: A function to be fit by Taylor expansion near point A at the star.

a, Counting: *How many coastlines are there, that is how many solutions to $h(x) = 0$ are there?*
b, Constant: *If the function is fit with a Taylor series and truncated at the first (constant) term, and if the truncated approximation is denoted $\tilde{h}_0(x)$, then how many solutions are there to $\tilde{h}_0(x) = 0$?*
c, Variations: *If the Taylor series is instead truncated after two terms ($\tilde{h}_1(x)$) then how many solutions are there to $\tilde{h}_1(x) = 0$, and what is the approximate value of the solution x?*
d, How many?: *What is the minimum number of terms in the Taylor series that must be retained to approximate all of the coastlines in the real function $h(x)$? Why?*
e, Extremes: *Consider the function at large magnitudes of positive and negative x. If the truncated Taylor series matches this behavior at large $|x|$, predict the sign of the coefficient in the largest power of x and whether the power is even or odd.*

1.11.3 Scheming Schematics and Articulate Analysis

Exercise 1.8 (Who is the Oddest?) *Sometimes it is said that* sin *is an* odd function *and* cos *is an* even function*. Examine the* sin *and* cos *functions in (1.23) and explain what this means in terms of the exponents of x. Consider $\sin(x)$ versus $\sin(-x)$ and $\cos(x)$ versus $\cos(-x)$, how do they compare? How do odd and even functions compare under the sign reversal of their argument (i.e., the input to the function, x or $-x$) in general?*

Exercise 1.9 (Biggie Smalls) *Examine each factor in the product that makes up (1.16). What makes them large or small as n increases?*

Exercise 1.10 (Getting Seri-us) *The Taylor series and the Fourier series are methods to determine the coefficients in series expansions of a function $f(x)$. They apply in approximating the function near $x = a$ and over a given interval where the function is periodic. Suppose we have a function that is periodic when $-\pi \leq x - a \leq \pi$, then we can write the Taylor and Fourier series–for the* same function *$f(x)$ as,*

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \dots,$$
$$f(x) = d_0 + d_1 e^{i(x-a)} + d_2 e^{2i(x-a)} + d_3 e^{3i(x-a)} + \dots$$

Use these forms to answer the following questions. Hint:

$$e^{ni(x-a)} = 1 + in(x-a) - \frac{n^2}{2}(x-a)^2 - \frac{in^3}{6}(x-a)^3 + \cdots + \frac{(in)^m}{m!}(x-a)^m + \dots$$

a, Constant: *Is there a simple relationship between c_0 and d_0 very near $x = a$ for all $f(x)$?*
b, Variations: *Does a specific term in the Fourier series, say $d_2 e^{2i(x-a)}$ always correspond to a specific term in the Taylor series, regardless of what $f(x)$ is?*
c, Equivalence?: *Assuming both series converge, is there a unique (though perhaps complicated) relationship among all of the c_n and all of the d_n?*
d, Specifically: *If $f(x) = e^{i(x-a)}$, what are the values of all nonzero d_n and c_n?*

1.11.4 Jargon to Argot

Exercise 1.11 (Sequence vs. Series) *What's the difference between a sequence and a series?*

Exercise 1.12 (Asymptote, Asymptotic, Limit) *An asymptote is a line that continually approaches a given curve but does not meet it at any finite distance. Contrast this definition to that of the adjective* asymptotic *and the term* limit*.*

Chapter 2

Real and Imaginary are Complex

Reading: Boas (2006, 2.1-2.2)

Somewhere during middle school, you probably were first amused by the phrase "imaginary numbers." If your experience was like mine (or my kids') it was years until they were formally introduced. Plus, you were never told why this odd phrase was used to describe them. Worse yet, they were just used to quiz your memorization of algebraic manipulations using the one key relation $i = \sqrt{-1}$. The deeper mysteries like why you can get away with only $\sqrt{-1}$ and didn't also need $\sqrt[3]{-1}$ (which would very fairly be called complex!) are lost.

But, in geophysics, imaginary numbers are incredibly useful in a variety of ways. These notes will introduce three key ways, which feed into many applications.

2.1 Maps, Distances, and Polynomial Roots

2.1.1 The Complex Plane

Reading: Boas (2006, 2.3-2.5)

We will spend much time this semester mapping lines and shapes and places and forces into locations in space. The complex plane is a very simple way of laying out two-dimensional maps, and such maps are also useful for manipulating complex numbers! For any complex number $z = x + iy$, it can be mapped onto a plane with x as the horizontal coordinate and y as the vertical coordinate (Fig. 2.1a).

We can express the same number in polar coordinates, by exploiting the triangle with hypotenuse r and angle counterclockwise from the x axis θ, or

$$z = x + iy = r(\cos\theta + i\sin\theta) = re^{i\theta}. \tag{2.1}$$

The last form uses the Euler relation ($e^{i\theta} = \cos(\theta) + i\sin(\theta)$), which is true for any θ. This relationship can be taken for granted now, but it will be derived below using Taylor series. Look at how this formula works out in Fig. 2.1b. Notice how increasing the power of the exponent in the formula results in a rotation about the coordinate axis.

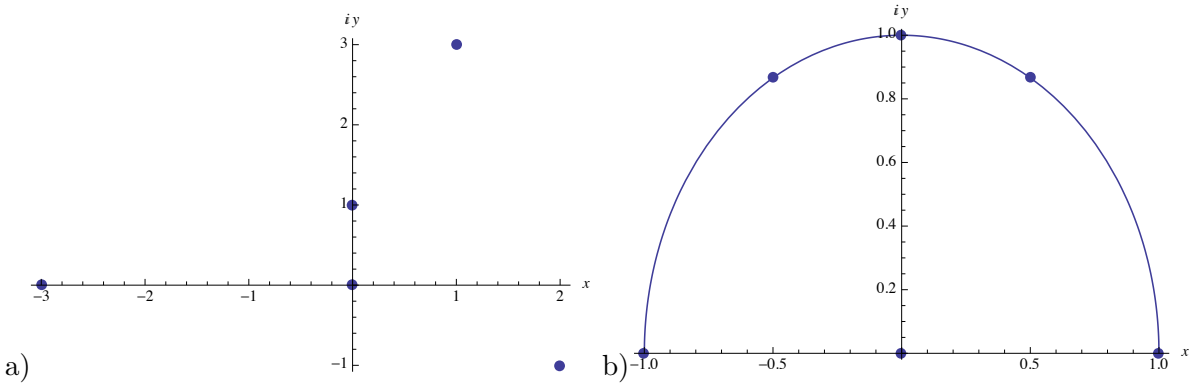


Figure 2.1: a) Points plotted on the complex plane: $0+0i, 1+3i, -3, 1i, 2-i$. b) Points plotted on the complex plane, based on polar representation: $0, 1, -1 = 1e^{i\pi}, i = 1e^{i\pi/2}, e^{i\pi/3}, e^{3i\pi/2}$.

Once we are in this framework of expressing points in the complex plane with single complex numbers, we can do many manipulations of these number to map and remap points on the plane to other points. If the remapping goes from the complex plane (or a subset) to the complex plane (or a subset) and is differentiable, then we say it is *analytic*. Complex analysis is the study of such operations, and it is a very powerful tool in 2-dimensional mapping. We will soon move to other systems (vectors, tensors) that are capable of handling higher dimensionality, but in 2d mapping the complex analysis is pretty tough to beat!

2.1.2 The Real Numbers

Reading: Boas (2006, 2.1)

Thus, there are really two real numbers (x and y) hiding inside of one complex number: z. In computer memory, complex numbers typically require twice as much storage space as real numbers (floating point number or floats).

So, let's return to why we don't need $\sqrt[3]{-1}$ in addition to $\sqrt{-1}$. In complex multiplication, the absolute value of the product of two numbers is the product of the absolute value, while the angle of each number away from the x-axis is summed together.

$$z_1 z_2 = r_1 e^{i(\theta_1)} r_2 e^{i(\theta_2)} = r_1 r_2 e^{i(\theta_1+\theta_2)} = |z_1||z_2| e^{i(\theta_1+\theta_2)}. \tag{2.2}$$

So, since $|-1| = 1$, the square root of -1 is just a unit distance away from the origin as 1, but halfway in angle from 1 to -1. That is i. Finding $\sqrt[3]{-1}$ is just the same, it is a unit distance away from the origin and a third of the way to -1. By this method you can figure out lots of things (the square root of positive real numbers are always real, for example).

Similarly we can do algebraic work to show that the magnitude of a complex number z is $\sqrt{zz*}$, where $z*$ is the complex conjugate of the number z, or $x - iy$, if $z = x + iy$.

2.1.3 How Did We Get So Complex? Root Finding

Reading: Boas (2006, 2.1)

More generally, we can wonder where all these roots come from and why we need them. This takes us back to the polynomial series we considered in the last chapter. Suppose we arrive at a physical process where y results from x at a given level of approximation (truncation) as,

$$y = az^2 + bz + c. \tag{2.3}$$

For a while we are happy with this relation, providing real z values and finding real y outputs. But, eventually we begin to wonder how we might predict what z will give us a particular y–that is, we want to invert the quadratic relation. Long ago, we were taught that this can be solved as

$$z = \frac{-b \pm \sqrt{b^2 - 4a(c-y)}}{2a}. \tag{2.4}$$

This gives two solutions (the + and the −) for every y where $b^2/4a - c > -y$. But what about if y becomes negative? What then? Then we need complex numbers so we can say for that case,

$$z = \frac{-b \pm i\sqrt{|b^2 - 4a(c-y)|}}{2a}. \tag{2.5}$$

We could have used a more complicated initial series, going up to z^2 (quadratic), z^3 (cubic), or z^4 (quartic). All of those polynomials can be solved without approximation with two, three, or four solutions. However, there are *less than or equal* to that many solutions that are real. That is, what complex numbers give us is the ability to keep track of how many solutions a polynomial will have, even when they are not real (and so may not be measurable, hence the name imaginary).

This situation is no problem when we know the formulae to find the roots (quintic or above). Then we must use numerical algorithms to find the roots. Typically, these methods are sophisticated repeated cycles (recursions) of guess, check, and correct the guess. They converge well when you guess an initial condition near a root, but erratically if not. It is very helpful in following such a method to know how many we need to find! This knowledge is only available when complex solutions are allowed.

2.2 Euler Rhymes with Boiler (not Ruler)

Reading: Boas (2006, 2.9, 2.11)

The Euler formula is sometimes taken to be one of the most astonishing results in mathematics, it is:

$$e^{i\theta} = \cos\theta + i\sin\theta. \tag{2.6}$$

It can be easily derived from examining the Taylor series of $\sin x, \cos x$, and e^x, given in (1.23-3.1).

It produces an even more amazing formula, which relates almost every high school mathematical number to one another:

$$e^{i\pi} + 1 = 0. \tag{2.7}$$

But, this is only one example of the wonders of complex numbers and their relation to maps, circles, and triangles.

The Euler formula for $e^{i\theta}$ and $e^{-i\theta}$ can be solved for a formula for sines and cosines.

$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}, \qquad \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}. \tag{2.8}$$

Note how the odd (sine) function comes out as odd under reversal of sign of θ and the even (cosine) function does not.

2.3 Sines, Sines, Everywhere Sines

Reading: Boas (2006, 2.11) The Euler formula is a very powerful one for compressing wavelike features, which we often take to be a sine wave. If the wave has amplitude a, wavelength λ, initial phase θ, and phase speed c in the x direction, we write for some wave variable ϕ

$$\phi = a\cos(2\pi[x - ct]/\lambda + \theta). \tag{2.9}$$

As time t increases, the wave moves more toward larger x. It is pretty hard from this formula to figure out how to measure θ if you happened to miss the moment at $x = 0, t = 0$. Consider the exponential form:

$$\phi = \mathrm{Re}\Big(ae^{i\theta}e^{2\pi i[x-ct]/\lambda}\Big) = \mathrm{Re}\Big(ze^{2\pi i[x-ct]/\lambda}\Big). \tag{2.10}$$

Now, both the amplitude and the initial phase are captured in the complex parameter z, and all of the manipulation of polar coordinate forms of complex numbers can be applied. Note that I use Re for real part and Im for imaginary part.

2.4 Sinh It Up Tight, Cosh It's a Rough Ride.

Reading: Boas (2006, 2.12)

How do we tell sine from cosine? Well

$$\sin(0) = 0, \qquad \sin'(0) = \cos(0) = 1 \tag{2.11}$$
$$\cos(0) = 1, \qquad \cos'(0) = -\sin(0) = 0. \tag{2.12}$$

As we will soon see, these simple boundary values make differential equations easy to solve because the solution at 0 is simple. Similarly, the hyperbolic sine and cosine obey,

$$\sinh\theta = \frac{e^{\theta} - e^{-\theta}}{2}, \qquad \cosh\theta = \frac{e^{\theta} + e^{-\theta}}{2}. \tag{2.13}$$

Similarly to the sines and cosines, these obey simple conditions at the origin.

$$\sinh(0) = 0, \qquad \sinh'(0) = \cosh(0) = 1 \tag{2.14}$$
$$\cosh(0) = 1, \qquad \cosh'(0) = \sinh(0) = 0. \tag{2.15}$$

Thus, we can relate the hyperbolic function of an imaginary number to the corresponding trigonometric function. Similarly, we can extend the definitions of logarithm.

2.5 Example Problems

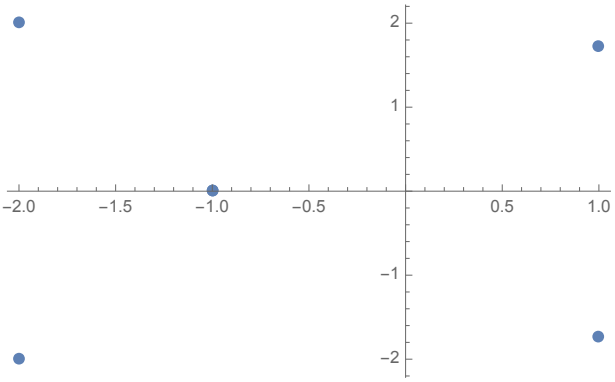
2.5.1 Jargon to Argot

Example 2.1 (Amplitude and Phase) *Look up the definitions of the amplitude and phase of a sinusoidal wave. Write equivalent real-valued waves, at a given time t and for all x, using only sine, cosine, and the real or imaginary part of a complex exponential. Label the phase and amplitude explicitly.*

Any sinusoidal wave can be written as $A\sin(kx - \omega t + \delta_s)$. The amplitude is A. According to Wikipedia, "Phase in sinusoidal functions or in waves has two different, but closely related, meanings. One is the initial angle of a sinusoidal function at its origin and is sometimes called phase offset or phase difference. Another usage is the fraction of the wave cycle which has elapsed relative to the origin." By the latter definition, we take the phase at the origin ($t = 0$, $x = 0$) to be δ_c. Alternatively, the same wave can be written as $A\cos(kx-\omega t+\delta_c)$ where $\delta_c = \delta_s - \pi/2$. These equivalences are made more clear by considering the complex exponential form of the waves (assuming amplitude A is still a real measured quantity), $A\mathrm{Re}\left[e^{i(kx-\omega t+\delta_c)}\right] = \mathrm{Re}\left[Ae^{i\delta_c}\right]\mathrm{Re}\left[e^{i(kx-\omega t)}\right]$. In this form, it is clear that any phase or amplitude can be selected by choosing the real and imaginary parts of the complex coefficient $Ae^{i\delta_c}$, and the sin form can be found by choosing it to be pure imaginary and the cos form can be found by choosing it to be pure real.

2.5.2 Manipulation

Example 2.2 (Plots on a Plane!) *Problems 2.4.3, 2.4.7, 2.4.9 of Boas (2006). You can plot them all on the same figure. Don't forget the complex conjugates!*



$2.4.3 : 1 - i\sqrt{3}; (x, y) = (1, -\sqrt{3}); (r, \theta) = (2, -\pi/3); 2\cos(\pi/3) - 2i\sin(\pi/3); 2e^{-i(\pi/3)}$

$2.4.3^* : 1 + i\sqrt{3}; (x, y) = (1, \sqrt{3}); (r, \theta) = (2, \pi/3); 2\cos(\pi/3) + 2i\sin(\pi/3); 2e^{i(\pi/3)}$

$2.4.7 : -1; (x, y) = (-1, 0); (r, \theta) = (1, \pi); \cos(\pi) + i\sin(\pi); e^{i(\pi)}$

$2.4.7^* : -1; (x, y) = (-1, 0); (r, \theta) = (1, \pi); \cos(\pi) + i\sin(\pi); e^{i(\pi)}$

$2.4.9 : 2i - 2; (x, y) = (-2, 2); (r, \theta) = (2\sqrt{2}, 3\pi/4); 2\sqrt{2}\cos(3\pi/4) + 2i\sqrt{2}\sin(3\pi/4); 2\sqrt{2}e^{i(3\pi/4)}$

$2.4.9^* : -2i - 2; (x, y) = (-2, -2); (r, \theta) = (2\sqrt{2}, -3\pi/4); 2\sqrt{2}\cos(3\pi/4) - 2i\sqrt{2}\sin(3\pi/4); 2\sqrt{2}e^{-i(3\pi/4)}$

Example 2.3 (Crunch your abs!) *Problems 2.5.5, 2.5.6, 2.5.26, 2.5.28, 2.5.29 of Boas (2006).*

$$2.5.5 : (i+\sqrt{3})^2 = -1+3+2i\sqrt{(3)} = 2+2i\sqrt{3},$$
$$2.5.6 : \left(\frac{1+i}{1-i}\right)^2 = \left(\frac{\sqrt{2}e^{i\pi/4}}{\sqrt{2}e^{-i\pi/4}}\right)^2 = \frac{2e^{i\pi/2}}{2e^{-i\pi/2}} = -1,$$
$$2.5.28 : \left|\frac{z}{z^*}\right| = \left|\frac{re^{i\theta}}{re^{-i\theta}}\right| = 1, \text{ as same distance from origin}$$
$$2.5.29 : \left|(1+2i)^3\right| = \left(\sqrt{5}\right)^3 = 5\sqrt{5}.$$

Example 2.4 (Complex Solving) *Problems 2.5.35, 2.5.43 of Boas (2006).*

2.5.35 : $x+iy = 3i-4$,
$x=-4, y=3,$ as both variables are real: no mixing between parts of the equations.
2.5.43 : $(x+iy)^2 = 2ix$,
trivial solution: $x=y=0,$ or
$2ixy = 2ix \rightarrow y=1$,
$x^2-y^2=0 \rightarrow x=\pm 1$.

2.5.3 Application

Example 2.5 (Series!) *Problems 2.11.1 of Boas (2006).*

$$
\begin{aligned}
\sin(z) &\equiv z - \frac{z^3}{6} + \dots, \\
\cos(z) &\equiv 1 - \frac{z^2}{2} + \frac{z^4}{24} + \dots. \\
e^{iz} &= 1 + iz + \frac{(iz)^2}{2} + \dots, \\
&= 1 + iz - \frac{z^2}{2} + \dots. \\
e^{-iz} &= 1 - iz + \frac{(iz)^2}{2} + \dots, \\
&= 1 - iz - \frac{z^2}{2} + \dots. \\
e^{iz} + e^{-iz} &= 1 - iz - \frac{z^2}{2} + \cdots + 1 + iz - \frac{z^2}{2} + \dots, \\
&= 2(1 - \frac{z^2}{2} + \frac{z^4}{24} + \dots), \\
&= 2(\cos(z)). \\
e^{iz} - e^{-iz} &= 1 - iz - \frac{z^2}{2} + \cdots - 1 - iz + \frac{z^2}{2} + \dots, \\
&= 2i(z - \frac{z^3}{6} + \dots), \\
&= 2i(\sin(z)).
\end{aligned}
$$

2.5.4 Evaluate & Create

Example 2.6 (Euler) *Problems 2.9.1 of Boas (2006).*

$$2.9.1 : e^{-i\pi/4} = (1-i)/\sqrt{2} = \frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}.$$

Example 2.7 *Problems 2.4.2 of Boas (2006).*

$2.4.2 : -1+i; (x,y) = (-1,1); (r,\theta) = (\sqrt{2}, 3\pi/4); \sqrt{2}\cos(3\pi/4) + i\sqrt{2}\sin(3\pi/4); \sqrt{2}e^{i(3\pi/4)}$
$2.4.2^* : -1-i; (x,y) = (-1,-1); (r,\theta) = (\sqrt{2}, -3\pi/4); \sqrt{2}\cos(3\pi/4) - i\sqrt{2}\sin(3\pi/4); \sqrt{2}e^{-i(3\pi/4)}$

Example 2.8 *Problems 2.4.4 of Boas (2006).*

$2.4.4 : -\sqrt{3}+i; (x,y) = (-\sqrt{3},1); (r,\theta) = (2, 5\pi/6); 2\cos(5\pi/6) + 2i\sin(5\pi/6); 2e^{i(5\pi/6)}$
$2.4.4^* : -\sqrt{3}-i; (x,y) = (-\sqrt{3},-1); (r,\theta) = (2, -5\pi/6); 2\cos(5\pi/6) - 2i\sin(5\pi/6); 2e^{-i(5\pi/6)}$

Example 2.9 *Problems 2.5.2 of Boas (2006).*

$$2.5.2: \frac{1}{i-1} = \frac{1}{\sqrt{2}e^{i3\pi/4}} = \frac{1}{\sqrt{2}}e^{-3i\pi/4} = \frac{1}{\sqrt{2}}\left[\cos(3i\pi/4) + i\sin(-3i\pi/4)\right] = \frac{-1}{2} - \frac{i}{2}.$$

2.6 Homework Problems

Exercise 2.1 (Plots on a Plane!) *Similar to problems in 2.4 of Boas (2006). Make a figure where you plot $z = 3 + 4i$ and its complex conjugate. Label these two points, as well as their real and imaginary parts, in the multiple ways shown in the figure below (i.e., convert between the forms).*

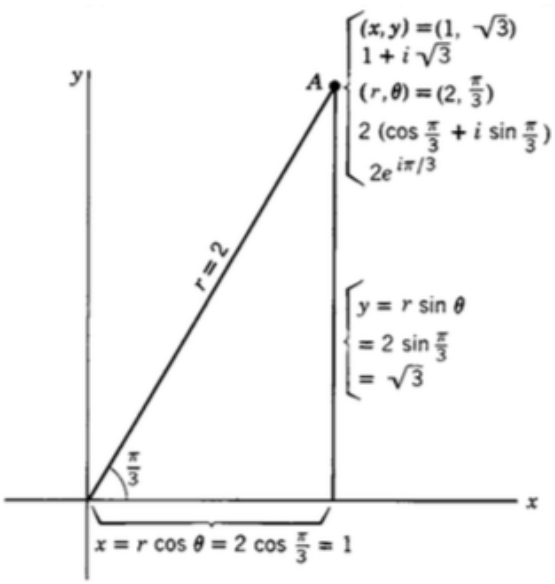


Figure 3.3

Exercise 2.2 (Crunch your abs!) *Similar to problems in 2.5 of Boas (2006). Simplify the following number, z, to the $z = x + iy$ form and to the $z = re^{i\theta}$ form. Then plot the number and all of its fourth roots in the complex plane (Boas, 2006, see Section 2.10).*

$$z = \left[\frac{1+i}{1-i}\right]^4 \tag{2.16}$$

Exercise 2.3 (Complex Solving) *Problem 2.5.44 of Boas (2006): Solve $x + iy = (1-i)^2$ for x, y.*

Exercise 2.4 (Series!) *Problems 2.11.2 of Boas (2006): Solve the equations $e^{i\theta} = \cos\theta + i\sin\theta, e^{-i\theta} = \cos\theta - i\sin\theta$, for $\cos\theta$ and $\sin\theta$ and so obtain equations (11.3) of Boas (2006).*

Exercise 2.5 (The Buoyancy, or Brunt-Väisälä, Frequency) *In a density stratified fluid, displacing a fluid parcel (without changing its density) upward or downward results in a restoring buoyancy force, because a parcel displaced upward will be denser than its neighbors and a parcel displaced*

downward will be more buoyant than its neighbors. The equation that describes the motion for the position of the parcel Z can be written

$$\frac{d^2Z}{dt^2} = -N^2 Z. \tag{2.17}$$

Where N is a function of the density (ρ) stratification in the vertical direction (z) as compared to a background density ρ_0 and gravitational acceleration g:

$$N^2 = \frac{-g}{\rho_0}\frac{d\rho}{dz}. \tag{2.18}$$

The frequency N is called the buoyancy frequency or Brunt-Väisälä frequency after David Brunt and Vilho Väisälä. Verify that $Z = e^{iNt}, Z = e^{-iNt}, Z = \cos(Nt)$, and $Z = \sin(Nt)$ satisfy this equation.

Chapter 3

Linear Algebra

3.1 Introduction–Babies, Bathwater, and Linearization

Reading: Boas (2006, 3.1, 3.2)

After discussing the roots of equations, and the complexity–and complex numbers–that arise from solving polynomial equations of order two and higher, it seems like a relief to consider only linear equations. But, what is lost? And, are linear equations interesting representations of the real world?

Let us revisit the Taylor series expansion in (1.16) and reproduced here in (3.1),

$$f(x) = \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!} f^{(n)}(a). \qquad (3.1)$$

As discussed in Chapter 1, this infinite series can be truncated, often to good accuracy if x is near a, at a finite N.

$$f(x) \approx \sum_{n=0}^{N} \frac{(x-a)^n}{n!} f^{(n)}(a). \qquad (3.2)$$

If $N = 1$, then

$$f(x) \approx f(a) + (x-a) f'(a), \qquad (3.3)$$

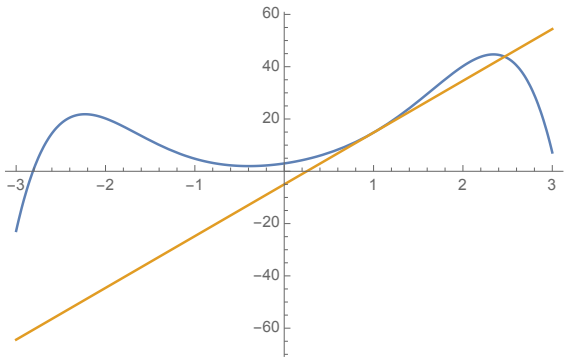


Figure 3.1: **Tangent to a function (3.3).**

which is a linear function in x. This approximation can be visualized as the tangent line to the function $f(x)$ at $x = a$ (Fig. 3.1). This approximation is as good as a linear approximation can get in a one-dimensional function of a single variable. As we saw from the Newton-Raphson method, such approximations can be valuable even in highly nonlinear systems. However, geophysical problems often consider multiple functions of multiple variables, which are even more complicated. The power of linear algebra shines in multivariate systems, because there the tools of linear algebra, matrix manipulation, and inversion techniques allow simultaneous and even-handed treatment of all of the variables and functions.

Often, the equations that govern a physical system are so complicated that we forgo a formal Taylor series expansion of a known solution and instead linearize the equations immediately, or assume that the whole dynamics are linear. This method will be discussed more in the dimensional analysis chapter, but the basic idea is to restrict consideration to situations where all the nonlinear terms are small (by selecting x very near a for example, or by choosing a region of space where $f''(a)$ is small). If they are small, then we just drop them and proceed to find a solution to the remaining (linear) set of equations. At the end, we can plug the solutions we find for the linear system back into the original (nonlinear) set of equations and evaluate the (in)accuracy of our solutions. There are even methods, called *perturbation series* or *weakly nonlinear analysis* where we use the linear solutions to estimate some nonlinear solutions (much like calculating the $n = 2$ in the Taylor series using the $n \leq 1$ terms). Very often, the key assumption in the linearization process is *small amplitude* of perturbations away from a known state, which in our simple example so far amounts to keeping x very near to a.

In practice, a linear approximation to a complicated nonlinear function is unlikely to be a very good approximation unless x is very near a (in terms of how far f remains near its tangent). But, linear equations are *so much easier to solve, especially in multivariate situations,* than nonlinear equations, that often we linearize the equations anyway. We can then study the whole class of linear solutions that do an excellent job of describing small amplitude phenomena and this study may build intuition useful in the full nonlinear system.

One key failing of small amplitude solutions (i.e., x near a) is that they lose precisely the terms in the equation that distinguish between small and large amplitude. We will see this in action many times, and exploit it to solve for all amplitudes (so long as they remain small) at the same time. However, many geophysical systems are quite sensitive to how large a perturbation to their state is (see Fig. 3.2 for a dramatic example), and to study such problems completely we need to return to incorporate amplitude-sensitive (i.e., nonlinear) terms in the equations.

3.2 Linear Systems–Line 'Em Up, Solve 'Em Down

Reading: Boas (2006, 3.2, 3.3, 3.4, 3.6, 3.7)

We can write any one-dimensional linear equation in a similar way,

$$ax = y, \tag{3.4}$$

which can be solved for x, so long as $a \neq 0$.

$$y = b/a. \tag{3.5}$$

If $a = 0$, then the linear equation does not constrain the value of x at all, and the system can only be solved if $y = 0$ as well. In that case, any value of x will do.[1] This problem has one

Figure 3.2: **A situation that's sensitive to perturbation amplitude.**

[1] I just looked at the Wikipedia page on linear equations and it is amazingly similar to what I wrote! Convergent evolution...

input or "source" (x) one coefficient (a) and one "response" or "output" y. Solving the equation amounts to asking "what source results in this output?"

You probably spent much of middle school staring at the equations for a line, such as

$$y = mx + b. \tag{3.6}$$

This equation defines a line, but *it is not a linear equation*! It is an *affine* relationship, because it doesn't intersect the origin. The equations for lines are discussed in detail in Boas (2006, 3.5).

Definition 3.1 (Linear Function or Linear Operator) *A function, or operator, f is linear if*

$$f(x_1 + x_2) = f(x_1) + f(x_2) \text{ and } f(ax) = af(x), \tag{3.7}$$

for any constant a and any pair of arguments x_1 and x_2.

Definition 3.2 (Linear Equation) *A linear equation is one consisting only of linear functions.*

Consider two points on the line, $y_1 = mx_1 + b$ and $y_2 = mx_2 + b$. Since $m(x_1 + x_2) + b = y_1 + y_2 - b$, instead of $y_1 + y_2$, this equation is not a linear one–at least when we take x and y as the only inputs and keep m and b fixed. The right hand side of this equation is an affine function, which would be linear if $b = 0$. To put it another way, linear functions can only be graphed as lines that pass through the origin. Ordinarily, it is easy to redefine the origin to make the function and equation linear.

Furthermore, we will often use the same definition of linear equation to describe differential equations–equations that involve derivatives in some of the terms. The operation of taking a derivative or integrating a function is linear (try it yourself and see!), so we can still use the definition of linear operators 3.1 and equations 3.2. However, the solutions to such equations are often not straight lines! Instead, they are commonly exponentials, logarithms, and power series.

3.3 Arrays, Vectors, Matrices, & Things in MATLAB

Reading: Boas (2006, 3.6, 3.7)

So far, we have considered linear equations with only one unknown. Linear algebra is a set of tools that generalizes the solution methods for the simple linear equation

$$ax = b \tag{3.8}$$

to the matrix equation

$$\mathbf{Ax} = \mathbf{b}. \tag{3.9}$$

Linear algebra is the study of linear functions, linear equations, their solutions, and the algebraic objects that help in finding these solutions: matrices and vectors.

In the homework for Boas (2006, 2.2), you will have opportunities to remind yourself how to solve sets of linear equations. The primary tool for solving them is adding and subtracting one equation from another and multiplying every term in an equation by a constant. An alternative to this

approach is also given in Boas (2006, 2.2), which is to express the system as a matrix equation. For example, the system of equations

$$ax + by = e, \tag{3.10}$$
$$cx + dy = f, \tag{3.11}$$

can be written as a matrix equation:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix} \tag{3.12}$$

or simply

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{3.13}$$

where the coefficient matrix is $\mathbf{A}$,

$$\mathbf{A} \equiv \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{3.14}$$

The vector of unknowns is $\mathbf{x}$,

$$\mathbf{x} \equiv \begin{bmatrix} x \\ y \end{bmatrix}. \tag{3.15}$$

And the source vector is $\mathbf{b}$,

$$\mathbf{b} \equiv \begin{bmatrix} e \\ f \end{bmatrix}. \tag{3.16}$$

Matrix algebra is a bit tricky, and involves patterns that are unfamiliar to the uninitiated. In particular, matrix multiplication is very different from ordinary multiplication–indeed, $\mathbf{AB} \neq \mathbf{BA}$! Boas (2006, 3.2, 3.3, 3.4, 3.6, 3.7) are an excellent tutorial and workout on these manipulations, and we will work through many examples in class.

The symmetries of matrices are often useful in simplifying calculations, and certain types of matrices are associated with certain types of physical phenomena. For example, symmetric matrices, where the transpose of the matrix is equal to the matrix, are very common in geophysics. Stress, diffusion, and viscosity are all associated with symmetric matrices. In our example, the matrix is symmetric if $b = c$. It is antisymmetric if it is equal to minus its transpose, which occurs here when $b = -c, a = d = 0$. Matrices involving complex entries are also possible, and often the combination of taking the complex conjugate (of each matrix component) and transpose occurs in place of the transpose for categorizing complex matrices.

In this specific example, the matrix is square, and so we are most interested in the inverse of the matrix (if it exists), which allows us to solve for the unknowns in terms of the source.

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}, \mathbf{A}^{-1} = \frac{1}{ad - cb} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{3.17}$$

We might also be interested in eigenvalues (λ_i) and eigenvectors ($\mathbf{v}_i$) of a square matrix, which are special in that when the matrix acts on an eigenvector or a vector proportional to one, it behaves as though the matrix multiplication was just multiplication by a scalar (the eigenvalue). There are only a limited number of eigenvector directions and eigenvalues, less than or equal to the number of dimensions of

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \tag{3.18}$$

$$\lambda_1 = \frac{1}{2}\left(a + d - q\right), \qquad \mathbf{v}_1 = \begin{bmatrix} \frac{a-d-q}{2c} \\ 1 \end{bmatrix}, \tag{3.19}$$

$$\lambda_2 = \frac{1}{2}\left(a + d + q\right), \qquad \mathbf{v}_2 = \begin{bmatrix} \frac{a-d+q}{2c} \\ 1 \end{bmatrix}, \tag{3.20}$$

$$q \equiv \sqrt{(a-d)^2 + 4bc} \tag{3.21}$$

As you can see, it would be easy to imagine that the eigenvalues and eigenvectors would be imaginary or complex. One useful result is that they are not for symmetric matrices. Can you see why?

Once we have the eigenvectors and eigenvalues, we can write the matrix as a product of matrices, where one of them is diagonal.

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}, \mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \mathbf{\Lambda}. \tag{3.22}$$

Here $\mathbf{V}$ is a matrix where each column has the same components as an eigenvector, and $\mathbf{\Lambda}$ is a diagonal matrix whose elements are the eigenvalues. For the 2×2 example,

$$\mathbf{V} = \begin{bmatrix} \frac{a-d-q}{2c} & \frac{a-d+q}{2c} \\ 1 & 1 \end{bmatrix}, \qquad \mathbf{V}^{-1} = \begin{bmatrix} \frac{-c}{q} & \frac{a-d+q}{2q} \\ \frac{c}{q} & \frac{d-a+q}{2q} \end{bmatrix}, \tag{3.23}$$

$$\mathbf{\Lambda} = \begin{bmatrix} \frac{1}{2}\left(a + d - q\right) & 0 \\ 0 & \frac{1}{2}\left(a + d + q\right) \end{bmatrix}. \tag{3.24}$$

This is called *diagonalizing* the matrix $\mathbf{A}$ or transforming to the eigenvector basis (when the eigenvectors $\mathbf{v}$ are rescaled to have unit length).

You may be aware that the numerical and scientific computing program MATLAB was built to exploit linear algebra to "speed the pace of discovery" as their slogan goes. Using MATLAB is, indeed, a brilliant way to get faster at solving matrix equations, inverting matrices, calculating determinants, etc. Mary Boas, on the other hand, seems quite skeptical of electronic computing instead of pencil-based computation. Expansion in minors, cofactors, and other algorithms make up much of her linear algebra chapter. I personally think this is like long division–most people don't remember how to do it and even if you did you'd still use a calculator or computer for any nontrivial problem. And, for small problems in linear algebra, learning the algorithm is often just as hard as memorizing the result. So, without further ado, here are the determinants, general solutions, and eigenvector/eigenvalue sets for 1×1 and 2×2 problems. For 3×3 problems, a few items are left out because they are messy.

3.3.1 1×1

$$ax = b, \qquad |a| = a, \qquad x = \left[\frac{1}{a}\right] b. \tag{3.25}$$

$$\lambda = a, \mathbf{v} = [1]. \tag{3.26}$$

3.3.2 2×2

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{3.27}$$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb, \tag{3.28}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{ad - cb} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix}, \tag{3.29}$$

$$\lambda_1 = \frac{1}{2}\left(a + d - \sqrt{(a-d)^2 + 4bc}\right), \quad \mathbf{v}_1 = \begin{bmatrix} a - d - \sqrt{(a-d)^2 + 4bc} \\ 2c \end{bmatrix}, \tag{3.30}$$

$$\lambda_2 = \frac{1}{2}\left(a + d + \sqrt{(a-d)^2 + 4bc}\right), \quad \mathbf{v}_2 = \begin{bmatrix} a - d + \sqrt{(a-d)^2 + 4bc} \\ 2c \end{bmatrix}. \tag{3.31}$$

3.3.3 3×3

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \tag{3.32}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11}, \tag{3.33}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{\begin{bmatrix} a_{22}a_{33} - a_{23}a_{32} & a_{13}a_{32} - a_{12}a_{33} & a_{12}a_{23} - a_{13}a_{22} \\ a_{23}a_{31} - a_{21}a_{33} & a_{11}a_{33} - a_{13}a_{31} & a_{13}a_{21} - a_{11}a_{23} \\ a_{21}a_{32} - a_{22}a_{31} & a_{12}a_{31} - a_{11}a_{32} & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11}}. \tag{3.34}$$

3.4 Vectors–Variables and Forcing

The unknowns $\mathbf{x}$ and the forcing $\mathbf{b}$ in the preceding section are both *vectors*. In the context of linear algebra, a vector is just a list of items that can be operated on by matrices, or equivalently

the input or output of a linear function. In terms of matrix algebra, you can just treat a vector as a single row or a single column matrix, but there is more to them than just that, because a collection of vectors defines a basis of possible inputs and outputs of matrices.

A vector has a dimension, which is how many components it has. As you have probably learned, a vector also can be thought of as having a direction and a magnitude. Imagine multiplying one vector by a scalar, you can make it longer, shorter, or zero; you can even make it point backward by multiplying it by a negative scalar. In this sense a vector defines not just a single magnitude and direction, but it can–together with multiplication by a scalar–be taken to *span* a whole range of vectors that all share the same direction.

Now consider taking two vectors, and seeing what space they span. If they share the same direction, then any comination of them span just the same one-dimensional space that either one of them at a time would. We may say that they are parallel, or anti-parallel, or linearly dependent, and we know that the following is true $\mathbf{u}\cdot\mathbf{v} = \pm|\mathbf{u}||\mathbf{v}|$. If this is not true, i.e., if the two vectors point in different directions, then there is a whole two-dimensional surface of vectors that can be formed from their combination together with two scalars, $a\mathbf{u} + b\mathbf{v}$. Continuing to add more vectors, you can fill up the whole space of possibilities (i.e., the space that has as many dimensions as the vectors has components). In this case, the span of the set of vectors is the whole space. Adding more vectors will only result in ones that are linearly dependent on the others. When a set of vectors spans the whole space, you could use them as a basis set, since every vector in the space can be formed by a combination of these vectors.

However, we often like our basis sets to be as simple as possible. That is, it is most useful when they have unit length (i.e., they are *normalized*) and are orthogonal to one another (i.e., $\mathbf{u}\cdot\mathbf{v} = 0$). When they are both, they are *orthonormal*. In this case, you don't get confused about which combination of vectors is which–you can just assign every component of a vector to one of the orthonormal basis vectors. In fact, that's the essence of how we write vectors as components!

$$\mathbf{v} = (\mathbf{v}\cdot\hat{\mathbf{x}}, \mathbf{v}\cdot\hat{\mathbf{y}}, \mathbf{v}\cdot\hat{\mathbf{z}}) = (\mathbf{v}\cdot\hat{\mathbf{x}})\hat{\mathbf{x}} + (\mathbf{v}\cdot\hat{\mathbf{y}})\hat{\mathbf{y}} + (\mathbf{v}\cdot\hat{\mathbf{z}})\hat{\mathbf{z}}.$$

3.5 Superpositions

Solutions to linear equations can be added together to form more solutions to linear equations, for example

$$\mathbf{A}\mathbf{x}_1 = \mathbf{b}_1, \qquad \mathbf{A}\mathbf{x}_2 = \mathbf{b}_2, \qquad \mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{b}_1 + \mathbf{b}_2. \tag{3.35}$$

We have been associating $\mathbf{b}$ with the source or forcing that results in $\mathbf{x}$. The fact that the matrix operations are linear means that we can add these two solutions together and get another which is the response to the combined forcing $\mathbf{b}_1 + \mathbf{b}_2$. This process is called superposition of the solutions, and it is useful both for considering the combined effects of two forcings and the kinds of solutions allowed by free modes (for which $\mathbf{b} = 0$).

If you think about it a bit, you will see that superpositions and the spaces spanned by input and output vectors are closely related! If you have forcings spanning a space, what possibilities are there for the size of the space of the unknowns? It turns out that the key question is what the *rank* of the matrix $\mathbf{A}$ is. A full-rank matrix has every row and every column linearly independent from one another. In this case, the size (i.e., number of dimensions) in the forcing space is equal

to the size of the space spanned by the unknowns. If the matrix is not full rank (i.e., some rows or columns are linear combinations of the others, or the matrix is rectangular and not square), then other possibilities arise.

3.6 Counting Infinity–Under- and Overdetermination

Reading: Boas (2006, 3.8)

Like polynomials, which have as many (possibly degenerate) roots as the order of the polynomial *if* complex solutions are allowed, linear solutions also have typical numbers of solutions.

If there are as many unknowns as equations, and none of the equations are repeated or linearly dependent (basically, just repeated combinations of the other equations), and all of the variables actually appear in the equations, then there is one solution of a specified magnitude (i.e., all variables are determined). In matrix language, if the coefficient matrix is square and nonsingular (the determinant is not zero), then there is a unique solution for $\mathbf{x}$ for a given $\mathbf{b}$. I call this situation just-determined.

However, if the matrix $\mathbf{A}$ is singular (determinant is zero), then either a) one of the equations is redundant or linearly dependent on the others (that is, one row can be reduced to all zeros by row reduction) or b) one of the variables is actually not constrained by the problem (one column can be reduced to all zeros by row operations). We then say the matrix is not of full rank. In this case, the matrix equation is not really one for a square matrix, and a row or a column doesn't do anything, and one of the cases below is the equivalent.

In the case where there are more unknowns that equations (or the matrix is not of full rank), there will be a combination of the unknowns that is *not constrained*. So, it can take on any value, and there are an infinite number of solutions. The dimension of the infinite set is governed by how many extra unknowns there are (or how much less than the size of the matrix is the rank). The infinite set is said to "span" a subspace of the given dimension, called the nullspace or kernel of the matrix. For any vector in the nullspace $\mathbf{v}$, $\mathbf{A}\mathbf{v} = 0$. It's clear from this relation, that adding any amount of $\mathbf{v}$ to a solution $\mathbf{x}$ won't affect $\mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{x} + a\mathbf{v}) = \mathbf{b}$. This result is why there are an infinite number of solutions based on how much $\mathbf{v}$ you add.

Inferences from many geophysical problems are underdetermined, because it is impossible to measure everything over a broad area of the Earth (especially so underground or underwater). Inverse methods are linear algebra methods for finding the possible solutions consistent with a set of observations that underdetermine the whole problem.

In the case where there are more equations than unknowns, the system is overdetermined. Boas (2006) discounts this case as being unrealistic, but in fact it is very common in geophysics where data tends to be noisy. Thus, we repeatedly measure the same thing imperfectly, and so find multiple equations with no solution that solves them all. This is when solution optimization or regression methods, such as least-squares, are used to help sort the signal from the noise. We will return to this topic later when we study statistics. Due to the power of matrix algebra, mathematically similar inverse methods can be applied to over- and under-determined problems.

An easy way to approach over-determined and under-determined systems is the Singular Value Decomposition (SVD). This approach shares many similarities to the eigenvectors and eigenvalues, but there is a right vector and a left vector and they are different sizes. In this method

3.6.1 From Triviality to Infinity: Cramer's Rule and Free Modes

Reading: Boas (2006, 3.3, 3.8) Cramer's rule is a technique to find a solution, based on the evaluation of determinants of the matrices and specialized submatrices that constitute an equation set. In the simplest, 2×2 case, it is just a rewriting of (3.29) in terms of determinants.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix}. \tag{3.36}$$

The 3×3 version is a more complicated algorithm, because the entries inside of the matrix depend on taking other determinants. Yet, note that the solution given in 3.34 is likewise inversely proportional to the determinant of the coefficient matrix. Thus, *if the coefficient matrix has a determinant of zero, then Cramer's rule doesn't apply and there is no direct solution!*

Given the typical desire to find a solution mathematically, you might suspect that just-determined systems where the determinant is nonzero would be the rule and the most interesting cases to study. However, singular matrices and linear equations are common and interesting in interpretation of under-determined and over-determined (but noisy) observations and in small perturbation theory. Cramer's rule gives the one solution for just-determined problems. Exercise 3.5 is an example of what occurs when Cramer's rule fails for singular matrices.

Let's schematically examine such a problem. Suppose we have our usual system of equations, but suppose it is *homogeneous*, which means that the sources are all zero. To put it another way, every term in the equations is proportional to at least one unknown. Then the matrix equation is

$$\mathbf{x} = \mathbf{0}. \tag{3.37}$$

If we can invert $\mathbf{A}$, then we get the trivial solution of every variable equals zero.

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{0} = \mathbf{0}. \tag{3.38}$$

However, if we fail to be able to invert $\mathbf{A}$, then there may be nontrivial (i.e., nonzero) solutions to this equation set. The common way for the inversion to fail is if the determinant is equal to zero. So, we can use this information to make a different statement. If $|\mathbf{A}| = 0$, then nontrivial solutions may exist to the equation $\mathbf{Ax} = \mathbf{0}$. It turns out that identifying wave solutions in linear systems always follows this pattern; waves are the nontrivial solutions that occur when the coefficient matrix of the wave equations obeys $|\mathbf{A}| = 0$. In this case, a singular determinant allows free (unconstrained) modes of the system, that arise to exploit and fill the nullspace.

3.7 Vector Bases

Reading: Boas (2006, 3.3, 3.8) As you may have inferred from the discussion of nullspaces, there is an intimate connection between our conception of space and our use of vectors. We can, for example, choose vectors that are not aligned and consider the space spanned by all linear combinations of them. One vector spans a one-dimensional line, two a plane, three a 3d space, etc. We will soon see how such a system is used in geophysics to describe positions, velocities, forces, and accelerations on Earth (vector and tensor analysis).

Describing such a space is most easily done when we choose a set of basis vectors, which are normalized (unit length) and orthogonal (when they are all dotted into each other, except when dotted with themself).

3.8 Eigenproblems and EOFs

One of the most important conceptual topics from this course is the idea of *eigenvectors and eigenvalues*.[2] In matrix algebra, a square matrix $\mathbf{M}$ typically has as many eigenvectors as its size. Each eigenvector $\mathbf{v}_i$ satisfies the equation

$$\mathbf{M}\mathbf{v}_i = \lambda_i \mathbf{v}_i, \tag{3.39}$$

where λ_i is the (scalar) eigenvalue for that particular eigenvector. We can make an equation for the eigenvalue pretty easily, by introducing the identity matrix $\mathbf{I}$.

$$\mathbf{M}\mathbf{v}_i = \lambda_i \mathbf{I}\mathbf{v}_i, \tag{3.40}$$

$$(\mathbf{M} - \lambda_i \mathbf{I})\,\mathbf{v}_i = 0. \tag{3.41}$$

For the last equation to be true with a nontrivial $\mathbf{v}_i$, Cramer's rule tells us that the determinant of $\mathbf{M} - \lambda_i \mathbf{I}$ must be zero. Calculating the polynomial of the determinant and setting it to zero provides an equation whose order is the number of columns or rows of M–and which therefore has this same number of (possibly complex) roots. These roots are the eigenvalues. Once you have the list of eigenvalues, you can solve (3.40) for the eigenvectors, which can be rescaled into an orthonormal basis of orthogonal vectors of length 1.

Consider how peculiar (3.40) is, based on your experiences in the homework problems multiplying matrices and vectors. How odd it would be to multiply a vector times a matrix and receive back the same vector multiplied by a scalar! But in a deep sense, this is fundamental to the way that square matrices work. They are a linear operation mapping from the space of vectors of a given size back to vectors of the same size. Apparently, some vectors are special for each matrix, and upon them the matrix acts only as a simple scalar multiplier.

We have seen one special kind of eigenvector already, although without acknowledging it. Vectors that lie in the nullspace of a matrix are eigenvectors with eigenvalue $\lambda = 0$. To see this, compare (3.40) with the definition of a vector $\mathbf{v}$ lying the nullspace of $\mathbf{M}$: $\mathbf{M}\mathbf{v} = 0$.

In fact, the eigenvectors and eigenvalues are so intricately related to a given matrix that you can reconstruct the matrix from the eigenvectors and eigenvalues. This process is related to diagonalizing the matrix, since a key matrix in the process is the similar diagonal. If we collect all of the eigenvectors–scaled to have length 1–as columns of a matrix $\mathbf{C}$, then when it is left-multiplied by $\mathbf{M}$, we find each column obeys (3.40), and

$$\mathbf{M}\mathbf{C} = \mathbf{C}\begin{bmatrix} \lambda_1 & 0 & 0 & \\ 0 & \lambda_2 & 0 & \cdots \\ 0 & 0 & \lambda_3 & \\ & \vdots & & \ddots \end{bmatrix} \tag{3.42}$$

[2] *Eigen* is a German adjective that roughly translates to proper or characteristic. Some books prefer characteristic to eigen, but I do not, since characteristic applies to many things in English whereas eigen seems to be used only with this particular meaning.

This can be understood in a number of ways. The diagonalization approach considers $\mathbf{C}$ as a rotation operation which transforms the matrix $\mathbf{M}$ into a different coordinate system where it is diagonal:

$$\mathbf{C}^{-1}\mathbf{M}\mathbf{C} = \mathbf{D}. \tag{3.43}$$

The other form of this equation that is useful is the one where $\mathbf{M}$ is constructed from the eigenvalues and eigenvectors. It is

$$\mathbf{M} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1}. \tag{3.44}$$

Thus, there is nothing about $\mathbf{M}$ that is unknown when all of the eigenvectors and eigenvalues are known.

One last consideration about (3.44) worth mentioning is the determinant. In Boas (2006, 3.9) it is shown that the determinant of the matrix product of square matrices is the product of the determinants. Since $\mathbf{C}$ is a matrix formed of orthonormal vectors, its determinant is 1 or -1, and thus the determinant of $\mathbf{C}^{-1}$ will be the inverse: $1/1$ or $-1/1$, respectively. Thus, (3.44) implies that $\mathbf{M}$ and $\mathbf{D}$ have the same determinant. Since $\mathbf{D}$ is diagonal, its determinant is just the product of the diagonal values. The important result, then, is that *the determinant of square matrices is the product of the eigenvalues.*

3.8.1 Singular Vectors, Empirical Orthogonal Functions, and Principal Component Analysis

Sometimes, when addressing under- or over-determined problems, we would like to take advantage of something similar to eigenvectors and eigenvalues. First, let's consider how we might form a square matrix ($\mathbf{M}$) related to a non-square one ($\mathbf{N}$). Well, $\mathbf{N}\mathbf{N}^\intercal$ is square, as is $\mathbf{N}^\intercal\mathbf{N}$, but they are not the same matrix (think about the size of each). But, we could form the eigenvectors and eigenvalues of $\mathbf{N}\mathbf{N}^\intercal$ and $\mathbf{N}^\intercal\mathbf{N}$ and think about what they mean.

This idea is the basis of Singular Vector Decomposition (or SVD). The eigenvectors of $\mathbf{N}\mathbf{N}^\intercal$ and $\mathbf{N}^\intercal\mathbf{N}$ can be used to deconstruct $\mathbf{N}$ into

$$\mathbf{N} = \mathbf{U}\mathbf{S}\mathbf{V}^\intercal, \tag{3.45}$$

where $\mathbf{S}$ is a diagonal matrix with the *singular values* on the diagonal, the columns of $\mathbf{U}$ and the columns of $\mathbf{V}$ are called the left-singular vectors and right-singular vectors of $\mathbf{N}$, respectively. Like the eigenvector matrix, $\mathbf{U}$ and $\mathbf{V}$ are both unitary matrices so their columns (the singular vectors) are an orthonormal set of vectors.

Consider $\mathbf{N}\mathbf{N}^\intercal$ and $\mathbf{N}^\intercal\mathbf{N}$. Since both product involve $\mathbf{U}\mathbf{U}^\intercal = \mathbf{I}$ or $\mathbf{V}\mathbf{V}^\intercal = \mathbf{I}$, the singular value decomposition times its transpose collapses into the eigenvector/eigenvalue equation (3.44), with $\mathbf{S}^2 = \mathbf{D}$. Thus, *the singular values are the square roots of the eigenvalues of* $NN^\intercal$ *or* $N^\intercal N$, and the singular vectors are the eigenvectors of one or the other as well.

One application of singular value decompositions deserves special mention here, because of its importance in geophysics. The determination of Empirical Orthogonal Functions, also known as Principal Component Analysis, is a popular way to compress large datasets and draw attention to certain types of variability.

Suppose $\mathbf{N}$ is an $(N_x N_y) \times N_t$ matrix of a variable interpolated onto a uniform grid at a series of snapshot times. Thus, a column of $\mathbf{N}$ contains all $(N_x N_y)$ of the grid points at a given time, and a row of $\mathbf{N}$ every value at a given location in each of the N_t snapshots. Then the EOFs are the left singular vectors of $\mathbf{N}$ (rearranged onto the spatial grid).

$$\mathbf{N} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \tag{3.46}$$

One of the matrices $\mathbf{U}$ and $\mathbf{V}$ is square of size $\min(N_x N_y, N_t)^2$, and the other will be rectangular of size $\max(N_x N_y, N_t) \times \min(N_x N_y, N_t)$.

EOF analysis is valuable because it provides the most efficient approximation to the matrix $\mathbf{N}$,

$$\mathbf{N}_{ij} \approx \sum_{k=1}^{K} \mathbf{U}_{ik}\mathbf{S}_{kk}\mathbf{V}^T_{kj}, \tag{3.47}$$

where only the singular vectors (columns of $\mathbf{U}$ and $\mathbf{V}$) corresponding to the the K largest singular values are used. Since these are the most important singular values, these are also the most important singular vectors. Thus, the most important spatial patterns of variability, and their variability in time is captured quickly and efficiently by EOF analysis. A little care is needed in interpreting these empirical modes as dynamical modes (which might be the eigenvectors of the governing system of equations), because the empirical modes are generated from only one realization of the data not the principles that really govern the system.

The EOFs may be equivalently defined as the orthogonal spatial basis functions which, when the variability is projected onto them, maximize the amount of covariance explained. In this framework, the reduction of the basis is called the Karhunen-Loève decomposition. However, the relationship between the eigenvectors/values and the singular vectors/values means that these two definitions are equivalent.

3.9 Special Matrices

Reading: Boas (2006, 3.9)

While I will not detail them here, the special matrices and operations in Boas (2006, 3.9) are important for you to be familiar with for building on later.

3.10 Logarithms Make Multiplication into Addition and Power Laws into Linear Systems–LogLog Plotting

Back in the days of sliderules, everyone had lots of intuition about how logarithms work. Sliderules are marked in logarithmic distances, which is why they work! Now that we use computers, we are not as familiar with them.

Consider the following derivatives, which you may recall results in formulae involving the "natural

logarithm" ln,

$$\frac{d}{dx}\ln(x) = \frac{1}{x}, \tag{3.48}$$

$$\frac{d}{dx}2^x = 2^x \ln(2), \tag{3.49}$$

$$\frac{d}{dx}10^x = 10^x \ln(10). \tag{3.50}$$

That differentiation rule is easy to apply, but there is a special case to consider. Note that $\ln 2 \approx 0.6931 < 1$ and $\ln 10 \approx 2.302 > 1$. So, the natural logarithm of big numbers is greater than one and the natural logarithm of little numbers is smaller than one. Which leads us to the magic number $e \approx 2.718$, the Euler number. It's the breakpoint in natural logarithm between little and big, for which $\ln e = 1$. This implies the critically important rule about e^x, which is that it is equal to its derivative.

$$\frac{d}{dx}e^x = e^x \ln(e) = e^x, \tag{3.51}$$

$$\text{as } \ln(e) = 1. \tag{3.52}$$

However, logarithms are still extremely useful, especially when combined with the tools of linear algebra. For example, suppose we measure a physical quantity that is the product of three factors, each with a different power:

$$f(A, B, C) = A^a B^b C^c. \tag{3.53}$$

Inverting for A, B, C from measurements of f seems difficult. However, suppose we take the logarithm of this equation first, and use the rules that govern them? Then,

$$\ln f(A, B, C) = \ln\left(A^a B^b C^c\right) = a\ln(A) + b\ln(B) + c\ln(C). \tag{3.54}$$

Thus, if we first take the logarithm, we arrive at a linear problem instead of a complicated nonlinear one! Futhermore, there was nothing special about using the natural logarithm, $\log_2$ or $\log_{10}$ would have worked in just the same way. In the next chapter, we will exploit this conversion from products of powers to linear functions to derive a deep and useful theorem about dimensional analysis.

Relatedly, when plotting linear or nearly so functions, it is a good idea to use carefully chosen ranges for the axes to show the features one desires. However, when plotting functions that are exponential or products of powers of the independent variables, it is a good idea to use semi-log or log-log plots (or log paper if you're doing it old school). On semi-log plots (where one axis has each power of ten at equal distance), exponentials come out as linear, as

$$y = A10^{mx}, \text{ will appear as } y' = mx' + \log_{10} A. \tag{3.55}$$

On log-log plots, you get the same result as if you take the logarithm of both sides of the equation, so as in (3.54), power law relationships are converted to straight lines. That is, by stretching the coordinates on each axis, it has the same effect as plotting $y' = \log_1 0(y), x' = \log_1 0(x)$ on linear coordinates.

3.11 Derivatives are Linear–Intro. to Diff. Eq.

Many of the equations of geophysics are differential equations, which relate the rate of change in space or time of variables to combinations of the variables themselves. One reason why linear algebra is so important is that differential equations are often linear equations, as *differentiation and integration are linear operations.*

Thus, in many ways linear differential equations can be treated with similar concepts as in linear algebra. We can think of functions spanning parts of function space, as functions being linearly dependent (e.g., $\sin t, \cos t$ and $10 \sin t + 20 \cos t$) and functions being linearly independent ($\sin t$ and $\cos t$). A key tool in this translation is the Wronskian, which is a matrix of functions defined on a set of functions $f_1(x), f_2(x), \ldots, f_n(n)$ as

$$\begin{bmatrix} f_1(x) & f_2(x) & \ldots & f_n(n) \\ f_1'(x) & f_2'(x) & \ldots & f_n'(n) \\ & & \ldots & \\ f_1^{(n)}(x) & f_2^{(n)}(x) & \ldots & f_n^{(n)}(n) \end{bmatrix} \tag{3.56}$$

The eigenvalues of the Wronskian help describe the space spanned by these functions and their derivatives. The determinant of the Wronskian will reveal linear independence, dependence, and whether a set of functions and their derivatives can be used a basis of dimension n.

3.12 Example Problems

3.12.1 Manipulation

Example 3.1 (Solving Equations) *Problems 3.2.3 and 3.2.5 of Boas (2006). For these, write the equation set as a matrix and row reduce. Determine the solutions, or if there are no solutions or an infinite set of solutions.*

3.2.3:

$$\begin{aligned} x - 2y &= -13 \\ -4x + y &= 17 \end{aligned}$$

Or,

$$\begin{bmatrix} 1 & -2 & -13 \\ -4 & 1 & 17 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & -13 \\ 0 & -7 & -35 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -13 \\ 0 & 1 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 5 \end{bmatrix}$$

Unique solution of $x = -3, y = 5$. 3.2.5:

$$\begin{aligned} 2x + y - z &= 2 \\ 4x + 2y - 2z &= 3 \end{aligned}$$

Or,

$$\begin{bmatrix} 2 & 1 & -1 & 2 \\ 4 & 2 & -2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & -1 & 2 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

No solutions–inconsistent equations.

Example 3.2 (Cramer's Rule) *Problems 3.3.16 of Boas (2006).*

The equation set is

$$\begin{aligned} A \quad & -B & &= -1 \\ ikA \quad & -KB & &= ik. \end{aligned}$$

Cramer's rule gives

$$A = \frac{\begin{vmatrix} -1 & -1 \\ ik & -K \end{vmatrix}}{\begin{vmatrix} 1 & -1 \\ ik & -K \end{vmatrix}} = \frac{K+ik}{-K+ik},$$

$$|A|^2 = \frac{K+ik}{-K+ik}\left[\frac{K+ik}{-K+ik}\right]^* = \frac{K+ik}{-K+ik}\frac{K-ik}{-K-ik} = \frac{K^2+k^2}{K^2+k^2} = 1.$$

Example 3.3 *Problem 3.4.3 of Boas (2006). Prove that the diagonals of a parallelogram bisect each other.*

3.4.3: Prove that the diagonals of a parallelogram bisect each other. Proof: Denote the vectors

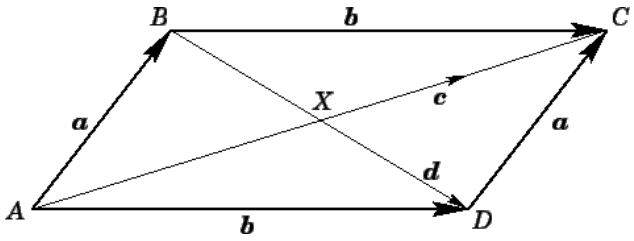


Figure 3.3: Diagram to 3.4.3 of Boas (2006).

by the order from which the go from a vertex to another, e.g., $\vec{AB}$ goes from point A to point B, then

$$\vec{AX} + \vec{XB} = \vec{AB} = \vec{DC} = \vec{DX} + \vec{XC}.$$

The central equality is by noting that those vectors are equal in magnitude and direction (although we draw them originating from different points), the others result from vector addition rules. Now, rearrange the equation,

$$\vec{AX} - \vec{XC} = \vec{DX} - \vec{XB}.$$

The two vectors on the left side of the equation point in one direction, while those on the right side point in a different direction. Thus, the only way this equality can hold is if both sides are zero, thus

$$\vec{AX} = \vec{XC}, \qquad \vec{DX} = \vec{XB}.$$

Quod erat demonstrandum.

Example 3.4 (Eigenproblems) *Problem 3.11.12 of Boas (2006).*

We could use the formulas for the 2×2 eigenproblem in (3.30). However, the related homework exercise requires you to perform the evaluation algorithm on a 3×3 problem, so the general algorithm will be exemplified here instead. 3.11.12:

$$\begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \mathbf{v} = \lambda \mathbf{v},$$

$$\begin{vmatrix} 1-\lambda & 3 \\ 2 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda) - 6 = \lambda^2 - 3\lambda - 4 = 0,$$

which is solved by factoring or using the quadratic equation to yield two solutions.

$$\begin{aligned} \lambda_1 &= -1 \\ \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \mathbf{v} &= -1\mathbf{v}, \\ v_x + 3v_y &= -v_x, \\ 2v_x + 2v_y &= -v_y, \\ \mathbf{v}_1 &= (3, -2), \end{aligned} \qquad \begin{aligned} \lambda_2 &= 4, \\ \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \mathbf{v} &= 4\mathbf{v}, \\ v_x + 3v_y &= 4v_x, \\ 2v_x + 2v_y &= 4v_y, \\ \mathbf{v}_2 &= (1, 1) \end{aligned}$$

3.13 Homework Problems

3.13.1 Manipulation

Exercise 3.1 (Solving Equations) *Problems 3.2.7 and 3.2.9 of Boas (2006). For these, write the equation set as a matrix and row reduce. Determine the solutions, or if there are no solutions or an infinite set of solutions.*

Exercise 3.2 (Matrices) *Problem 3.4.1 of Boas (2006).*

Exercise 3.3 (Vectors) *Problems 3.4.9, 3.4.12, 3.4.15 of Boas (2006).*

Exercise 3.4 (Eigenproblems) *Problems 3.11.15 of Boas (2006).*

3.13.2 Application

Exercise 3.5 (Nontrivial Cramer's) *In Pedlosky (1987), Cramer's rule is repeated used to determine the dispersion relation for waves and instabilities that solve complex linear systems of equations. One example is the derivation of Kelvin and Poincaré waves in a channel (x is along-channel distance and y is the cross-channel distance, and L is the channel width). The waves (in displacement of the ocean surface, or η) are assumed to have the form*

$$\eta = Re\left(A \cos \alpha y + B \sin \alpha y\right) e^{i(kx - \sigma t)} \tag{3.57}$$

The parameter k is the wavenumber in x, σ is the frequency, and α is the wavenumber in y. A and B are amplitudes. In the derivation, the wave equations were used to show that must α depend

on other parameters (C_0, a typical wave speed and f, the Coriolis parameter) in the following way: $\alpha^2 = \frac{\sigma^2 - f^2}{C_0^2} - k^2$. The remaining equations (the boundary conditions at the walls of the channel) were boiled down to the following linear equations on A and B.

$$\alpha A + \frac{fk}{\sigma} B = 0, \tag{3.58}$$

$$\left[\alpha \cos \alpha L + \frac{fk}{\sigma} \sin \alpha L\right] A + \left[\frac{fk}{\sigma} \cos \alpha L - \alpha \sin \alpha L\right] B = 0. \tag{3.59}$$

Using Cramer's rule, prove that: a) If the determinant of the coefficients of A and B doesn't vanish, then the only solution is $A = 0$, $B = 0$. b) That a nontrivial solution is possible if the determinant vanishes, and show that a vanishing determinant is equivalent to the condition (called the dispersion relation which is used to solve for frequency given wavenumber or vice versa):

$$(\sigma^2 - f^2)(\sigma^2 - C_0^2 k^2) \sin \alpha L = 0. \tag{3.60}$$

Finally, c) the equations for A and B are linear, but the dispersion relation between σ and k is not. Which operation in the use of Cramer's rule will virtually guarantee nonlinear polynomials? (Hint: the order of the polynomials will be closely related to the number of columns or rows in the coefficient matrix)

3.13.3 Scheming Schematics and Articulate Analysis

Exercise 3.6 (Linear Independence) *Problem 3.8.8 Boas (2006). Show that $\sin x$ and $\cos x$ are linearly independent. To do so, evaluate the Wronskian (the matrix of functions and their derivatives), and show that it's determinant is not zero.*

Exercise 3.7 (Special Operations) *Problem 3.9.3 Boas (2006).*

Chapter 4

Dimensional Analysis

4.1 Introduction–As a Rule, You Need a Ruler

Unlike mathematics, science is grounded in observations. Observations are taken with reference to scale, and in some cases those references can be fundamental (e.g., the speed of light in a vacuum) or conventional (Système International d'Unités) or convenient (whatever stick or piece of string comes to hand) or based on the problem (the depth of a relevant body of water). Since these scales are totally chosen by us, no physical law can depend on the choice, since another scientist might have measured exactly the same objects with a different choice of ruler–same phenomenon, same physics, different numbers to measure it.

Dimensional analysis is a formal way of handling these issues, and varies in utility depending on the experiment. Sometimes, it is only a useful way to double-check your calculations. Sometimes, it helps avoid repeating experiments that are not fundamentally different from one another. Sometimes it can reveal underlying aspects of a problem that can be used to develop surprising scaling laws. Always it is useful to help in designing experiments.

4.1.1 Dimensional and Dimensionless

Since measurements are taken with reference to a particular scale, they are typically dimensional quantities. Depending on what is being measured the units will vary. Units, such as meters, furlongs, astronomical units, etc., are conventional. Dimensions, such as length, time, temperature, can be measured with a variety of units, but every term in an equation must have the same dimensions. Physical constants or processes are the only way to relate among different dimensions. Thus, it is fair to compare apples to apples, or a dozen apples to an apple, but not apples to oranges (unless you first convert them into a common currency, e.g., price in $). It is generally a good idea to keep all of the units the same for a particular dimension, i.e., always use meters or feet to measure length. If you mix up feet, inches, meters, centimeters, you will have to convert often.[1]

The common system of units for science is the SI (Système International d'Unités), whose consistency and wide acceptance has led to its use being required by many scientific journals. In the SI,

[1] Or else: http://mars.jpl.nasa.gov/msp98/news/mco990930.html

there are seven base units for seven dimensions:

1. Meters (m) measure distance, L.
2. Seconds (s) measure time, T.
3. Kilograms (kg) measure mass, M.
4. Kelvin (K) measure temperature, θ.
5. Moles (mol) measure quantity, N.
6. Amperes (A) measure electric current, I.
7. Candelas (cd) measure luminous intensity, I_v

The most commonly encountered in geophysics are the first 5. All other units accepted in SI are either derived from these quantities (such as the Newton, $1\,\mathrm{N} = 1\,\mathrm{kg\,m\,s^{-2}}$, which measures force and the Pascal for pressure or stress: $1\,\mathrm{Pa} = 1\,\mathrm{N\,m^{-2}} = 1\,\mathrm{kg\,s^{-2}\,m^{-1}}$), or are measurable meaningful quantities (a day, a year, an astronomical unit)

You can write equations whose units do not match, but whose dimensions do, for example,

$$100\,\mathrm{cm} = 1\,\mathrm{m}. \tag{4.1}$$

You cannot meaningfully write equations whose dimensions do not match, such as $1\,\mathrm{m} = 4.3\,\mathrm{d}$.[2] Once you have written an equation such as (4.1), you can rearrange it into a fancy way for writing 1,

$$1 = \frac{1\,\mathrm{m}}{100\,\mathrm{cm}}. \tag{4.2}$$

In this form, it is extremely useful, as you can multiply anything in an equation with units of centimeters with this number and arrive at meters–it is the conversion factor. For example,

$$15.2\,\mathrm{cm} = 15.2\,\cancel{\mathrm{cm}}\frac{1\,\mathrm{m}}{100\,\cancel{\mathrm{cm}}} = 0.152\,\mathrm{m}.$$

Many others exist, and it is useful to think of them all as writing fancy versions of 1.

If the units can't control the physics, then what can? Dimensionless ratios that compare the amount of different physical effects. For example, we might wish to compare the rate at which momentum is delivered to a location by incoming fluid to the rate at which it is dissipated by viscosity at that point. Both effects are part of the momentum equation, and their ratio is the *Reynolds number*. Some common dimensionless parameters are given in the following table.

[2] One must be careful at the extremes, however. It was once thought that neither mass nor energy could be destroyed. We now know they can be converted from one to another in nuclear and other high-energy reactions. Furthermore, length and time are related at relativistic speeds (the speed of light is not just an ordinary speed, it actually converts space to time). Einstein was smart enough to break the rules and still be right, but even $E = mc^2$ is dimensionally consistent! Einstein (1956) describes what a powerful motivator the apparent dimensional coincidence of inertial mass equalling gravitational mass was; hence general relativity.

Name	Definition	Description
Reynolds (Re)	$\frac{UL}{\nu}$	Inertia to Friction
Rossby (Ro)	$\frac{U}{fL}$	Inertia to Coriolis
Froude (Fr)	$\frac{U}{c}$	Velocity to Wave Speed
Mach (Ma)	$\frac{U}{c_s}$	Velocity to Sound Speed
Strouhal (St)	$\frac{\omega L}{U}$	Oscillation to Velocity
Prandtl (Pr)	$\frac{\nu}{\kappa}$	Viscosity to Diffusivity
Richardson (Ri)	$\frac{-gd\rho/dz}{\rho_0(dU/dz)^2}$	Stratification to Shear
Ekman (Ek)	$\frac{\nu}{fL^2} = \frac{\mathrm{Ro}}{\mathrm{Re}}$	Viscosity to Coriolis
Grashof (Gr)	$\frac{g\Delta\rho L^3}{\nu^2\rho_0}$	Convection to Viscous
Rayleigh (Ra)	GrPr	Conduction versus Convection (Diffusivity)
Euler (Eu)	$\frac{\Delta p}{\rho U^2}$	Pressure Change versus Kinetic Energy

The value of the dimensionless parameters quantifies how important each physical effect is. Ideally, a numerical model or laboratory experiment would duplicate all of the dimensionless parameters of the real system, this situation is called *dynamical similitude*. In practice, numerical accuracy and stability, and the size of the laboratory equipment, as well as the mechanical properties of available fluids and solids (e.g., viscosity, diffusivity, Lameé parameters), mean that only some dimensionless parameters can be brought to the same values as the natural system.

Dimensional consistency is the proper handling of dimensions and units in an equation, where every term has the same dimensions, functions operate only on dimensionless quantities, and all units are clear.

4.1.2 Scaling Equations

Each term in an equation must have the same dimensions, and the units must be specified. Some terms scale naturally with one set of units and others with another. Each term will formed from a product of quantities, each with dimensions and units. It is typically the case that some terms are important under certain parameter ranges and others under other parameter ranges. For example, viscosity in a liquid is much more important on small scales than on large scales. How do we analyze such relationships in a meaningful (and dimensionally consistent) manner?

All physical laws have some of the following fundamental dimensions: mass (M), length (L), time (T), temperature (θ), electric current (I), and luminous intensity (I_v). Units, however, can be chosen arbitrarily, so long as they are the appropriate dimensions for the scales to be measured. Every term in an equation must have the same dimensions: we can no more say that length = time than apples = oranges. A different observer might choose to change the units of length independently of the units of time, and invalidate the equation. If, on the other hand, we introduce a particular *scale*, such as a velocity like the speed of sound c_s, then we can write length = $c_s \times$ time. Now both sides of the equation have dimensions of L, and thus each scale for distance is associated with a timescale (the time it takes sound to travel a distance L).

4.1.3 A Handy Set: Scaling an O(1) Theory

Consider the x momentum equation for a rotating, constant density fluid.

$$\underbrace{\frac{\partial u}{\partial t}}_{U/T} + \underbrace{u\frac{\partial v}{\partial x} + v\frac{\partial u}{\partial y}}_{U^2/L} + \underbrace{fu}_{fU} = -\underbrace{\frac{1}{\rho_0}\frac{\partial p}{\partial y}}_{P/\rho_0 L} + \underbrace{\nu\frac{\partial^2 u}{\partial z^2}}_{\nu U/L^2}. \tag{4.3}$$

We do not have to understand the equation fully to begin to consider which terms may be important by considering how large each is likely to be. This is called scaling the terms in the equations, and it is not meant to solve an equation, merely to identify which terms are most important to measure accurately and which are likely to be negligible. This equation describes the rate of change of the u velocity, based on u itself, as well as velocity in the other directions (v, w), the Coriolis parameter f, the density ρ_0, the pressure p, and the viscosity ν. We consider the units to be measured in U for velocity, T for time, L for length, P for pressure, and we know the dimensions of Coriolis parameter $[f] = 1/T$, density $[\rho_0] = M/L^3$, and kinematic viscosity $[\nu] = L^2/T$. Here we use square brackets to denote the fundamental dimensions of any quantity. Thus, each of the terms has the units of L/T^2, so the equation is dimensionally consistent. If we can estimate some of these parameters, based on observations or the statement of a particular problem, then we can make progress into understanding the scales of the terms in this equation.

Let us suppose that we are given estimates of $L, U, T \approx L/U$. Then, we can divide through by these common factors to yield:

$$\frac{L}{U^2}\Bigg[\underbrace{\frac{\partial u'}{\partial t'}}_{1} + \underbrace{u'\frac{\partial v'}{\partial x'} + v\frac{\partial u'}{\partial y'}}_{1} + \underbrace{\frac{1}{\mathrm{Ro}}u'}_{\mathrm{Ro}^{-1}} = -\underbrace{\mathrm{Eu}\frac{\partial p'}{\partial y'}}_{\mathrm{Eu}} + \underbrace{\frac{1}{\mathrm{Re}}\frac{\partial^2 u'}{\partial z'^2}}_{\mathrm{Re}^{-1}}\Bigg], \tag{4.4}$$

where we have used the common geophysical dimensionless ratios from the table above, and all primed variables are the dimensionless versions of the original variables, e.g., $x' = x/L, u' = u/U$. Without further information, we don't know which terms are important, but we have estimates of U, L, and we know the fluid we are studying so we may have an idea of ν as well. The Coriolis parameter, f, is a function of the length of day and latitude (Exercise 1.5), so that is known. The Euler number is likely to balance whatever the largest other term in the equation is, so here it is likely to be $O(1)$. At small scales, such as in the lab, we expect $\mathrm{Ro} \gg 1$, $\mathrm{Re} < 1$, and $\mathrm{Eu} \sim \mathrm{Re}$.[3] At large scales, such as those for oceanic and atmospheric motions, we expect $\mathrm{Ro} \ll 1$, $\mathrm{Re} \gg 1$, and $\mathrm{Eu} \sim \mathrm{Ro}^{-1}$. Note that as long as Ro and Eu and Re are not equal to zero, we can multiply all terms in the equations by these factors or their inverses as needed.

However, this approach presumes we already know a lot about the problem and its governing dynamics. Many problems are much less well known, and dimensional analysis is even more important in those cases. Consider the maximum speed of a sailboat of length ℓ. We know the boat makes waves as it passes, so gravity is likely to be important, so we consider also the gravitational acceleration g. The density of the water might be important, so we consider ρ_0. Finally, viscosity may be important, so we consider ν. We are looking for a velocity U, with dimensions of L/T, so let's consider the most general product of these factors we can, which we might consider as a term in

[3] The $\sim$ symbol means that the two quantities are of the same size, or scale together, though they are not necessarily proportional.

an equation for U

$$U \propto g^a \rho_0^b \nu^c \ell^d, \tag{4.5}$$

$$[U] = L^1 T^{-1} = [g]^a [\rho_0]^b [\nu]^c [L]^d = L^a T^{-2a} M^b L^{-3b} L^{2c} T^{-c} L^d. \tag{4.6}$$

Now, consider the exponents of each dimension. They must be equal on both sides of the equation, and they form a linear set of equations! So, we can use the tools we found in the last chapter to analyze how many solutions there are, and what they imply. If we order these conditions by all of the possible contributions to L, T, and M, respectively,

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 2 & 1 \\ -2 & 0 & -1 & 0 \\ 0 & -3 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}. \tag{4.7}$$

We can immediately see that we have fewer equations than unknowns. In fact, we have as many equations as we have fundamental dimensions $d = 3$ and as many unknowns as we have potential dimensional parameters $p = 4$. So, we have an underdetermined system and expect at least one free parameter.

Proceeding toward solving, we see quickly that $b = 0$, so density cannot play a role in the velocity. This result is easy to understand, since no other variable has dimensions of mass, so no equation can be written including ρ_0 which would be independent of a choice of units of mass. If we neglect viscosity, setting $c = 0$, then there we find a solution, $U \propto g^{1/2}\ell^{1/2}$, or $\mathrm{Fr} \sim 1$, where we recognize a Froude number formed from the surface wave speed $g^{1/2}\ell^{1/2}/(2\pi)$. If, on the other hand, we neglect g, setting $a = 0$, then we find the solution $U \propto \nu/\ell$, or $\mathrm{Re} \sim 1$. If we keep them both, then we arrive at a solution of the form,

$$U \propto g^{1/2} l^{1/2} \mathrm{Re}^{-c}, \text{or} \tag{4.8}$$

$$U \propto \nu \ell^{-1} \mathrm{Fr}^{-a/2}. \tag{4.9}$$

That is, we can use one dimensionless ratio to provide the basic scaling, but then we cannot say how much the other may play a role, since the linear system for the coefficients is underdetermined. The convenient named dimensionless ratios were singled out here, but any combination or power of Fr and Re is also dimensionless, so could be used instead (of course the true physical laws would counter any arbitrary choices in definitions).

Measurements in a wave tank or of boat speeds over a range of realistic Froude and Reynolds numbers reveal that the speed of a large boat in water is very nearly $U = 0.4 g^{1/2}\ell^{1/2}$, because the production of waves by pushing water out of the way of the bow is the key process requiring power to move the boat. Note that the dimensionless number 0.4 cannot be determined by dimensional analysis techniques, only by measurements. Note also that the fact that Reynolds number correction is negligible for large boats in water (rather than smaller boats in a more viscous fluid) cannot be revealed by dimensional analysis.

4.2 Buckingham Pi

In the preceding problem, we found two dimensionless parameters in a problem with 4 dimensional parameters p plus $r = 1$ dimensional results and three fundamental dimensions d. Let's consider

a generic formula for a product of parameters, in a generic dimensional equation, then determine how many dimensionless parameters are possible. We write the generic equation as

$$F_0 = F_1 + F_2 + \dots \tag{4.10}$$

We suppose that solving this equation will deliver r dimensional results (maybe more than one for generality), involving p dimensional parameters, and d fundamental dimensions. Every term in the equation will have the same dimensions, which can generically be represented by the $d \leq 7$ dimensions that appear $[F_0] = T^{a_1} L^{a_2} M^{a_3} \theta^{a_4} N^{a_5} I^{a_6} {I_v}^{a_7}$ (square brackets denote dimensions of a quantity, not value). It will also have a scale, composed of a combination of the p dimensional parameters D_k: $\{F_0\} = \prod_{k=1}^{p} {D_k}^{c_k}$ (curly brackets denote scale of a quantity, not value). We will suppose that we can always arrange the combination of scales to be a product: if a more complicated function than a product of scales is needed, then a Taylor series expansion can be used and the result will still be in the form of (4.10). Likewise, any term in the equation is dimensionally equivalent with F_0, so it can be written as:

$$\{F_j\} = \{F_0\} \prod_{i=1}^{N} P_i^{b_{ij}}. \tag{4.11}$$

The *pi factors* P_i are dimensionless, so any combination of them retains the dimensional covariance of the F_j. Dividing through by the scale of F_0, we can make a dimensionless equation

$$\frac{F_0}{\{F_0\}} = \frac{F_1}{\{F_0\}} + \frac{F_2}{\{F_0\}} + \dots + \frac{F_J}{\{F_0\}}, \tag{4.12}$$

$$\tilde{F}_0 = \tilde{F}_1 \frac{\{F_1\}}{\{F_0\}} + \tilde{F}_2 \frac{\{F_2\}}{\{F_0\}} + \dots + \tilde{F}_J \frac{\{F_J\}}{\{F_0\}} \tag{4.13}$$

$$\tilde{F}_0 = \tilde{F}_1 \prod_{i=1}^{N} P_i^{b_{i1}} + \tilde{F}_2 \prod_{i=1}^{N} P_i^{b_{i2}} + \dots + \tilde{F}_J \prod_{i=1}^{N} P_i^{b_{iJ}}. \tag{4.14}$$

The tildes denote dimensionless variables. Without a loss of generality one term can always be made a dimensionless constant, and the rest are dimensionless variables–which each have a scale given by a product of P_i factors (4.11). Each term in this equation is scaled by a potentially different combination of exponents (different b_{ij} matrix) of the P_i factors.

How many different P_i factors can there be? Consider the generalized form of (4.7). There will be d equations involving r results and p parameters. But, which is the solution and which is the unknown?

It is easier to consider each product in (4.14) or (4.11) as a function of each dimension. Each P_i is composed of the p dimensional parameters and r dimensional results introduced in the problem, such as ρ_0, g, U, ℓ, ν above. The d independent dimensions I could be as many as seven, but typical problems involve fewer. To boil this down to a linear matrix equation, consider the logarithm of

$$\ln[P_i] = 0 = \sum_{j=1}^{p+r} c_{ij} \ln[D_j] = \sum_{j=1}^{p+r} \sum_{k=1}^{d} c_{ij} d_{jk} \ln[I_k]. \tag{4.15}$$

which holds for every k from 1 to d (since the units of each dimension are arbitrary) and every i from 1 to N. The matrix c_{ij} has as many rows as there are P_i, and $p + r$ columns since every

dimensional parameter and result might appear, so at most there could be $p+r$ linearly independent rows or $p+r$ independent P_i factors. The matrix d_{jk} has $p+r$ rows and d columns.

The indicial equation (4.15) is homogeneous (equal to zero), so either the I_k is zero, which is not meaningful, so the matrix multiplication in $\sum_{j=1}^{p+r} c_{ij}d_{jk}$ must equal zero for every P_i. Thus, the $n+p$ possible dimensions must be orthogonal to d constraints, so there are $p+r-d$ independent combinations of P_i factors possible. It is possible that fewer P_i will result, if the rows of c_{ij} are not linearly independent. Note that a set is not unique–one can recombine products and powers of P_i factors to make another set of $p+r-d$ for convenience or to produce some of the P_i factors as familiar ones.

What does knowing the number of P_i factors tell us about our generic dimensional (4.10) and dimensionless (4.12) equations? Suppose there is only one dimensionless grouping. Then, (4.14) can only be:

$$\text{constant} = \tilde{F}_1 \frac{\{F_1\}}{\{F_0\}} = \tilde{F}_1 P_1, \tag{4.16}$$

$$\{\text{constant}\} \approx P_1. \tag{4.17}$$

Thus, the one dimensionless grouping is equal to a dimensionless constant $\frac{F_0}{[F_0]}$. If one thinks of all of the possible values of the one grouping as being a line, then the physical result is a point on that line. Indeed, for the sailboat speed, we found that the Froude number was effectively a constant. If there are two dimensionless groupings, then one may seek a relationship between them consisting of some universal function. For three groupings, the function becomes more complex, possibly a plane of physical values within a 3-dimensional space of all possible values, etc.

Theorem 4.1 (Buckingham Pi) *In any physical problem with p dimensional parameters and r dimensional results and d fundamental dimensions, there will be at most $p+r-d$ independent dimensionless Pi factors possible*

4.3 Homework Problems

4.3.1 Jargon to Argot

Exercise 4.1 *What is the difference between a parameter, a unit, and a dimension?*

4.3.2 Manipulation

Exercise 4.2 *Show that Ro, Re, and Ra are dimensionless by expanding their parameters into their dimensions.*

4.3.3 Application

Exercise 4.3 *(From 2.8 of Wilcox, 1997)*
The shape of a hanging drop of liquid satisfies the following empirical equation

$$\frac{(\rho - \rho_a) g d^3}{C} = \sigma \tag{4.18}$$

Where ρ, ρ_a are the densities of the drop and air, g is gravitational acceleration, d is drop diameter, σ is surface tension (units of Newtons per meter) and C is an empirical constant. What are the units of C?

Exercise 4.4 *(From 2.36 of Wilcox, 1997)*
Because of a phenomenon called vortex shedding, a flagpole will oscillate at a frequency ω when the wind blows at velocity U. The diameter of the flagpole is D and the kinematic viscosity of air is ν. Using dimensional analysis, develop an equation for ω as the product of a quantity independent of ν with the dimensions of ω and a function of all relevant dimensionless groupings.

4.3.4 Evaluate & Create

Exercise 4.5 *Use dimensional analysis on a problem of interest to your life or research, and provide at least one relevant dimensionless grouping.*

Chapter 5

Partial Derivatives

5.1 Derivatives Are Linear Operators

Reading: Boas (2006, 4.1, 4.3, 4.5)

We have been learning about linear equations and linear operators, and now we are preparing to move into *differential equations*, which involve derivatives as functions within the equations.

Definition 5.1 (ODE) *An ordinary differential equation is an equation involving derivatives with respect to one variable.*

Definition 5.2 (PDE) *A partial differential equation is an equation involving derivatives with respect to a number of variables.*

Very often, geophysical quantities depend on more that one variable. The most common example is a *field*, which is a scalar or a vector that takes on a different value at every point in space and time. Examples are temperature in the mantle (a scalar field) or velocity in the ocean (a vector field). We certainly want to know how these variables change from point to point in space and time, so we'd like to consider their derivatives with respect to directions and time. As they may vary in all directions, knowing how they vary in one direction only gives us part of the answer. Thus, these derivatives with respect to one independent variable at a time are called *partial derivatives*. To know the total change in a field, we'd need to know how all of the different partial derivatives add up.

I am assuming that you all recall your basic calculus rules for differentiation. However, we can prove the linearity of the derivative even without memorized rules.

$$\frac{\mathrm{d}}{\mathrm{d}x}\, af(x) = \lim_{h\to 0} \frac{af(x+h) - af(x)}{h} = a \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = af'(x),$$

and

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left[f(x) + g(x)\right] &= \lim_{h\to 0} \frac{f(x+h) + g(x+h) - f(x) - g(x)}{h} \\
&= \lim_{h\to 0} \left[\frac{f(x+h) - f(x)}{h} + \frac{g(x+h) - g(x)}{h}\right] \\
&= f'(x) + g'(x).
\end{aligned}$$

Some of you may have done many such expansions using the definition of the derivative in your calculus classes. Alternatively, we might have recalled the chain rule for functions of functions,

$$\big(f(g(x))\big)' = f'(g(x))g'(x). \tag{5.1}$$

The linearity of derivatives is easily shown to be a consequence of the chain rule.

Boas (2006) goes on to point out that we can consider strings of nested functions in this way.

$$\frac{\partial y}{\partial x} = \frac{\partial y}{\partial u}\frac{\partial u}{\partial v}\frac{\partial v}{\partial x}, \qquad \text{if } y = u(v(x)).$$

It is not a great leap to consider a function of two variables, $\theta(x, y)$. Suppose, for example, θ represents the temperature of a volcanic ash ejecta as it flies through the air. Tracking all of the trajectories, we might make a map of temperature at every point in space (x, y). Alternatively, we could follow one particular piece of ash as it flies, and seek its temperature as a function of time $\theta(t)$. Of course, if we knew the locations as a function of time $x(t), y(t)$, then we could use the $\theta(x, y)$ function to infer the function $\theta(t)$. Now, the idea of partial differentiation allows one to respect these relationships, so that

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial\theta}{\partial x}\frac{dx}{dt} + \frac{\partial\theta}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t},$$
$$\mathrm{d}\theta = \frac{\partial\theta}{\partial x}\mathrm{d}x + \frac{\partial\theta}{\partial y}\mathrm{d}y,$$

Notice how the partial derivative, such as $\partial/\partial x$, differs from the total derivative, d/dt, and the differential $d\theta$. The total derivative calls into play all of the arguments of the function θ, since θ depends on these variables, along with partial derivatives with respect to them and then their total derivative following the chain rule. The differential form is equivalent, but without calling upon the extra independent variable t, which may not be explicitly needed. For example, in the volcanic ash problem, one might be interested in how temperature changes with x and y without needing to know the time dependence.

5.2 Multivariate Taylor Series

Reading: Boas (2006, 4.2)

Just as the Taylor series is among the most important kinds of series for applications, the multivariate Taylor series is among the most important uses for partial differentiation. Indeed, it exposes the fundamental reason why one would want to use partial derivatives, to expose the underlying dependences of a function on other variables.

5.2.1 Multivariate Taylor Series

The Taylor series for a function depending on a number of variables $(x_1, x_2, \ldots, x_k)$ near $(a_1, a_2, \ldots, a_n)$ is

$$f(x_1, x_2, \ldots, x_k) = \sum_{n=0}^{\infty} \frac{1}{n!}\left([x_1 - a_1]\frac{\partial}{\partial x_1} + [x_2 - a_2]\frac{\partial}{\partial x_2} + \cdots + [x_k - a_k]\frac{\partial}{\partial x_k}\right)^n f(a_1, a_2, \ldots a_k). \tag{5.2}$$

The notation in the preceding is a bit messy. Note that the derivatives inside the parentheses only act on the function outside, not on the $(x_i - a_i)$ factors inside the parentheses. Furthermore, the function outside is written as being evaluated at $a_1, a_2, \ldots$, but as in the Taylor series, it is the nth derivative that is evaluated at this location so take the derivative before plugging in the values.

It is a little bit cleaner if written using the binomial theorem and considering only two variables. In this case, the expansion can be written as

$$\begin{aligned}
\text{Since } (x+y)^n &= \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k, \text{ where } \binom{n}{k} = \frac{n!}{k!(n-k)!}, \\
f(x,y) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left([x-a] \frac{\partial}{\partial x} + [y-b] \frac{\partial}{\partial y} \right)^n f(a,b) \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{[x-a]^{n-k} [y-b]^k}{k!(n-k)!} \left(\frac{\partial^n f(x,y)}{\partial x^{n-k} \partial y^k} \bigg|_{x=a, y=b} \right).
\end{aligned}$$

5.3 Approximations Using Differentials

Reading: Boas (2006, 4.8)

The Taylor series lends itself naturally to truncated series, which are often quite accurate so long as the series is convergent.

5.3.1 Linear and Near-Linear Taylor Series–Asymptotics

As we now have powerful tools to solve linear systems of equations, it is very tempting to truncate the Taylor series at the linear order $n = 1$. However, there are some things to keep in mind in the multivariate form that did not arise in the univariate form of the series.

1. We discussed how being close to $x = a$ affected the accuracy of the truncated series. In the multivariate case, the distance involves all of the variables. Thus, care is needed to be sure that the error in truncating at $(x_1 - a_1)^1$ is not very different from the error in truncating $(x_3 - a_3)^1$. If it is, then it may be necessary to go to higher order in some directions so that the overall error is correctly balanced. That is the accuracy of the truncated multivariate Taylor series will depend on how far from the estimation point the series is applied, measured in terms of how wiggly the function is *along each independent variable*.

2. Instead of finding a tangent line near a point, the truncated multivariate Taylor series will be a plane or a hyperplane tangent to the function near a point.

3. Perturbation and asymptotic methods to handle multiple variables and multiple parameters controlling how rapidly the series is likely to converge usually involve assumptions about the ordering of magnitudes of these variables and parameters. This ordering is an important part of the resulting approximate series, and different series may result from different orderings.

Similarly, it is easy to understand when a function that depends on only one function reaches a maximum or a minimum. In multiple variables, it may be a maximum in one variable and a

mimimum in another (called a saddle point because of the saddle-shaped function nearby). Folds, multi-valued functions, cusps, and all kinds of other issues arise in understanding the shape of multivariate nonlinear functions. Interestingly, many of these features arise not only in principle, but in practice with real systems. We will discuss some of these issues when we discuss chaos and dynamical systems.

5.3.2 Error Propagation in Products

Logarithms are still extremely useful for error propagation. For example, just as in (3.54) suppose we measure a physical quantity that is the product of three factors, each with a different power:

$$f(A,B) = A^a B^b. \tag{5.3}$$

We take the logarithm of this equation first, and then the find the differential relationship,

$$\frac{\mathrm{d}f}{f} = a\frac{\mathrm{d}A}{A} + b\frac{\mathrm{d}B}{B}. \tag{5.4}$$

This form is extremely useful for error propagation, as the differentials can be interpreted as the errors in each factor of the product. If the different factors are random, and their differentials can take on positive or negative values, we might be more interested in the squared value of $\mathrm{d}f$, or

$$\frac{(\mathrm{d}f)^2}{f^2} = a^2\frac{(\mathrm{d}A)^2}{A^2} + b^2\frac{(\mathrm{d}B)^2}{B^2} + 2ab\frac{\mathrm{d}A\,\mathrm{d}B}{AB}. \tag{5.5}$$

If $\mathrm{d}A$ and $\mathrm{d}B$ are random and uncorrelated, then the last term will vanish on average, so then we can write the relationship for the root mean square value of $\mathrm{d}f$ (here I use angle brackets to denote averaging),

$$\frac{\sqrt{\langle(\mathrm{d}f)^2\rangle}}{f} = \sqrt{\left\langle a^2\frac{\mathrm{d}A^2}{A^2}\right\rangle + \left\langle b^2\frac{\mathrm{d}B^2}{B^2}\right\rangle}. \tag{5.6}$$

There is normally a modest numerical difference between (5.4) and (5.6), with the latter being more accurate. The critical issue, correct in both forms, is that the higher the power of the factor the greater its contribution to the error.

5.4 Change of Variables

Reading: Boas (2006, 4.11)

The chain rule provides a mechanism for evaluating functions of functions, and one common application where functions get stacked in this way is when a change of variables is chosen. Often the change of variables is for mathematical convenience, which usually reduces the number of steps in the chain rule. Sometimes, however, the change of variables is to bring the variables closer to easy measurement or to allow a physical principle to be clear–in this cases the mathematics may be more difficult.

The D'Alembert solution to the wave equation (Example 1 of section 4.11 of Boas) is an important concept and mathematical result. We will reuse this and related forms often when we study partial differential equations.

5.5 Differentiation of Integrals

Reading: Boas (2006, 4.12) With one variable, the fundamental theorem of calculus states

$$\int_u^v \left[\frac{\mathrm{d}}{\mathrm{d}t} f(t)\right] \mathrm{d}t = f(v) - f(u). \tag{5.7}$$

It doesn't make much sense to consider taking the derivative outside of the integral,

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_u^v f(t)\, \mathrm{d}t = 0, \tag{5.8}$$

because in evaluating at the bounds u and v the dependence on t is lost. Indefinite integrals can be differentiated sort of this way, but the effect usually is just to drop the integral and the derivative (as they invert one another).

However, in multiple variables, it is quite common that the integrand or the bounds of an integral might depend on another variable. Then, one might consider differentiating with respect to that variable, which is where the Leibniz formula is useful:

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_{u(x)}^{v(x)} f(x,t)\, \mathrm{d}t = f(x,v)\, \frac{\mathrm{d}v}{\mathrm{d}x} - f(x,u)\, \frac{\mathrm{d}u}{\mathrm{d}x} + \int_u^v \frac{\partial f(x,t)}{\partial x}\, \mathrm{d}t. \tag{5.9}$$

This formula is the first of many generalizations of the fundamental theorem of calculus in multiple variables that we will find useful.

5.6 Example Problems

5.6.1 Jargon to Argot

Example 5.1 *Explain why you think the terms "partial" derivative and "total" derivative apply as they do.*

Partial derivatives describe only part of the change in a dependent variable with respect to a set of independent variables–in fact just the change associated with one independent variable at a time. Total derivatives describe the change associated with all of the independent variables in turn.

Example 5.2 *Problem 4.1.6 of Boas (2006). If $u = e^x \cos y$, prove that* $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = 0.$

$$\begin{aligned}
u &= e^x \cos y, \\
\frac{\partial u}{\partial x} &= e^x \cos y, \qquad \frac{\partial u}{\partial y} = -e^x \sin y, \\
(a)\ \frac{\partial^2 u}{\partial x \partial y} &= -e^x \sin y, \qquad \frac{\partial^2 u}{\partial y \partial x} = -e^x \sin y, \\
(b)\ \frac{\partial^2 u}{\partial x^2} &= e^x \cos y, \qquad \frac{\partial^2 u}{\partial y^2} = -e^x \cos y \longrightarrow \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.
\end{aligned}$$

Example 5.3 *Problems 4.2.4 of Boas (2006). Use the multivariate Taylor series to evaluate the expansion of e^{xy}.*

4.2.4:

$$f(x,y) = e^{xy},$$

$$\left.\frac{\partial e^{xy}}{\partial x}\right|_{x=y=0} = \left. ye^{xy}\right|_{x=y=0} = 0, \qquad \left.\frac{\partial^n e^{xy}}{\partial x^n}\right|_{x=y=0} = \left. y^n e^{xy}\right|_{x=y=0} = 0,$$

$$\left.\frac{\partial e^{xy}}{\partial y}\right|_{x=y=0} = \left. xe^{xy}\right|_{x=y=0} = 0, \qquad \left.\frac{\partial^n e^{xy}}{\partial y^n}\right|_{x=y=0} = \left. x^n e^{xy}\right|_{x=y=0} = 0.$$

Therefore, we only need to worry about mixed derivatives.

$$\left.\frac{\partial^2 e^{xy}}{\partial x \partial y}\right|_{x=y=0} = \left.\frac{\partial \left(ye^{xy}\right)}{\partial y}\right|_{x=y=0} = \left. e^{xy} + xye^{xy}\right|_{x=y=0} = 1,$$

$$\left.\frac{\partial^3 e^{xy}}{\partial x^2 \partial y}\right|_{x=y=0} = \left.\frac{\partial \left(y^2 e^{xy}\right)}{\partial y}\right|_{x=y=0} = \left. 2ye^{xy} + xy^2 e^{xy}\right|_{x=y=0} = 0,$$

$$\left.\frac{\partial^4 e^{xy}}{\partial x^2 \partial y^2}\right|_{x=y=0} = \left.\frac{\partial^2 \left(y^2 e^{xy}\right)}{\partial y^2}\right|_{x=y=0} = \left. 2e^{xy} + 2xye^{xy} + 2xye^{xy} + x^2 y^2 e^{xy}\right|_{x=y=0} = 2.$$

In fact, we only need to worry about mixed derivatives where the number of x-derivatives equals the number of y derivatives, or only the even differential orders. The value of the $\frac{\partial^{2n} f(x,y)}{\partial x^n \partial y^n}$ derivative will be $n!$. We can use the factorial form of the binomial coefficient to expand, noting that only the matched derivatives survive,

$$\begin{aligned} f(x,y) = e^{xy} &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{[x-a]^{n-k}\,[y-b]^k}{k!(n-k)!} \left(\left.\frac{\partial^n e^{xy}}{\partial x^{n-k} \partial y^k}\right|_{x=a,y=b} \right), \\ &= \sum_{n=0}^{\infty} \frac{x^n y^n}{n!n!} n! \\ &= \sum_{n=0}^{\infty} \frac{x^n y^n}{n!} \end{aligned}$$

Which is the same form as the expansion for e^z with $z = xy$.

Example 5.4 *Problem 4.11.3 of Boas (2006). Suppose that $w = f(x,y)$ satisfies $\frac{\partial^2 w}{\partial x^2} - \frac{\partial^2 w}{\partial y^2} = 1$ Put $x = u + v$, $y = u - v$, and show that w satisfies $\frac{\partial^2 w}{\partial u \partial v} = 1$. From this point, the equation is easily solved by integrating in u and v.*

$$
\begin{aligned}
w &= f(x,y), \\
\frac{\partial^2 w}{\partial x^2} - \frac{\partial^2 w}{\partial y^2} &= 1, \\
x = u+v, y = u-v \longrightarrow u &= \frac{1}{2}x + \frac{1}{2}y, v = \frac{1}{2}x - \frac{1}{2}y, \\
\frac{\partial w}{\partial x} &= \frac{\partial w}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial w}{\partial v}\frac{\partial v}{\partial x} = \frac{1}{2}\frac{\partial w}{\partial u} + \frac{1}{2}\frac{\partial w}{\partial v}, \\
\frac{\partial w}{\partial y} &= \frac{\partial w}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial w}{\partial v}\frac{\partial v}{\partial y} = \frac{1}{2}\frac{\partial w}{\partial v} - \frac{1}{2}\frac{\partial w}{\partial v}, \\
\frac{\partial^2 w}{\partial x^2} &= \frac{\partial}{\partial u}\left[\frac{1}{2}\frac{\partial w}{\partial u} + \frac{1}{2}\frac{\partial w}{\partial v}\right]\frac{\partial u}{\partial x} + \frac{\partial}{\partial v}\left[\frac{1}{2}\frac{\partial w}{\partial u} + \frac{1}{2}\frac{\partial w}{\partial v}\right]\frac{\partial v}{\partial x} = \frac{1}{4}\frac{\partial^2 w}{\partial u^2} + \frac{1}{2}\frac{\partial^2 w}{\partial u \partial v} + \frac{1}{4}\frac{\partial^2 w}{\partial v^2}, \\
\frac{\partial^2 w}{\partial y^2} &= \frac{\partial}{\partial u}\left[\frac{1}{2}\frac{\partial w}{\partial u} - \frac{1}{2}\frac{\partial w}{\partial v}\right]\frac{\partial u}{\partial y} + \frac{\partial}{\partial v}\left[\frac{1}{2}\frac{\partial w}{\partial u} - \frac{1}{2}\frac{\partial w}{\partial v}\right]\frac{\partial v}{\partial y} = \frac{1}{4}\frac{\partial^2 w}{\partial u^2} - \frac{1}{2}\frac{\partial^2 w}{\partial u \partial v} + \frac{1}{4}\frac{\partial^2 w}{\partial v^2}, \\
\frac{\partial^2 w}{\partial x^2} - \frac{\partial^2 w}{\partial y^2} &= \frac{\partial^2 w}{\partial u \partial v} = 1
\end{aligned}
$$

The last line is easily integrated with respect to u, v to solve the equation.

5.6.2 Scheming Schematics and Articulate Analysis

Example 5.5 *Problem 4.8.2 of Boas (2006). Using the two-variable Taylor series [say (2.7)] prove the following "second derivative tests" for maximum or minimum points of functions of two variables. If $f_x = f_y = 0$ at (a,b), then (In this problem, all subscripts imply partial derivatives), (a,b) is a minimum point if at (a,b),$f_{xx} > 0, f_{yy} > 0$, and $f_{xx}f_{yy} > f_{xy}^2$; (a,b) is a maximum point if at (a,b),$f_{xx} < 0, f_{yy} < 0$, and $f_{xx}f_{yy} > f_{xy}^2$; (a,b) is neither a maximum nor a minimum point if $f_{xx}f_{yy} < f_{xy}^2$. (Note that this includes $f_{xx}f_{yy} < 0$, that is, f_{xx} and f_{yy} of opposite sign.) Hint: Let $f_{xx} = A, f_{xy} = B, f_{yy} = C$; then the second derivative terms in the Taylor series are $Ah^2 + 2Bhk + Ck^2$; this can be written $A(h + Bk/A)^2 + (C?B^2/A)k^2$. Find out when this expression is positive for all small h, k [that is, all (x,y) near (a,b)]; also find out when it is negative for all small h, k, and when it has both positive and negative values for small h, k.*

We begin with the Taylor series, and we are at a point with vanishing first derivatives,

$$
\begin{aligned}
f(x,y) =& \sum_{n=0}^{\infty} \frac{1}{n!}\left([x-a]\frac{\partial}{\partial x} + [y-b]\frac{\partial}{\partial y}\right)^n f(a,b) \\
=& f(a,b) + \cancel{\left([x-a]\frac{\partial f}{\partial x} + [y-b]\frac{\partial f}{\partial y}\right)\Big|_{x=a,y=b}} \\
&+ \frac{1}{2}\left([x-a]^2\left[\frac{\partial^2 f}{\partial x^2}\right]_{x=a,y=b} + 2\,[x-a]\,[y-b]\left[\frac{\partial^2 f}{\partial x \partial y}\right]_{x=a,y=b} + [y-b]^2\left[\frac{\partial^2 f}{\partial y^2}\right]_{x=a,y=b}\right) \\
&+ \frac{1}{6}\left([x-a]^3\left[\frac{\partial^3 f}{\partial x^3}\right]_{x=a,y=b} + 3\,[x-a]^2\,[y-b]\left[\frac{\partial^3 f}{\partial x^2 \partial y}\right]_{x=a,y=b} + 3\,[x-a]\,[y-b]^2\left[\frac{\partial^3 f}{\partial x \partial y^2}\right]_{x=a,y=b} + \ldots\right.
\end{aligned}
$$

Very near the point, all terms of higher order will typically be negligible in comparison to the

second-order terms, so

$$f(x,y) \approx f(a,b) + \frac{1}{2}\left([x-a]^2 \left[\frac{\partial^2 f}{\partial x^2}\right]_{x=a,y=b} + 2\,[x-a]\,[y-b] \left[\frac{\partial^2 f}{\partial x \partial y}\right]_{x=a,y=b} + [y-b]^2 \left[\frac{\partial^2 f}{\partial y^2}\right]_{x=a,y=b} \right).$$

If $f_{xx} = \left[\frac{\partial^2 f}{\partial x^2}\right]_{x=a,y=b} > 0$ and $f_{yy} > 0$, then the first and third terms in the parentheses are positive definite. Only the middle term might cause the function to decrease. So, we consider the following:

$$[x-a]^2 f_{xx} + 2\,[x-a]\,[y-b]\,f_{xy} + [y-b]^2 f_{yy} = f_{xx}\left([x-a] + [y-b]\,\frac{f_{xy}}{f_{xx}} \right)^2 + \left(f_{xx} f_{yy} - f_{xy}^2 \right) \frac{[y-b]^2}{f_{xx}}$$

The first term on the right has the same sign as f_{xx}, and the second term has the same sign as f_{xx} if $f_{xx} f_{yy} > f_{xy}^2$. The three rules for extrema follow directly from this result.

5.7 Homework Problems

5.7.1 Manipulation

Exercise 5.1 *Problem 4.1.1 of Boas (2006). If $u = \frac{x^2}{x^2+y^2}$, find the partial derivatives of u with respect to x and y.*

Exercise 5.2 *Problems 4.2.6 of Boas (2006). Use the multivariate Taylor series to evaluate the expansion of e^{x+y}.*

Exercise 5.3 *Problem 4.4.8 of Boas (2006).*

Exercise 5.4 *Problem 4.5.1 of Boas (2006).*

Exercise 5.5 *Problem 4.12.11 of Boas (2006).*

5.7.2 Application

Exercise 5.6 (Streamfunction and Velocity Potential) *As we come to understand how fluids move, we will be interested in breaking up the velocity field into a part that results from a convergence or divergence of flow and a part that results from a swirling or rotating flow. This can be done with the Helmholtz Decomposition, which we will address in a later chapter on vector analysis. For now, we'll just test some examples.*

The relationship between velocity (u, v) and streamfunction (ψ) and velocity potential (ϕ), in 2D, is just

$$u = -\frac{\partial \psi}{\partial y} - \frac{\partial \phi}{\partial x}, \tag{5.10}$$

$$v = \frac{\partial \psi}{\partial x} - \frac{\partial \phi}{\partial y}. \tag{5.11}$$

Both velocity components, as well as ψ and ϕ should be interpreted as spatial fields–that is, they are all functions of x and y and have a value at every point in space.

The Jacobian is a useful function for evaluating the advection by flow due to a streamfunction alone, which in 2D is just

$$J(A,B) = \begin{vmatrix} \frac{\partial A}{\partial x} & \frac{\partial A}{\partial y} \\ \frac{\partial B}{\partial x} & \frac{\partial B}{\partial y} \end{vmatrix} = \frac{\partial A}{\partial x}\frac{\partial B}{\partial y} - \frac{\partial A}{\partial y}\frac{\partial B}{\partial x}. \tag{5.12}$$

a) If $\phi = 0$, show that $J(\psi, f(x,y)) = u\dfrac{\partial f(x,y)}{\partial x} + v\dfrac{\partial f(x,y)}{\partial y}$.
b) If A is a function of B, rather than independently varying in x and y, show that $J(B, A(B)) = 0$.

5.7.3 Evaluate & Create

Exercise 5.7 (The D'Alembert Wave) *Problem 4.7.23 of Boas (2006).*
If $u = f(x-ct) + g(x+ct)$, show that $\dfrac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\dfrac{\partial^2 u}{\partial t^2}$.

Chapter 6

Multivariate Integrals and Change of Variables

6.1 Integrals Are Linear Operators

Reading: Boas (2006, 5.1)

As with differentiation, it is useful to know that integrals are linear operators (What does this mean? Can you prove it?). Thus, equations containing integrals can still be handled as we have been learning to handle linear equations. Concepts such as eigenvalues and eigenvectors continue to be useful.

6.1.1 Integrals are Antiderivatives

You might have expected that if differentiation was linear, then integration would be linear, too, since they are each others inverse. The key nugget of calculus that tells us this fact is the fundamental theorem of calculus.

Theorem 6.1 *Fundamental Theorem of Calculus. If f is a continuous function that is defined over an interval including a and x, then*

$$F(x) = \int_a^x f(t)\,\mathrm{d}t, \tag{6.1}$$

Then F is also continuous and differentiable on the same interval, and

$$\frac{\mathrm{d}}{\mathrm{d}x} F(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_a^x f(t)\,\mathrm{d}t\right] = f(x). \tag{6.2}$$

Theorem 6.2 *Fundamental Theorem of Calculus–Definite Integral Corollary. If $\frac{d}{dx}F$ is a continuous function that is defined over an interval including a and x, then*

$$\int_a^b \frac{\mathrm{d}}{\mathrm{d}t} F(t)\,\mathrm{d}t = F(a) - F(b). \tag{6.3}$$

This corollary works even if $\frac{d}{dt}F(t)$ is not continuous, but nearly so (e.g., has a finite number of discontinuities).

This chapter will be devoted to integrals over multiple variables, e.g., over areas and volumes. It will take some time before we can build up to the equivalents of the fundamental theorem of calculus in multiple variables, because we need to know a bit more about vector analysis, but we will get there! When we do, we will sometimes be able to use integrals to solve equations involving derivatives, just as we sometimes used the inverse of a matrix to solve linear systems of equations.

Odd & Even

As we saw from the Taylor series and from sines and cosines, the derivatives of even functions are odd and derivatives of odd functions are even. As the (indefinite) integral is the antiderivative, the (indefinite) integral of an even function is odd, and vice versa. Thus, a very useful trick for thinking about functions that may cancel themselves out when integrated is:

Theorem 6.3 *The definite integral of an odd function is zero if the bounds of the integral are symmetric about zero.*

6.2 Multiple Integrals

Reading: Boas (2006, 5.2)

It is likely that you were taught to think of the integral of a function as the area under the curve, probably by construction as the sum of lots of little rectangles under the curve and then considering the limit as they get very thin. Therefore, let's begin to think about integrals over areas in the same way.

6.2.1 Integrals over Areas

Consider the function $z(x, y)$. It might be the elevation of a mountain, for example. What if we wanted to know the volume of the mountain? Well, volume under the surface area of the mountain is a lot like the area under a curve, and this is precisely what the multiple variable integral does. If we approximate the function $z(x, y)$ with a grid of boxes of width Δx and length Δy and height that approximates z over the interval x to $x + \Delta x$ and y to $y + \Delta y$, then this is an approximation to the integral. The integral is the limit as Δx and Δy go toward zero. For an area integral over a simple rectangular area, this limit is just

$$\int_0^{L_y} \int_0^{L_x} z(x, y)\, \mathrm{d}x\, \mathrm{d}y = \lim_{\Delta y \to 0} \lim_{\Delta x \to 0} \sum_{j=1}^{L_y/\Delta y} \sum_{i=1}^{L_x/\Delta x} \left[z(i\Delta x, j\Delta y) \right] \Delta x \Delta y. \tag{6.4}$$

This is indeed how one might approximate the integral in a numerical model with a discrete grid. As the grid boxes get smaller, there are more and more of them. Thus, the limit converges. It doesn't matter if we sum first along the x direction and then sum along the y direction, the integral is the same in this case.

We might have been more sophisticated in our selection of the area of integration. For example, we might have considered the volume under a semi-sphere, defined by the function $r^2 = x^2 + y^2 + z^2$,

or $z = \sqrt{r^2 - x^2 - y^2}$ with limits at the circle circumscribed by $x^2 + y^2 = r^2$. We would write this as

$$\int_{-r}^{r} \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} \sqrt{r^2 - x^2 - y^2}\, \mathrm{d}x\, \mathrm{d}y = \int_{-r}^{r} \left[\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} \sqrt{r^2 - x^2 - y^2}\, \mathrm{d}x \right] \mathrm{d}y, \tag{6.5}$$

where I've used square brackets to emphasize the order of integration. Instead, if we integrate in y first,

$$\int_{-r}^{r} \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \sqrt{r^2 - x^2 - y^2}\, \mathrm{d}y\, \mathrm{d}x = \int_{-r}^{r} \left[\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \sqrt{r^2 - x^2 - y^2}\, \mathrm{d}y \right] \mathrm{d}x. \tag{6.6}$$

Notice how the limits on the inner integral involve the variable that's being integrated over in the outer integral. In the rectangular case (6.4), we didn't have to worry about the shape of the rectangle because it was just as wide in x at every y value. But, the circle and the sphere have different dimensions at the edges, and the limits of the integrals in (6.5-6.6) keep track of these edges.

Instead of thinking of the volume under a surface, another common way to use an integral is to find the average value of a function. First, consider the integral

$$\int_0^{L_y} \int_0^{L_x} 1\, \mathrm{d}x\, \mathrm{d}y = L_x L_y. \tag{6.7}$$

This is the volume under a function of height 1, so it is just the area of the top (or bottom) of the rectangular solid. Similarly, the area integral over a cylinder top is

$$\int_{-r}^{r} \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} 1\, \mathrm{d}x\, \mathrm{d}y = \pi r^2. \tag{6.8}$$

Thus, the double integral can be used to find areas. But, what if we consider a more interesting integrand than 1, for example, $z(x, y)$. We've already seen that would give us the volume under the surface $z(x, y)$. Now the ratio of the two gives

$$\langle z \rangle = \frac{\displaystyle\int_0^{L_y} \int_0^{L_x} z(x, y)\, \mathrm{d}x\, \mathrm{d}y}{\displaystyle\int_0^{L_y} \int_0^{L_x} \mathrm{d}x\, \mathrm{d}y}, \tag{6.9}$$

which is just the volume divided by the area, or *the average height*. More generally, the area-weighted average of any function can be written as

$$\langle f \rangle = \frac{\displaystyle\iint_A f(x, y)\, \mathrm{d}x\, \mathrm{d}y}{\displaystyle\iint_A \mathrm{d}x\, \mathrm{d}y}, \tag{6.10}$$

Where the subscript A implies that the limits of the integral should be chosen to span the whole area. We might also consider the an average, where being high up counted more. For example, potential energy is one such quantity, where mass at higher elevations has more energy than mass at

lower elevations. To take such as weighted average, we just multiply the integrand by the weighting factor, and also multiply the denominator's integrand (to keep the correct units),

$$\frac{\langle fz \rangle}{\langle z \rangle} = \frac{\iint\limits_A f(x,y)z(x,y)\,\mathrm{d}x\,\mathrm{d}y}{\iint\limits_A z(x,y)\,\mathrm{d}x\,\mathrm{d}y}. \tag{6.11}$$

Many physical concepts rely on averages and moments, especially area averages and volume averages. Moments are averages weighted with a variable or coordinate, e.g., the first moment of f with x is $\langle xf \rangle$, the second moment is $\langle x^2 f \rangle$, etc. Mass, the center of mass, and rotational inertia are moments of density. Mean, variance, skewness, and kurtosis of a random variable are the moments of its probability distribution.

6.2.2 Integrals over Volumes

Like area averages, we can conceptualize volume integration by thinking about slicing up the volume we want to integrate. This works well when we want to conceptualize, for example, the relationship between density and mass. Density is the mass per unit volume, so if we mark boxes of size Δx, Δy, Δz with a density, then the sum over all the boxes of their density times their volume will be the total mass.

However, unlike the area under the function interpretation of the one-dimensional integral and the volume under the area interpretation of the two-dimensional integration, there is no (familiar) equivalent in the three dimensional integration. The hypervolume between a three dimensional surface and another doesn't cut it.

So, we think about three-dimensional (and higher) integrals using the other intuitions built in the last section. The average over a volume is

$$\langle f \rangle = \frac{\iiint\limits_V f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{\iiint\limits_V \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}. \tag{6.12}$$

Like with areas, when the volumes take complex shapes, care is needed to set up the bounds on the integrals. We can also think about moments and weighted averages over the volume.

One example that is particularly apt is the calculation of volume and the calculation of mass from density. The calculation of mass from density ρ is just

$$M = \iiint\limits_V \rho(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{6.13}$$

The calculation of a volume is just the same as the density calculation, except with a density equal to 1.

$$V = \iiint\limits_V \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{6.14}$$

So, the average density is just

$$\langle \rho \rangle = \frac{\iiint_V \rho(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{\iiint_V \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}. \tag{6.15}$$

You should go over the methods and examples described in Boas (2006, 5.2, 5.3)

6.2.3 How Much Stuff?

The other major application of the volume integral (and also area integrals and line integrals sometimes) is to figure out how much of something is inside a region. This integration is needed when we have a density of something, that is, amount of stuff per unit volume (e.g., density, melt fraction, concentration) or per unit area (e.g., wind stress) or per unit length (e.g., force applied is the work done per unit length displaced). Integrating the density results in the amount of stuff in the volume of integration, e.g., mass, mass of a constituent, force applied, or work done.

6.2.4 Order of Integration

Near (6.5), it was noted that care is needed with the integral bounds when reversing the order of integration. Similarly, care is needed when the integrands depend on some of the variables, but not others. As a general rule, you can reverse the order of integrations as convenient, but only if you don't take any variables outside of the integrals corresponding to their differential. That is, when you integrate in x, with the corresponding differential $\mathrm{d}x$ and bounds on the values of x, *there should be no x dependence left over*. So, no later integral bounds should reintroduce x, for example. Keep checking this rule and you'll be OK.[1]

6.3 Change of Variables

Reading: Boas (2006, 5.4)

One final crucial aspect of multiple integrals is what to do when we want to change coordinate systems. You may have noticed that $\mathrm{d}x$, $\mathrm{d}y$, and $\mathrm{d}z$ all have the units of length. What about integration in polar coordinates, or in latitudes and longitudes? How do we keep track?

A homework problem from last week touched on the idea of the Jacobian, and it comes back here. The Jacobian is the determinant of the matrix of all of the partial derivatives of one set of coordinates with respect to the other. For example, the earth coordinates (geopotential height, latitude, longitude, or $(\mathfrak{z}, \phi, \vartheta)$) are similar to spherical coordinates (radius, colatitude, longitude or (r, θ, ϕ)), but with latitude instead of colatitude[2] and with height referenced to the mean sea level (or preferably the mean geoid elevation) $r_0(\phi, \vartheta)$ to include the effects of variation in gravity and Centrifugal force from the earth's rotation. The order of the variables is altered to keep the

[1] There is a more subtle reversal of integrals that worries mathematicians more than physicists–if the integrals all exist (i.e., if the sequence with shrinking Δx converges). This issue is rarely our problem, but you might have discussed it in great detail in a multivariable calculus class.

[2] Physics books like Boas (2006) just love to mess with us by using colatitude.

coordinates right-handed, which is a symmetry we will want for vector analysis.[3] If, for simplicity, we approximate a spherical earth with r_0 as a constant, then

$$\begin{aligned} x &= (\mathfrak{z} + r_0)\cos\phi\cos\vartheta, \\ y &= (\mathfrak{z} + r_0)\sin\phi\cos\vartheta, \\ z &= (\mathfrak{z} + r_0)\sin\vartheta. \end{aligned} \tag{6.16}$$

The Jacobian is,

$$\begin{aligned} \frac{\partial(x,y,z)}{\partial(\mathfrak{z},\phi,\vartheta)} &= \begin{vmatrix} \frac{\partial x}{\partial \mathfrak{z}} & \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \vartheta} \\ \frac{\partial y}{\partial \mathfrak{z}} & \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \vartheta} \\ \frac{\partial z}{\partial \mathfrak{z}} & \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial \vartheta} \end{vmatrix} = \begin{vmatrix} \cos\phi\cos\vartheta & -(\mathfrak{z}+r_0)\sin\phi\cos\vartheta & -(\mathfrak{z}+r_0)\cos\phi\sin\vartheta \\ \sin\phi\cos\vartheta & (\mathfrak{z}+r_0)\cos\phi\cos\vartheta & -(\mathfrak{z}+r_0)\sin\phi\sin\vartheta \\ \sin\vartheta & 0 & (\mathfrak{z}+r_0)\cos\vartheta \end{vmatrix} \\ &= (\mathfrak{z}+r_0)^2\cos\vartheta \end{aligned} \tag{6.17}$$

To convert between integrals expressed in x, y, z to those in $\mathfrak{z}, \phi, \vartheta$, we use the functions (6.16) to substitute for all of the occurrences of the variables in the integrands and integral bounds, and then we switch the differentials at the end from $\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ to the Jacobian times the new differentials, or $(\mathfrak{z}+r_0)^2\cos\vartheta\,\mathrm{d}\mathfrak{z}\,\mathrm{d}\phi\,\mathrm{d}\vartheta$. Or,

$$\iiint\limits_{V(x,y,z)} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \iiint\limits_{V(\mathfrak{z},\phi,\vartheta)} f(x(\mathfrak{z},\phi,\vartheta), y(\mathfrak{z},\phi,\vartheta), z(\mathfrak{z},\phi,\vartheta))(\mathfrak{z}+r_0)^2\cos\vartheta\,\mathrm{d}\mathfrak{z}\,\mathrm{d}\phi\,\mathrm{d}\vartheta. \tag{6.18}$$

This transformation makes the units and dimensions correct, and consistently handles the volume transformations that go along with the coordinate transformation. Boas (2006) gives the Jacobians and transformation rules for polar, cylindrical, and spherical coordinates, but earth coordinates are the most common coordinates used in geophysics aside from Cartesian.

6.4 Example Problems

6.4.1 Manipulation

Example 6.1 (Areas) *Set up and evaluate integrals to calculate the area of an $L_x \times L_y$ rectangle and a radius R circle.*

$$\begin{aligned} A &= \int_0^{L_y}\int_0^{L_x} \mathrm{d}x\,\mathrm{d}y = L_x L_y, \\ A &= \int_0^{R}\int_0^{2\pi} r\,\mathrm{d}r\,\mathrm{d}\phi = \pi R^2. \end{aligned}$$

[3] For now, think of a right-handed system as one where the coordinate directions at a point in space can be formed with the thumb, first, and middle fingers of a right hand. If we switch from a right-handed to a left-handed coordinate system, the Jacobian of the transformation will be negative.

Example 6.2 (Masses) *Set up and evaluate the mass and the center of mass (density-weighted average of x, y, z) of the rectangular solid and cylinder objects in exercise 6.1, and for a sphere of radius R, when the density is a) constant $\rho(x, y, z) = \rho_0$ and b) a linear function of z, $\rho(x, y, z) = \rho_0 + a(z - z_0)$ where z_0 is the location of the midpoint along the z-axis of the cylinder. You do not have to evaluate the integral for the sphere (unless you want to). (Hint: Arrange the objects so they are centered on $x = 0, y = 0, z = 0$, and then use observations about odd and even functions to integrate without integrating).*

Rectangular Solid:

$$
\begin{aligned}
M &= \int_{-L_z/2}^{L_z/2}\int_{-L_y/2}^{L_y/2}\int_{-L_x/2}^{L_x/2} \rho_0 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \rho_0 L_x L_y L_z, \\
\overline{\mathbf{x}}\cdot\hat{\mathbf{i}} &= \frac{\int_{-L_z/2}^{L_z/2}\int_{-L_y/2}^{L_y/2}\int_{-L_x/2}^{L_x/2} x\rho_0 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{M}, \\
\overline{\mathbf{x}}\cdot\hat{\mathbf{j}} &= \frac{\int_{-L_z/2}^{L_z/2}\int_{-L_y/2}^{L_y/2}\int_{-L_x/2}^{L_x/2} y\rho_0 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{M}, \\
\overline{\mathbf{x}}\cdot\hat{\mathbf{k}} &= \frac{\int_{-L_z/2}^{L_z/2}\int_{-L_y/2}^{L_y/2}\int_{-L_x/2}^{L_x/2} z\rho_0 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{M}, \\
\overline{\mathbf{x}} &= (0,0,0).
\end{aligned}
$$

The last step is clear because each integral involves are of an odd function (x, y, z) over symmetric bounds (we choose to do the easy–vanishing–integral first). Noticing this makes it fast to evaluate the difference in the variable density case:

$$
\begin{aligned}
M &= \int_{-L_z/2}^{L_z/2}\int_{-L_y/2}^{L_y/2}\int_{-L_x/2}^{L_x/2} (\rho_0 + az) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \rho_0 L_x L_y L_z, \\
\overline{\mathbf{x}}\cdot\hat{\mathbf{i}} &= \frac{\int_{-L_z/2}^{L_z/2}\int_{-L_y/2}^{L_y/2}\int_{-L_x/2}^{L_x/2} x(\rho_0 + az) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{M}, \\
\overline{\mathbf{x}}\cdot\hat{\mathbf{j}} &= \frac{\int_{-L_z/2}^{L_z/2}\int_{-L_y/2}^{L_y/2}\int_{-L_x/2}^{L_x/2} y(\rho_0 + az) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{M}, \\
\overline{\mathbf{x}}\cdot\hat{\mathbf{k}} &= \frac{\int_{-L_z/2}^{L_z/2}\int_{-L_y/2}^{L_y/2}\int_{-L_x/2}^{L_x/2} \rho_0 z + az^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{M}, \\
\overline{\mathbf{x}} &= (0,0,\frac{a}{12M}L_z^3) = (0,0,\frac{a}{12\rho_0}\frac{L_z^2}{L_x L_y}).
\end{aligned}
$$

Cylinder (keeping center of mass in x, y, z coordinates):

$$
\begin{aligned}
M &= \int_{-h/2}^{h/2}\int_0^{2\pi}\int_0^R \rho_0 r \,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}z = \rho_0\pi R^2 h,\\
\overline{\mathbf{x}}\cdot\hat{\mathbf{i}} &= \frac{\int_{-h/2}^{h/2}\int_0^{2\pi}\int_0^R \rho_0 r^2\cos\phi\,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}z}{M},\\
\overline{\mathbf{x}}\cdot\hat{\mathbf{j}} &= \frac{\int_{-h/2}^{h/2}\int_0^{2\pi}\int_0^R \rho_0 r^2\sin\phi\,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}z}{M},\\
\overline{\mathbf{x}}\cdot\hat{\mathbf{k}} &= \frac{\int_{-h/2}^{h/2}\int_0^{2\pi}\int_0^R \rho_0 z r\,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}z}{M},\\
\overline{\mathbf{x}} &= (0,0,0).
\end{aligned}
$$

Now for variable density, choosing coordinates so that $z_0 = 0$:

$$
\begin{aligned}
M &= \int_{-h/2}^{h/2}\int_0^{2\pi}\int_0^R (\rho_0 + az) r \,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}z = \rho_0\pi R^2 h,\\
\overline{\mathbf{x}}\cdot\hat{\mathbf{i}} &= \frac{\int_{-h/2}^{h/2}\int_0^{2\pi}\int_0^R (\rho_0 + az) r^2\cos\phi\,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}z}{M},\\
\overline{\mathbf{x}}\cdot\hat{\mathbf{j}} &= \frac{\int_{-h/2}^{h/2}\int_0^{2\pi}\int_0^R (\rho_0 + az) r^2\sin\phi\,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}z}{M},\\
\overline{\mathbf{x}}\cdot\hat{\mathbf{k}} &= \frac{\int_{-h/2}^{h/2}\int_0^{2\pi}\int_0^R (\rho_0 + az) z r\,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}z}{M},\\
\overline{\mathbf{x}} &= (0,0,\frac{a\pi h^3 R^2}{12M}) = (0,0,\frac{ah^2}{12\rho_0})
\end{aligned}
$$

Sphere (keeping center of mass in x, y, z coordinates):

$$
\begin{aligned}
M &= \int_0^{2\pi}\int_0^{\pi}\int_0^R \rho r^2\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi = \rho\frac{4}{3}\pi R^2,\\
\overline{\mathbf{x}}\cdot\hat{\mathbf{i}} &= \frac{\int_0^{2\pi}\int_0^{\pi}\int_0^R \rho r^3\cos\phi\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi}{M},\\
\overline{\mathbf{x}}\cdot\hat{\mathbf{j}} &= \frac{\int_0^{2\pi}\int_0^{\pi}\int_0^R \rho r^3\sin\phi\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi}{M},\\
\overline{\mathbf{x}}\cdot\hat{\mathbf{k}} &= \frac{\int_0^{2\pi}\int_0^{\pi}\int_0^R \rho r^3\cos\theta\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi}{M},\\
\rho &= (\rho + ar\cos\theta).
\end{aligned}
$$

We note that all of these center of mass integrals would vanish if ρ were a constant ρ_0, but if it varies, then the vertical (z) direction has a center of mass not located at the origin.

6.5 Homework Problems

6.5.1 Manipulation

Exercise 6.1 (Volumes) *Set up and evaluate integrals to calculate the volume of an $L_x \times L_y \times L_z$ rectangular solid, a cylinder of radius R and height h, and a sphere of radius R.*

6.5.2 Application

Exercise 6.2 (Earth) *Redo the cylinder and sphere volume calculations from exercise 6.1, but with all integrals and integrands expressed* in earth coordinates. *Hint: It is easiest to consider the cylinder as sitting with its base on the origin (rather than centered on the origin). Then break up the integral into two parts. First, there is the conic section that extends from $\vartheta = \tan^{-1}(h/R)$ to $\vartheta = \pi/2$ and is bounded at the surface of the top of the cylinder. This surface can be described by $\mathfrak{z} = h/\sin\vartheta$. The other surface to consider is the outer shell of the cylinder. This surface can be described by the function $\mathfrak{z} = R/\cos\vartheta$, and it is relevant for $\vartheta = 0$ to $\vartheta = \tan^{-1}(h/R)$.*

6.5.3 Scheming Schematics and Articulate Analysis

Exercise 6.3 *Problem 5.4.25 Boas (2006). The volume inside a sphere of radius r is $V = \frac{4}{3}\pi r^3$. Then $\mathrm{dV} = 4\pi\mathrm{r}^2\mathrm{dr} = \mathrm{Adr}$, where A is the area of the sphere. What is the geometrical meaning of the fact that the derivative of the volume is the area? Could you use this fact to find the volume formula given the area formula?*

Chapter 7

Fourier Analysis

In Chapter 1, the Taylor series was used to expand functions into a series of polynomials. The formula is:

$$f(x) = \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!} f^{(n)}(a). \tag{7.1}$$

Thinking of functions in this way is useful, as each term in the polynomial has different behavior. When x is near to a, the terms of lower n are usually more important, while the larger n terms become more important as x moves farther away from a. Also, when a function is odd or even, only odd or even polynomial powers appear in the series. Examining the behavior of each term, or collections of terms, in the series leads to a deeper understanding of the function. Experiments may be designed specifically to determine only a few of the terms in the series for a function that is unknown but describes a measurable quantity of interest.

For a finite length series, we can think of the list of coefficients as a vector, the list of polynomial powers as another vector $\mathbf{p}$, and we can add and subtract functions using vector arithmetic:

$$f(x) \approx \mathbf{c} \cdot (1, (x-a), (x-a)^2, \ldots, (x-a)^N) = \mathbf{c} \cdot \mathbf{p}, \tag{7.2}$$

$$g(x) \approx \mathbf{d} \cdot \mathbf{p}, \tag{7.3}$$

$$\alpha f(x) + \beta g(x) \approx (\alpha \mathbf{c} + \beta \mathbf{d}) \cdot \mathbf{p} \tag{7.4}$$

Each polynomial power acts like a coordinate axis in this vector system, and the different polynomial powers do not mix under linear operations on the functions, so they form an orthogonal coordinate basis. We can therefore think of a function as a*superposition* of the different polynomial powers, with the coefficient vector determining how much of each power goes into the mix. We can add together different superpositions describing different functions using linear algebra rules, and the superposition of superpositions will be the sum of the functions.

The Fourier series expansion is similar in many ways, except instead of expanding into polynomial powers, the expansion is into sines, cosines, or complex exponentials that oscillate more and more rapidly as more and more terms are retained in the series. The list of coefficients can be treated as a vector, and there are even close relationships between the length of these coefficient vectors and the variability of the approximated function.

7.1 Introduction–Scales

Reading: Boas (2006)[7.1-7.3].

When observations or experiments are conducted, often the results are lists of measurements separated in time or space. The collection of these results is intended to *sample* a phenomenon sufficiently often and spanning over a total interval of time or space as to categorize its variability. Fourier analysis is particularly well-suited to analyze such data.

The underlying idea of Fourier analysis is that oscillations or repeated patterns exist within the data. These oscillations or repeated patterns are common in nature, vibrations and waves are typical in many circumstances. Even when the dataset does not obviously contain such patterns, it may be thought to be composed of a *superposition of repeating oscillations over a variety of different frequencies*. Just as the Taylor series approximates a function by a superpositions of powers of $(x-a)$, which works better and better the more powers are considered, the Fourier series approximations a function by a superposition of oscillations, and it works better and better the more frequencies of oscillations are considered. The *band* or range of frequencies that go into the Fourier analysis determines the scales of oscillations that are being measured accurately.

Consider first a collection of data at a fixed location but over a span of time. A sediment core or an oceanographic mooring are geophysical examples. Suppose that the samples are taken at regular intervals, T_s, so that we can consider a sampling frequency of $f_s = 1/T_s$. The fastest variability we could hope to measure with this strategy would be a signal that went from a value greater than average on one measurement to a value below average on the next. Thus, the period of the oscillation, that is, the time from a greater than average value to the next greater than average value would be $2T_s$, and thus the highest frequency of oscillation that can be measured is $0.5f_s = 1/(2T_s)$. In Fourier analysis, this frequency is called the Nyquist frequency. On the other hand, all observations must come to an end, so there is a total duration over which the observations occur. If N observations are taken, then this period is $T_r = NT_s$, and thus the minimum frequency that can be measured is $f_r = 1/T_r = 1/(NT_s)$.

Fourier series can be used to examine variability on all frequencies from f_r to f_n, usually broken down into intermediate frequencies at intervals of f_r. Variability over longer or shorter periods than those in the measured band may affect the Fourier analysis, and this contamination is called *aliasing*.

7.2 Expansions

Reading: Boas (2006)[7.4-7.5, 7.7-7.8].

The Taylor series was constructed by differentiation of the function to be approximated and the generic polynomial that was to become the Taylor series. After differentiation, evaluating the result at $x = a$ made every term in the series vanish except one. By this construction method, each coefficient in the series can be found one at a time.

The Fourier series in constructed in a similar fashion, by using "Fourier's Trick." The trick is built

on the following averages over a period:

$$\langle 1 \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} 1 \, \mathrm{d}x = 1, \tag{7.5}$$

$$\langle \cos nx \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos nx \, \mathrm{d}x = 0, \tag{7.6}$$

$$\langle \sin mx \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin mx \, \mathrm{d}x = 0, \tag{7.7}$$

$$\langle \cos^2 nx + \sin^2 nx \rangle = \langle 1 \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} 1 \, \mathrm{d}x = 1, \tag{7.8}$$

$$\langle \cos^2 nx \rangle = \frac{1}{4\pi} \int_{-\pi}^{\pi} \cos^2 nx \, \mathrm{d}x = \frac{1}{2}, \text{ unless } n = 0 \text{ when it is } = 1, \tag{7.9}$$

$$\langle \sin^2 nx \rangle = \frac{1}{4\pi} \int_{-\pi}^{\pi} \sin^2 nx \, \mathrm{d}x = \frac{1}{2}, \text{ unless } n = 0 \text{ when it is } = 0. \tag{7.10}$$

Note that the last two results can be derived from the second one by noting that since sine and cosine differ only in phase over this interval, each must contribute equally to the sum in $\langle \cos^2 nx + \sin^2 nx \rangle$, unless $n = 0$ when sine vanishes everywhere. These averages, along with some less familiar extensions, can be summarized

$$\langle \sin mx \cos nx \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin mx \cos nx \, \mathrm{d}x = 0. \tag{7.11}$$

$$\langle \sin mx \sin nx \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin mx \sin nx \, \mathrm{d}x = \begin{cases} 0, & \text{if } m = n = 0, \\ 0, & \text{if } m \neq n, \\ \frac{1}{2}, & \text{if } m = n \neq 0. \end{cases} \tag{7.12}$$

$$\langle \cos mx \cos nx \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos mx \cos nx \, \mathrm{d}x = \begin{cases} 1, & \text{if } m = n = 0, \\ 0, & \text{if } m \neq n, \\ \frac{1}{2}, & \text{if } m = n \neq 0. \end{cases} \tag{7.13}$$

Each of these averages is easily found by manipulation of the sines and cosines angle addition rules, integration of sines and cosines, and (7.5-7.10). The angle addition rules are

$$\sin(a + b) = \sin a \cos b + \sin b \cos a, \tag{7.14}$$

$$\cos(a + b) = \cos a \cos b - \sin b \sin a. \tag{7.15}$$

The complex exponential equivalent averages also come in handy.

$$\langle e^{imx} e^{inx} \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{imx+inx} \, \mathrm{d}x = \begin{cases} 1, & \text{if } m = n = 0, \\ 0, & \text{if } m \neq n, \\ 1, & \text{if } m = -n \neq 0. \end{cases} \tag{7.16}$$

These are easily derived using Euler's formula and (11.40-11.42) and (11.45), or (11.45) can be used to show (11.40-11.42).

Fourier's Trick is to convert any function that is periodic on $-\pi$ to π to a Fourier series by examining the average value of that function times a sine or cosine. This trick will isolate one coefficient in the Fourier expansion series. For example, if we assume a Fourier series of the form (the $\frac{1}{2}$ is for convenience later):

$$f(x) = \frac{1}{2} a_0 + a_1 \cos x + a_2 \cos 2x + a_3 \cos 3x + \cdots + b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \ldots \tag{7.17}$$

We know that our function $f(x)$ might be odd or even or a combination of an odd and an even part ($f(x) = f_o(x) + f_e(x)$, where $f(-x) = -f_o(x) + f_e(x)$). Thus, we need both the sine series for the odd part (which is $f_o(x) = [f(x) - f(-x)]/2$) and the cosine series for the even part (which is $f_e(x) = [f(x) + f(-x)]/2$). Then we find the following averages to hold by using the rules above.

$$\langle f(x)\rangle = \frac{1}{2}a_0, \tag{7.18}$$

$$\langle f(x)\cos x\rangle = \frac{1}{2}a_1, \tag{7.19}$$

$$\langle f(x)\sin x\rangle = \frac{1}{2}b_1, \tag{7.20}$$

$$\langle f(x)\cos nx\rangle = \frac{1}{2}a_n, \tag{7.21}$$

$$\langle f(x)\sin nx\rangle = \frac{1}{2}b_n. \tag{7.22}$$

We can construct any sine and cosine expansion over $-\pi$ to π by this method.

The complex exponential Fourier series is even simpler because the factors of $\frac{1}{2}$ are not needed.

$$f(x) = c_0 + c_1 e^{ix} + c_2 e^{2ix} + c_3 e^{3ix} + \ldots, \tag{7.23}$$

$$\langle f(x)e^{-inx}\rangle = c_n. \tag{7.24}$$

Note, however, that the exponent of the multiplier is $-inx$ while the exponent of the selected term in the series is $+inx$.

7.2.1 Other Intervals

It is not necessary for $f(x)$ to be periodic on $-\pi \le x \le \pi$, it could be periodic on any other interval $-\frac{L}{2} \le x \le \frac{L}{2}$ and we would just need to rescale using a coefficient of x. For example, if $y = kx$, and $k = \frac{2pi}{L}$, then $f(y)$ will be periodic on $-\pi \le x \le \pi$ if $f(x)$ is periodic on $-\frac{L}{2} \le x \le \frac{L}{2}$. The Fourier expansion will then have arguments of $y, 2y, 3y, \ldots$, which is equivalent to $kx, 2kx, 3kx, \ldots$

7.2.2 Fourier in Time

When a time series is expanded into its Fourier components, the coefficients of time in the argument of the sines and cosines are called frequencies. They are *angular frequencies*, and generally denoted ω or σ, if the sine and cosine arguments are periodic every $\omega t = 2\pi$. For example, $\sin\omega t = 0$ and is increasing with t whenever $\omega t = 0, 2\pi, 4\pi, \ldots$ We can think of $\omega = \frac{2\pi}{T}$ as playing the same role as the k factor in Section 7.2.1, where T is the period of the oscillation with frequency ω.

Sometimes, we prefer to drop the 2π in the definition of the frequency. These frequencies are called *regular frequencies* and are usually denoted f. The sines and cosines will be periodic at times when $ft = 1, 2, 3, \ldots$ For example, $\sin\omega t = \sin(2\pi ft) = 0$ and is increasing every time ft is an integer. We can convert between angular and regular frequencies using $2\pi f = \omega$.

High frequencies correspond to rapid oscillations and short periods. Low frequencies correspond to slow oscillations and long periods.

An Example: Milankovitch Forcing

Milankovitch (1930) performed detailed calculations projecting the variability of the orbit of the Earth. He rationalized that these variations might have an impact on climate. The three primary signals he identified–precession, ellipticity, and obliquity–differ in magnitude and period.

More modern analyses, e.g., (Hays et al., 1976) and (Imbrie and Imbrie, 1980), attack the problem using Fourier analysis. By this method, the total forcing and paleorecord of ice ages–as recorded in isotopic signatures–are decomposed by period and then analyzed. Fig. 7.1 is an example of one such analysis.

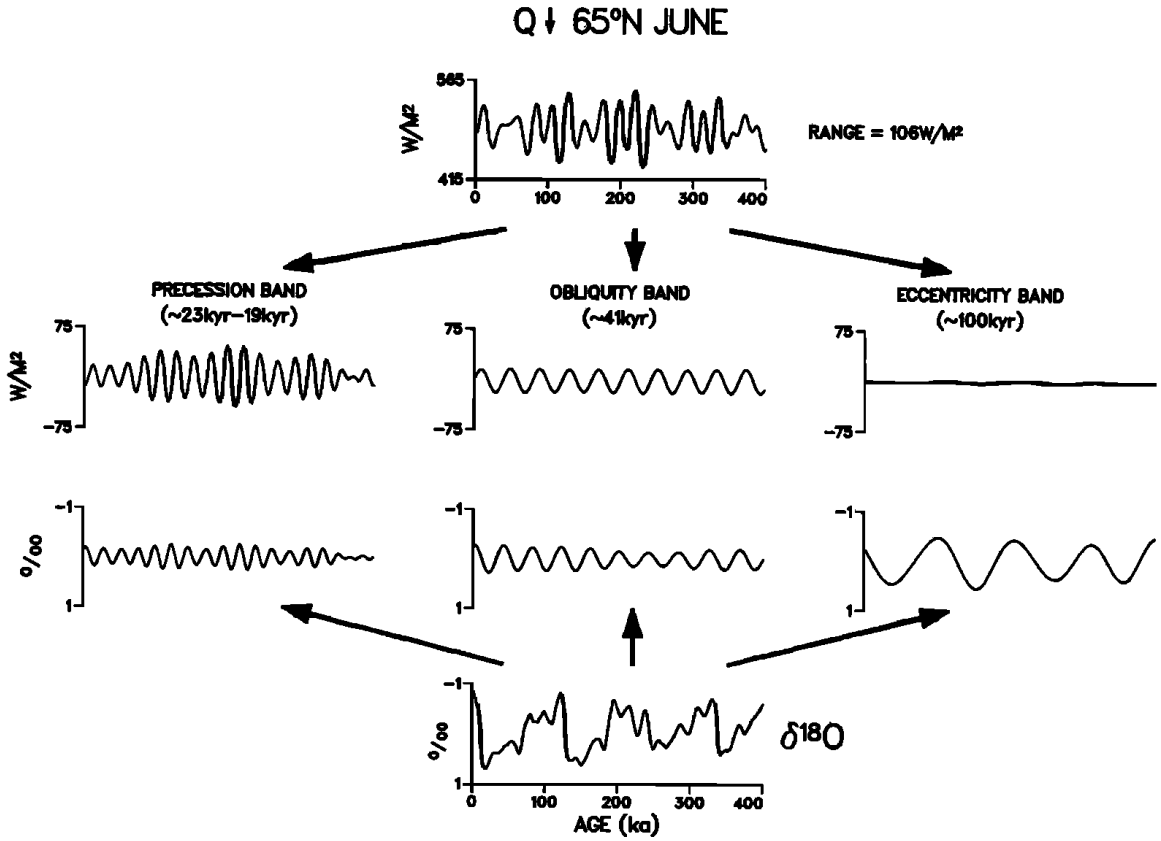


Fig. 1. The 100-kyr cycle problem as seen by partitioning radiation and climate time series into their dominant periodic components (in the precession, obliquity, and 100-kyr eccentricity bands). Radiation time series are from Berger [1978*a*]; $\delta^{18}O$ data are from Imbrie et al. [1984]. Partitioning is done using Hamming band-pass filters with a bandwidth of 0.019 kyr^{-1} for the 41- and 100-kyr bands and 0.036 kyr^{-1} for the 23-kyr band [Jenkins and Watts, 1968]. The $\delta^{18}O$ cycles at periods near 23, 41, and 100 kyr are so strongly correlated with astronomically driven radiation cycles as to suggest a causal linkage in all three bands. But these correlations for the 23-, 41-, and 100-kyr bands (coherencies of 0.95, 0.90, and 0.91, respectively, in Table 2) hide an intriguing physical problem. Why is the system's response so strong in the 100-kyr band? There the amplitude of the radiation signal (2 W m^{-2}) is 1 order of magnitude smaller than in the other two bands.

Figure 7.1: An example of Fourier decomposition of a signal–in this case the radiation variability of orbital motions and the isotope variability of glaciation from Imbrie et al. (1993).

By decomposing the paleorecord in different locations by period, the effects of obliquity, precession

and ellipticity can be individually assessed.

7.2.3 Fourier in Space

When a set of spatial observations are expanded into Fourier components, the coefficients of the spatial coordinate in the argument of the sines and cosines are called wavenumbers. Wavenumbers play the same role as angular frequencies, and are usually defined as $k = \frac{2\pi}{\lambda}$ where λ is called the *wavelength*. The sines and cosines are periodic each time kx is a multiple of 2π.

High wave numbers correspond to rapid oscillations and short wavelengths. Low wave numbers correspond to slow oscillations and long wavelengths.

An Example: The Spectrum of Turbulence

Taylor (1938) made a study of the power spectrum of turbulence. That is, he examined the amount of variability in the Fourier transform of velocities in a turbulent flow, as a function of spatial scale and temporal frequency. He hypothesized (Taylor's Hypothesis) that if the turbulence were carried along by a flow that was much faster than the rate at which turbulence evolved, then the power spectrum of the frequency in time could be related to the power spectrum of the frequency in space. That is, you could either sit still and wait for the turbulent eddies to be advected by you, or you could examine multiple locations along the flow where eddies existed at the same time. If the turbulence were "frozen in" to the flow, then the two would be interchangeable. He examined experimental evidence to support his hypothesis.

This conception inspired Kolmogorov (1941) to examine the spatial power spectrum of turbulence, and to predict that the kinetic energy power spectrum decays with spatial scale (wavenumber k, which is 2π over wavelength) according to $E(k) \propto k^{-5/3}$. This paper is likely the most important paper on turbulence so far.

7.2.4 Discrete and Continuous

Reading: Boas (2006)[7.11-7.12].

7.3 Derivatives of Expansions

Reading: Boas (2006)[7.9].

7.3.1 Application in Differential Equations

One very powerful aspect of the Fourier transform is that it can be used to convert differential equations to algebraic equations, where every derivative is replaced by the frequency or wavenumber being considered. Thus, the differential equation can be solved one frequency at a time, and then the results can be recombined into the total solution by superposition. We will return to this method when as we learn to solve differential equations.

7.4 Homework Problems

7.4.1 Scheming Schematics and Articulate Analysis

Exercise 7.1 (Square Wave) *Problem 7.5.1 of Boas (2006).*

Exercise 7.2 (Sines, Cosines, Exponentials) *Problem 7.5.12 of Boas (2006). Show that in (5.2) the average values of* $\sin(mx)\sin(nx)$ *and of* $\cos(mx)\cos(nx)$, $m \neq n$, *are zero (over a period), by using the complex exponential forms for the sines and cosines as in (5.3).*

7.4.2 Evaluate & Create

Exercise 7.3 (Derivatives) *a) Show that the following function* $f(x)$ *and Fourier series* $g(x)$ *are equivalent on the interval from* $-\pi$ *to* π *up to order of* $\sin(2x)$. *To do so, multiply the* $f(x)$ *and* $g(x)$ *functions by each of the following in turn:* $\sin(x), \sin(2x)$ *and* $\cos(0x), \cos(x), \cos(2x)$. *Show that the average value of the product from* $-\pi$ *to* π *is the same, for example that* $\langle f(x)\sin(2x)\rangle = \langle g(x)\sin(2x)\rangle$. *(see Boas, 2006, pg. 351).*

$$\begin{aligned} \forall : -\pi \leq x \leq \pi, \\ f(x) &= x(\pi - x)(\pi + x), \\ g(x) &= \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^3} 12 \sin(nx). \end{aligned}$$

b) Take the first derivative of $f(x)$ *and* $g(x)$ *(by taking the derivative of the generic term in the series). Show that the resulting derivatives are equivalent, using the same method as in a).*

Chapter 8

Vector & Tensor Analysis

Reading: Boas (2006, 6.1)

Vector analysis is at the heart of all geophysical analyses, models, and theories. It is through the rules of vector analysis that we can quickly assess whether a theory is correctly framed to apply regardless of coordinates, and it is the language we use to express the principles of conservation of energy, momentum, mass, and other fundamental concepts of continuum mechanics. It avoids many of the detailed trigonometry calculations of three-dimensional geometry while retaining the capability to describe and analyze data in that geometry.

8.1 Introduction–Products, Derivatives, and Vector Fields

We have already learned how to multiply a vector by a scalar and how to sum vectors. We have even learned to solve systems of equations involving these operations, as well as the more powerful (and obscure) matrix multiplication operation.

Vector analysis adds new operations to the list, which are particularly useful in continuum mechanics–especially in three dimensions. The new operations are a vector product of two vectors, or cross product, and derivatives of vectors. The derivatives of vectors that most interest us are those that provide insight beyond a particular choice of coordinates, and they are the gradient, the divergence, and the curl. These operations are full of symmetries, which make them powerful, but also complicate our job of learning to use them. Thus, we will have to learn intuition for what they represent, but also how they interact with one another (these rules are called triple product and calculus identities).

8.2 Fields–Scalars and Vectors

Reading: Boas (2006, 6.2, 6.5)

Whatever we plan to take the spatial derivative of needs to be defined at many points in space. A *field* is such a variable. A field can be a scalar field (a scalar at every point, such as temperature,

pressure, or chemical concentration), a vector field (a vector at every point, such as velocity, acceleration, momentum, gravitational force), or even a tensor field (a tensor at every point, such as stress tensor, anisotropic viscosity, etc.) which we will discuss in later chapters.

Many physical phenomena are expressed using fields, resulting in field theories to encapsulate the phenomena. Continuum mechanics, which describes the motions of solids and fluids, is a field theory. On very small scales, these materials are made up of atoms and molecules, with space in between, and then the description of their properties using fields breaks down. However, continuum mechanics prohibits consideration of these small scales, because (as the name suggests) the materials considered are assumed to be continual in space and time on the scales of interest.

We have discussed how all vectors have only a magnitude and a direction, so how do we describe a field? There are two parts to the field, the *position* which is a vector pointing from the origin to a given location, and the *value* which may be a scalar or a vector. Thus, the field is a function of position, and often also time. When a field is vector valued, the value continues to only have a magnitude and direction. Generally, there must be a good reason to mix up the position with the value, so generally we compare neighboring values with each other, not values and positions.

8.3 If It Quacks Like a Duck...

We are going to begin considering how vectors change under change of coordinates (rotation, reflection in a mirror, more general changes), as well as how vectors are differentiated. We have been considering vectors as objects in linear algebra, or linear systems of equations, that have a direction and magnitude and combine according to the rules of linear operations. However, an alternative, and nearly equivalent way to define vectors is by their behavior under coordinate transformations.

What do I mean by "behavior under coordinate transformations"? Well, lets consider a scalar, say temperature. When we redefine a new coordinate system, the temperature at a point does not change, but the position vector describing where this temperature is located may change. When we consider a vector field, *both the position and value* will change under a change of coordinates. Keeping track of such changes is a big part of what we will do, and finding operations, such as differentiation operators, that respect the way that vectors transform will be key to expressing our physical laws in a way that is not arbitrarily dependent on a choice of coordinates.

Taking this idea to the extreme, it is possible to define a vector by what it does under rotation or other change of coordinates, and then use this definition to prove that such an object obeys the linear relations that we have taken to define vectors. This exercise is a diversion in pure math from our perspective, but the take-away message is that understanding how vectors relate to coordinates is fundamental to understanding vector fields.

We learned in the linear algebra chapter of Boas (2006) that any rotation of a coordinate system or vector can be expressed by multiplication by a matrix, $\mathbf{R}$, which can be composed of cosines and sines, such as a rotation around the x axis:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}. \tag{8.1}$$

More generally, any orthonormal matrix (rows and columns are orthogonal and of magnitude 1) can be considered to be a combination of rotation (although it is sometimes hard to figure out what the rotation axis is) and possibly also a mirror reflection. These matrices have their transpose equal to their inverse, $\mathbf{R}^\intercal = \mathbf{R}^{-1}$. A combination of a rotation and a reflection has a determinant of -1, whereas a pure rotation has a determinant of $+1$.

Two important examples of orthonormal matrices are the generic rotation matrix by angle θ about a unit vector $\mathbf{n}$, and the reflection matrix along the unit vector $\mathbf{n}$ direction:

$$R_{ij}(n,\theta) = \cos\theta\delta_{ij} + \sin\theta\epsilon_{ijk}n_j + (1-\cos\theta)n_i n_j, \tag{8.2}$$

$$\Re_{ij}(n) = \delta_{ij} - 2n_i n_j. \tag{8.3}$$

Note that the trace of the rotation matrix gives an easy way to calculate the angle: $Tr(R_{ij}) = 1 + 2\cos\theta$.

The rotated, reflected version of a vector $\mathbf{v}$ is $\mathbf{Rv}$. We can consider the length of a vector before rotation as a matrix product of a row vector (its transpose) and a column vector, or $\mathbf{v}^\intercal\mathbf{v} = |v|^2$. Recalling that transposing a matrix product reverses the order of the factors, we see that

$$(\mathbf{Rv})^\intercal\mathbf{Rv} = \mathbf{v}^\intercal\mathbf{R}^\intercal\mathbf{Rv} \tag{8.4}$$

$$= \mathbf{v}^\intercal\mathbf{R}^{-1}\mathbf{Rv} \tag{8.5}$$

$$= \mathbf{v}^\intercal\mathbf{v} \tag{8.6}$$

$$|v|^2. \tag{8.7}$$

Thus, rotation and reflection do not change the length of a vector. Or, more generally, any scalar product $\mathbf{A}^\intercal\mathbf{B}$ will be unaffected by a rotation or reflection of the coordinate system.

Matrices can also be rotated, based on reproducing the correct result as in the unrotated case. Note

$$\mathbf{RMR}^\intercal\mathbf{Rv} = \mathbf{R}(\mathbf{Mv}). \tag{8.8}$$

That is, it takes two rotation matrices to rotate a matrix, and the rotated form is $\mathbf{RMR}^{-1}$. [1] It takes one matrix to rotate a vector and none to rotate a scalar.

8.3.1 Vector Covariance

I might prefer a different coordinate system than you do, but physics doesn't choose a favorite between you and me. Thus, we cannot write equations that are subject to a choice of coordinates. Scalar fields are independent of coordinate system, thus it is acceptable to write: scalar function=other scalar function. Likewise once we have chosen a coordinate system, we know how to express vectors and their components. So, we can write: vector function=other vector function. These two examples are *vector covariant*, that is, when one side changes due to a change of coordinates so does the other. We cannot meaningfully write scalar function=vector function, because when we change the choice of coordinates, one side will change while the other will not.

Obviously, we cannot cheat by moving terms from side to side, so the two addends must also be vector covariant. Factors in products need not be the same vector type, which leads to a multiplicity of vector products that will be discussed next.

[1] The resemblance of this rotation formula to the diagonalization by eigenvectors formula (3.22) is not a coincidence! Diagonalization involves rotation of the coordinates and perhaps also stretching or squeezing in some directions and maybe complex numbers in eigenvalues or eigenvectors.

8.4 Multiplying Vectors

Reading: Boas (2006, 6.3)

We also learned that one can multiply a vector times a scalar, which just changes the magnitude of the vector but not the direction. This is one vector covariant form of vector multiplication,

We have also struck on the idea of the inner product, a.k.a. the dot product: $\mathbf{A}\cdot\mathbf{B} = |\mathbf{A}||\mathbf{B}|\cos\theta = A_xB_x + A_yB_y + A_zB_z = \mathbf{A}^\intercal\mathbf{B}$. This product compares two vectors and the result is a scalar. The special case of a vector dotted with itself give the magnitude of the vector squared (in fact, this result is often taken to define magnitude).

Another ways to multiply vectors is to take the vector product, a.k.a. the cross product. For example, in three dimensions, it takes two vectors and results in a third vector, perpendicular to the first two and with magnitude $|\mathbf{A}\times\mathbf{B}| = |\mathbf{A}||\mathbf{B}|\sin\theta$, which makes it seem like a natural complement to the dot product. However, the vector that results from a cross product also has a direction, which is taken to be perpendicular to both of its arguments (i.e., it lies in perpendicular to the plane spanned by the input vectors),

The cross product is actually a bit of an oddity, as we'll see when we express it as a tensor operation. Its result is actually not a vector, but a pseudo vector, which means that it does not behave like a vector when its image in a mirror is considered (vectors reverse under mirror reflection, pseudovectors do not). Furthermore, in two dimensions, the cross product results in a vector that is *not contained in the two dimensions of the space*! Generally, in two dimensions, the cross product is therefore considered not as a vector times a vector equals a vector. In four dimensions, the cross product is even messier.

Oddity though it may be, the cross product is very useful in three dimensions. It is fundamental to the understanding of the rotation of the Earth, the rotation of a fluid element (vorticity), angular momentum, torque, and non-conservative forces. It even has its own fundamental theorem of calculus, called Stokes' Theorem.

8.4.1 Application: Casting a Shadow

One very useful application of the dot product is in calculating the amount of energy delivered to a surface by incoming and outgoing radiation, such as the Sun shining on the Earth. When we hold up a disc perpendicular to the Sun's rays, all of the photons streaming toward it are absorbed (heating the disc) or reflected. However, as we tilt the disc, it only blocks an elliptical shape of the Sun's light. Thus, the orientation of the object matters (hinting at a vector), even though energy is a scalar. We can use this method to determine the intensity of the Sun.

The "solar constant" is the measure of this incoming radiation, and it is about $1361\,\mathrm{W\,m^{-2}}$. That is, the amount of energy leaving the Sun is best measured in energy per unit time per square meter. The reason for the "per square meter" is because the Sun's energy fills out into larger and larger space with distance from the Sun. The easy way to imagine this is as imaginary concentric spherical shells centered on the Sun. If the Sun's energy production is constant in time, then the same amount of energy passes through each shell per unit time. But, as the shells farther out have more area than the shells closer in (by a factor of radius ratio squared), the power is constant per unit area, rather than being constant at all points in space. This result is the "inverse square law."

But even when we consider the inverse square law, the average radiation arriving at the Earth's surface is *not* the solar constant per square meter of Earth's surface. Consider the shadow cast by the sphere (Area: $4\pi r^2$) of the Earth–it is a circle of Area πr^2, since only the front side of the Earth is exposed and most of that surface is at an angle to the Sun, just as the tilted disc just described. Thus, we integrate $\mathbf{I}\cdot\hat{\mathbf{n}}$ instead of $|\mathbf{I}|$, where $\mathbf{I}$ is the energy of the sun propagating along a sunbeam, and $\hat{\mathbf{n}}$ is a unit vector normal to Earth's surface.[2] The sine of the angle between the sunbeams and the surface is just what is needed to reduce by a factor of 2, and the fact that only one side of the Earth is illuminated gives the other factor of two.

This problem is an example of a larger set of examples of projection problems. The projection of a vector $\mathbf{A}$ onto $\mathbf{B}$ is $(\mathbf{A}\cdot\mathbf{B})\mathbf{B}$, so it involves the dot product.

8.4.2 Application: Finding an Axis

The velocity of the Earth as it rotates can be written down as $\mathbf{v}_{rot} = \frac{2\pi|r_\perp|}{\text{day}}\hat{\phi}$, where $r_\perp$ is the distance from the rotation axis to the point in question.

A much simpler formula for the rotation is

$$\mathbf{v}_{rot} = \mathbf{\Omega}\times\mathbf{r}. \tag{8.9}$$

$\mathbf{\Omega}$ is a vector aligned with the rotation axis, with a magnitude of the angular frequency of the rotation ($\approx 2\pi/(24\,\text{h}) = 7.272\times10^{-5}\,\text{s}^{-1}$).[3] Thus, the velocity is clearly perpendicular to both the axis of rotation and the position from the origin (taken to lie somewhere on the axis). Note how the timing of the circumference ($2\pi r_\perp$ per day) comes in, and note how the fact that the poles do not have a velocity of rotation occurs.

8.4.3 Triple Products

Sometimes, after multiplying two vectors, you may then multiply by a third one. These "triple products" are a bit complicated to do step by step, so the result is sometimes tabulated or memorized. Here they are for reference:

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}, \tag{8.10}$$

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \mathbf{B}\,(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}\,(\mathbf{A}\cdot\mathbf{B})\,. \tag{8.11}$$

The former gives us a vector expression for a determinant, and the latter is often called the "back cab" rule because of the pattern of letters in this standard way of writing it.

[2] I usually use the $\hat{\cdot}$ symbol to indicate unit vector length, but Boas (2006) does not.

[3] You may wonder why this frequency is only approximate. It is because once per day is actually not the rotation period of the Earth, it is the time that it takes for the Sun to return to noon. As the Earth is moving around the Sun, it takes just a little more than one rotation of the Earth to get the Sun back to noon. The real rotation rate of the Earth is $2\pi/(sidereal day) = 23.934\,469\,6\,\text{h} = 7.292\,12\times10^{-5}\,\text{s}^{-1}$.

8.5 Differentiating Fields

Reading: Boas (2006, 6.5-6.7)

A field has a value at every point in space. It is natural to consider the change in a field from one point to another, just as we use calculus to consider the change in functions from one point to another. It is also natural to expect partial derivatives to play a key role in this analysis. The symmetries of vectors, discussed above, will help choose the particular combinations of partial derivatives that dominate this analysis of vectors.

8.5.1 Differentiating Scalars with the Gradient: Ch-ch-ch-ch-ch-ch-Changes

Reading: Boas (2006, 6.6)

The gradient of a scalar field results in a new vector. The gradient can be thought of as the "del" operation ∇ on the scalar. "Del" can be thought of as,

$$\nabla = \hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z} \tag{8.12}$$

$$= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right), \tag{8.13}$$

$$= \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix} \tag{8.14}$$

The gradient is the first use of ∇. The gradient of a scalar field $f(x, y, z, t)$ is a vector where the components are the partial derivatives of $f(x, y, z, t)$ in each direction:

$$\nabla = \hat{\mathbf{i}} \frac{\partial f(x,y,z,t)}{\partial x} + \hat{\mathbf{j}} \frac{\partial f(x,y,z,t)}{\partial y} + \hat{\mathbf{k}} \frac{\partial f(x,y,z,t)}{\partial z} \tag{8.15}$$

$$= \left(\frac{\partial f(x,y,z,t)}{\partial x}, \frac{\partial f(x,y,z,t)}{\partial y}, \frac{\partial f(x,y,z,t)}{\partial z} \right), \tag{8.16}$$

$$= \begin{bmatrix} \frac{\partial f(x,y,z,t)}{\partial x} \\ \frac{\partial f(x,y,z,t)}{\partial y} \\ \frac{\partial f(x,y,z,t)}{\partial z} \end{bmatrix} \tag{8.17}$$

The gradient of a scalar is a vector, because it rotates and reflects like a vector (how can you tell?), because it has a magnitude like a vector (which is?), and because it undergoes general linear operations like a vector (how can you tell?). I keep the dependence of f on time t, as an example of a representation of other parameters which are not the coordinate directions. *In normal, non-relativistic, physics, the gradient only consists of the partial derivatives with respect to spatial coordinates, holding time and other parameters upon which the scalar field depends constant.*

The gradient of a scalar field points at every location toward the direction where the scalar field increases most rapidly, with a magnitude that corresponds to how rapidly the scalar field changes

per unit distance. The pressure gradient force per unit mass $-\nabla p$ is an excellent example of the gradient operator. It always point from high pressure toward low pressure (note the minus sign), and is greater in magnitude when the pressure change is more rapid in space.

One must be careful in evaluating the gradient in other coordinate systems, such as cylindrical, spherical, or earth coordinates. Note that the coordinates in these systems don't even have the same units, so it is clear that you have to be careful when forming the components of a vector with objects of different units. The gradient in each of these coordinates is

Cylindrical: (8.18)

$$\nabla f(r,\phi,z,t) = \hat{\mathbf{r}}\,\frac{\partial f(r,\phi,z,t)}{\partial r} + \frac{\hat{\phi}}{r}\,\frac{\partial f(r,\phi,z,t)}{\partial \phi} + \hat{\mathbf{z}}\,\frac{\partial f(r,\phi,z,t)}{\partial z}\,,$$

Spherical: (8.19)

$$\nabla f(r,\theta,\phi,t) = \hat{\mathbf{r}}\,\frac{\partial f(r,\theta,\phi,t)}{\partial r} + \frac{\hat{\theta}}{r}\,\frac{\partial f(r,\theta,\phi,t)}{\partial \theta} + \frac{\hat{\phi}}{r\sin\theta}\,\frac{\partial f(r,\theta,\phi,t)}{\partial \phi}\,,$$

Earth: (8.20)

$$\nabla f(\mathfrak{z},\phi,\vartheta,t) = \hat{\mathfrak{z}}\,\frac{\partial f(\mathfrak{z},\phi,\vartheta,t)}{\partial \mathfrak{z}} + \frac{\hat{\phi}}{(\mathfrak{z}+r_0)\cos\vartheta}\,\frac{\partial f(\mathfrak{z},\phi,\vartheta,t)}{\partial \phi} + \frac{\hat{\vartheta}}{(\mathfrak{z}+r_0)}\,\frac{\partial f(\mathfrak{z},\phi,\vartheta,t)}{\partial \vartheta}\,.$$

When excursions in the vertical are small in comparison to the radius of the earth, then the factor in the denominator in earth coordinates simplifies: $(\mathfrak{z}+r_0) \approx r_0$.

In 2-dimensional (or 4D or 10D, etc.) problems, the gradient is interpreted in much the same way, with differentiation in each Cartesian coordinate direction representing a component of the Cartesian vector. In special and general relativity, where time is treated as a fourth dimension, one may include the partial derivative with respect to time as part of the gradient, and then the speed of light is needed to match the units (and meaning) of the spacetime vector that results. Rarely, this approach is used for non-relativistic wave problems where the wave speed is constant.

8.5.2 Differentiating Vectors

Reading: Boas (2006, 6.7)

So far, we have considered only taking the derivative of a scalar field. What about the derivatives of a vector field? There are two main flavors: the divergence and the curl. The divergence takes a derivative of a vector field and results in a scalar. The curl takes the derivative of a vector and results in a (pseudo-)vector. The ∇ notation proves particularly powerful here, since it makes the divergence appear as a dot product and the curl appear as a cross product.

Divergence

The divergence takes a derivative of a vector field and results in a scalar. The ∇ notation proves particularly powerful here, since it makes the divergence appear as a dot product.

$$\begin{aligned}
\text{div}(\mathbf{v}(x,y,z,t)) &= \nabla\cdot\mathbf{v}(x,y,z,t), && (8.21)\\
&= \left(\frac{\partial}{\partial x}\,,\frac{\partial}{\partial y}\,,\frac{\partial}{\partial z}\right)\cdot\big(v_x(x,y,z,t),v_y(x,y,z,t),v_z(x,y,z,t)\big)\,, && (8.22)\\
&= \frac{\partial v_x(x,y,z,t)}{\partial x}+\frac{\partial v_y(x,y,z,t)}{\partial y}+\frac{\partial v_z(x,y,z,t)}{\partial z}\,. && (8.23)
\end{aligned}$$

Like the gradient, the divergence needs care in other non-Cartesian coordinate systems. The divergence in each of our standard systems is (with subscripts indicating vector coordinate values)

Cylindrical: (8.24)

$$\nabla\cdot\mathbf{v}(r,\phi,z,t)=\frac{1}{r}\frac{\partial(rv_r(r,\phi,z,t))}{\partial r}+\frac{1}{r}\frac{\partial v_\phi(r,\phi,z,t)}{\partial\phi}+\frac{\partial v_z(r,\phi,z,t)}{\partial z}\,,$$

Spherical: (8.25)

$$\nabla\cdot\mathbf{v}(r,\theta,\phi,t)=\frac{1}{r^2}\frac{\partial\left(r^2v_r(r,\theta,\phi,t)\right)}{\partial r}+\frac{1}{r\sin\theta}\frac{\partial\left(\sin\theta v_\theta(r,\theta,\phi,t)\right)}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial v_\phi(r,\theta,\phi,t)}{\partial\phi}\,,$$

Earth: (8.26)

$$\begin{aligned}
\nabla\cdot\mathbf{v}(\mathfrak{z},\phi,\vartheta,t) &= \frac{1}{(\mathfrak{z}+r_0)^2}\frac{\partial\left[(\mathfrak{z}+r_0)^2v_\mathfrak{z}(\mathfrak{z},\phi,\vartheta,t)\right]}{\partial\mathfrak{z}}+\frac{1}{(\mathfrak{z}+r_0)\cos\vartheta}\frac{\partial v_\phi(\mathfrak{z},\phi,\vartheta,t)}{\partial\phi}\\
&\quad+\frac{1}{(\mathfrak{z}+r_0)\cos\vartheta}\frac{\partial\left[\cos\vartheta v_\vartheta(\mathfrak{z},\phi,\vartheta,t)\right]}{\partial\vartheta}\,.
\end{aligned}$$

When excursions in the vertical are small in comparison to the radius of the earth, then the factor in the denominator in earth coordinates simplifies: $(\mathfrak{z}+r_0)\approx r_0$. The factor in the numerator inside the derivative, however, should be differentiated before this approximation is made, that is

$$\frac{1}{(\mathfrak{z}+r_0)^2}\frac{\partial\left[(\mathfrak{z}+r_0)^2v_\mathfrak{z}(\mathfrak{z},\phi,\vartheta,t)\right]}{\partial\mathfrak{z}}=\frac{2v_\mathfrak{z}(\mathfrak{z},\phi,\vartheta,t)}{(\mathfrak{z}+r_0)}+\frac{\partial v_\mathfrak{z}(\mathfrak{z},\phi,\vartheta,t)}{\partial\mathfrak{z}}\approx\frac{2}{r_0}v_\mathfrak{z}(\mathfrak{z},\phi,\vartheta,t)+\frac{\partial v_\mathfrak{z}(\mathfrak{z},\phi,\vartheta,t)}{\partial\mathfrak{z}}\,.$$

Curl

The curl takes the derivative of a vector and results in a (pseudo-)vector. The ∇ notation proves particularly powerful here, since it makes the curl appear as a cross product.

$$\begin{aligned}
\text{curl}(\mathbf{v}(x,y,z,t)) &= \nabla\times\mathbf{v}(x,y,z,t), && (8.27)\\
&= \left(\frac{\partial}{\partial x}\,,\frac{\partial}{\partial y}\,,\frac{\partial}{\partial z}\right)\times\big(v_x(x,y,z,t),v_y(x,y,z,t),v_z(x,y,z,t)\big)\,, && (8.28)\\
&= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_x(x,y,z,t) & v_y(x,y,z,t) & v_z(x,y,z,t) \end{vmatrix} && (8.29)\\
&= \hat{\mathbf{i}}\left[\frac{\partial v_z}{\partial y}-\frac{\partial v_y}{\partial z}\right]+\hat{\mathbf{j}}\left[\frac{\partial v_x}{\partial z}-\frac{\partial v_z}{\partial x}\right]+\hat{\mathbf{k}}\left[\frac{\partial v_y}{\partial x}-\frac{\partial v_x}{\partial y}\right]. && (8.30)
\end{aligned}$$

Note that here the position of the vector components has been dropped, but it is important to remember that the vector is still a field, otherwise differentiation with respect to space would be meaningless.

Like the gradient and divergence, the curl needs care in other non-Cartesian coordinate systems. The curl in each of our standard systems is

Cylindrical: (8.31)

$$\nabla \times \mathbf{v}(r,\phi,z,t) = \hat{\mathbf{r}}\left[\frac{1}{r}\frac{\partial v_z}{\partial \phi} - \frac{\partial v_\phi}{\partial z}\right] + \hat{\phi}\left[\frac{\partial v_r}{\partial z} - \frac{\partial v_z}{\partial r}\right] + \frac{\hat{\mathbf{z}}}{r}\left[\frac{\partial (r v_\phi)}{\partial r} - \frac{\partial v_r}{\partial \phi}\right].$$

Spherical: (8.32)

$$\nabla \times \mathbf{v}(r,\theta,\phi,t) = \frac{\hat{\mathbf{r}}}{r\sin\theta}\left[\frac{\partial (\sin\theta v_\phi)}{\partial \theta} - \frac{\partial v_\theta}{\partial \phi}\right] + \frac{\hat{\theta}}{r}\left[\frac{1}{\sin\theta}\frac{\partial v_r}{\partial \phi} - \frac{\partial (r v_\phi)}{\partial r}\right] + \frac{\hat{\phi}}{r}\left[\frac{\partial (r v_\theta)}{\partial r} - \frac{\partial v_r}{\partial \theta}\right].$$

Earth: (8.33)

$$\begin{aligned}\nabla \times \mathbf{v}(\mathfrak{z},\phi,\vartheta,t) =& \frac{\hat{\mathfrak{z}}}{(\mathfrak{z}+r_0)\cos\vartheta}\left[\frac{\partial v_\vartheta}{\partial \phi} - \frac{\partial (\cos\vartheta v_\phi)}{\partial \vartheta}\right] + \frac{\hat{\phi}}{(\mathfrak{z}+r_0)}\left[\frac{\partial v_\mathfrak{z}}{\partial \theta} - \frac{\partial\left[(\mathfrak{z}+r_0)v_\vartheta\right]}{\partial \mathfrak{z}}\right] \\ &+ \frac{\hat{\vartheta}}{(\mathfrak{z}+r_0)}\left[\frac{\partial\left[(\mathfrak{z}+r_0)v_\phi\right]}{\partial \mathfrak{z}} - \frac{1}{\cos\vartheta}\frac{\partial v_\mathfrak{z}}{\partial \phi}\right].\end{aligned}$$

Special care is needed in interpreting the curl in fewer or more dimensions than 3. In two dimensions, the curl results in a (pseudo)scalar instead of a vector. The magnitude of this scalar is just the component of the three dimensional curl in the direction perpendicular to the plane of motion. In four dimensions, the curl is better considered by another method which will be discussed later.

8.5.3 Application: Angular Velocity and Solid Body Rotation

More generally than in section 8.4.2, any solid body rotation can be expressed as $\mathbf{v}_{rot} = \boldsymbol{\Omega} \times (\mathbf{r} - \mathbf{r}_{\text{axis}})$, where r_{axis} is just the constant distance from the origin to any point on the axis and $\boldsymbol{\Omega}$ is the angular frequency aligned with the axis. This formula is quite useful for figuring out if the flow is just rotating or has shear.

The *vorticity* is the curl of a velocity, which is a crucial quantity in fluid dynamics $\omega = \nabla \times \mathbf{v}$. Notice that the vorticity of solid body rotation is (using a triple product rule from pp. 339 of Boas, 2006)

$$\begin{aligned}\omega &= \nabla \times \boldsymbol{\Omega} \times (\mathbf{r} - \mathbf{r}_{\text{axis}}) && (8.34)\\ &= \left([\mathbf{r} - \mathbf{r}_{\text{axis}}] \cdot \nabla\right)\boldsymbol{\Omega} - (\boldsymbol{\Omega} \cdot \nabla)(\mathbf{r} - \mathbf{r}_{\text{axis}}) - (\mathbf{r} - \mathbf{r}_{\text{axis}})(\nabla \cdot \boldsymbol{\Omega}) + \boldsymbol{\Omega}\left(\nabla \cdot [\mathbf{r} - \mathbf{r}_{\text{axis}}]\right) \\ &= -(\boldsymbol{\Omega} \cdot \nabla)\mathbf{r} + \boldsymbol{\Omega}(\nabla \cdot \mathbf{r}) && (8.35)\\ &= -\left(\Omega_x \frac{\partial}{\partial x} + \Omega_y \frac{\partial}{\partial y} + \Omega_z \frac{\partial}{\partial z}\right)\left(x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}\right) + \boldsymbol{\Omega}\left(\frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z}\right) \\ &= -\boldsymbol{\Omega} + 3\boldsymbol{\Omega} \\ &= 2\boldsymbol{\Omega}. && (8.36)\end{aligned}$$

So, the vorticity of solid body rotation is twice the angular frequency vector.

8.5.4 Oddballs and Second Derivatives

Reading: Boas (2006, 6.7)

The gradient of a vector is also a meaningful quantity, but the result is neither a vector nor a scalar, it is a dyad (or a matrix, or a second-rank tensor), with elements columns consisting of the gradient of each component of the vector. We will be in a better position to understand this object later, when we discuss tensor analysis.

You can invent other forms of differentiation of scalars and vectors, but they are unlikely to result in true scalars or vectors or pseudovectors. Thus, it will be very difficult for them to participate in vector covariant relationships.

The Laplacian, ∇^2, is the most important second derivative in vector analysis. It is the divergence of the gradient. There are other second derivatives as well, some of which always vanish. You can also perform the Laplacian of a vector, which results in a vector (i.e., the Laplacian acts on each component of the vector field).

$$\nabla^2 f(x,y,z,t) = \nabla\cdot\nabla f(x,y,z,t) = \left(\frac{\partial^2 f(x,y,z,t)}{\partial x^2}, \frac{\partial^2 f(x,y,z,t)}{\partial y^2}, \frac{\partial^2 f(x,y,z,t)}{\partial z^2}\right) \tag{8.37}$$

The Laplacian has a simple interpretation: the value of the Laplacian at a point measures the deviation of that point from the average value of its surroundings. This interpretation is why the diffusion of temperature (or other scalar tracer) is modeled as $\dfrac{\partial T}{\partial t} = \kappa\nabla^2 T$, a relationship called Fick's Law. Many forms of viscosity are also modeled using the Laplacian.

Like the gradient, divergence, and curl, the Laplacian takes a complicated form in other coordinate systems. Here they are:

Cylindrical: (8.38)

$$\nabla^2 f(r,\phi,z,t) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2},$$

Spherical: (8.39)

$$\nabla^2 f(r,\theta,\phi,t) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \phi^2},$$

Earth: (8.40)

$$\nabla^2 f(\mathfrak{z},\phi,\vartheta,t) = \frac{1}{(\mathfrak{z}+r_0)^2}\frac{\partial}{\partial \mathfrak{z}}\left((\mathfrak{z}+r_0)^2\frac{\partial f}{\partial \mathfrak{z}}\right) + \frac{1}{(\mathfrak{z}+r_0)^2\cos^2\vartheta}\frac{\partial^2 f}{\partial \phi^2} + \frac{1}{(\mathfrak{z}+r_0)^2\cos\vartheta}\frac{\partial}{\partial \vartheta}\left(\cos\vartheta\frac{\partial f}{\partial \vartheta}\right).$$

Other uses of ∇ are important to know, see Table 8.1. Perhaps the most important identities using second derivatives of ∇ are the following, which are true for *any scalar* f *and any vector* $\mathbf{v}$*:*

$$\nabla\cdot(\nabla\times\mathbf{v}) = 0, \tag{8.41}$$

$$\nabla\times(\nabla f) = 0. \tag{8.42}$$

The following few sections will exploit these identities.

8.5.5 Application: Vorticity and Divergence

Typical motions of a fluid are continually churning at all locations. The equations to govern this motion are complex, as we will see when we discuss them in the partial differential equations chapter. Thus, it would be nice to simplify the motion by breaking it apart into different classes of meaningful motion.

The divergent velocity is one type of meaningful motion. The equation for the conservation of mass in a compressible fluid of density ρ and velocity $\mathbf{u}$ is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0. \tag{8.43}$$

This equation states that if the density is to be constant at every point in space ($\frac{\partial \rho}{\partial t}$), then the divergence of the flux of density $\rho\mathbf{u}$ must vanish.

One interesting example is that of a constant density fluid. In this case, the conservation of mass becomes

$$\nabla \cdot \mathbf{u} = 0. \tag{8.44}$$

The divergence of the velocity is zero. Thus, if there is to be no increase or decrease in density, the velocity is constrained to not "pile up" anywhere. Divergent (or convergent) velocity fields tend to change the density.

Another kind of meaningful flow is vortical flow. The *vorticity*, ω of a fluid with velocity $\mathbf{u}$ is

$$\omega = \nabla \times \mathbf{u}. \tag{8.45}$$

The vorticity describes how much the fluid is "swirling" about a given location (hence the fact that solid body rotation has vorticity, as noted above). Does vortical flow lead to convergences? Does convergent flow lead to vorticity? That is the topic of the next subsection.

8.5.6 Streamfunction and Potential: Helmholtz Decomposition

Any smooth vector in an infinite domain can be written as the sum of two parts, $\mathbf{u} = \mathbf{u}_{\text{rot}} + \mathbf{u}_{\text{div}}$. The rotational part $\mathbf{u}_{\text{rot}}$ may be generated by a vector streamfunction $\mathbf{\Psi}$: $\mathbf{u}_{\text{rot}} = \nabla \times \mathbf{\Psi}$). The divergent part $\mathbf{u}_{\text{div}}$ can be generated using a scalar potential ϕ: $\mathbf{u}_{\text{div}} = -\nabla\phi$.

Note that the ∇ identities guarantee that

$$\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{u}_{\text{rot}} + \nabla \cdot \mathbf{u}_{\text{div}} = \nabla \cdot \mathbf{u}_{\text{div}} = -\nabla^2\phi, \tag{8.46}$$

$$\nabla \times \mathbf{u} = \nabla \times \mathbf{u}_{\text{rot}} + \nabla \times \mathbf{u}_{\text{div}} = \nabla \times \mathbf{u}_{\text{rot}} = \nabla(\nabla \cdot \mathbf{\Psi}) - \nabla^2\mathbf{\Psi}. \tag{8.47}$$

Later, we will see that these two equations can be solved for ϕ and $\mathbf{\Psi}$ up to a meaningless constant. Notice how the total flow field $\mathbf{u}$ is split up into a part that controls the curl and a part that controls the divergence. If either vanishes, then we can take the remaining of $\mathbf{u}_{\text{rot}}$ or $\mathbf{u}_{\text{div}}$ as the whole of the vector field. For example, above we showed that a constant density fluid obeys $\nabla \cdot \mathbf{u} = 0$. In general, a vector field whose divergence vanishes is called *solenoidal* or *rotational*. Thus, the whole velocity field in this case is given by $\mathbf{u}_{\text{rot}}$, which can in turn be generated by the curl of a streamfunction $\mathbf{\Psi}$. The way to find this streamfunction is to examine the curl of the velocity–which

we have already defined as the vorticity. Thus, the vorticity is the key indicator of flow in a constant density fluid.

You have reached the cutting edge of research! This topic is fresh and important enough that current publications feature extensions of the decomposition, like in my first paper (Fox-Kemper et al., 2003).

8.5.7 Conservative and Nonconservative Forces

Another use for the Helmholtz decomposition is categorizing forces into conservative and nonconservative forces. Boas (2006) discusses the relationship between work and potential energy. The relationship between work done W and force applied $\mathbf{F}$ is

$$W = \int_{\text{start}}^{\text{finish}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \mathrm{W}. \tag{8.48}$$

If the force and the work done result only from adding to or releasing energy from reversible process, e.g., releasing potential energy, then can consider the total energy budget as $W + \phi = 0$, where ϕ is the potential energy. In this case, the integral above is independent of path, and

$$W = -\phi, \tag{8.49}$$
$$\mathbf{F} = -\nabla\phi. \tag{8.50}$$

The (conservative) force resulting from a reversible process can be written as resulting from the gradient of the potential.

If both conservative and nonconservative (such as friction) forces act, then we can use our ∇ identities to identify how much of the force results from conservative forces ($\mathbf{F}_c = -\nabla\phi$) and how much results from nonconservative ones ($\mathbf{F}_n \neq$).

$$\mathbf{F} = \mathbf{F}_c + \mathbf{F}_n, \tag{8.51}$$
$$\nabla \times \mathbf{F} = \cancelto{0}{\nabla \times \mathbf{F}_c} + \nabla \times \mathbf{F}_n, \tag{8.52}$$
$$\nabla \cdot \mathbf{F} = \nabla \cdot \mathbf{F}_c + \cancelto{0}{\nabla \cdot \mathbf{F}_n} \tag{8.53}$$

So, the categorization of conservative and nonconservative forces is a Helmholtz decomposition on the forces.

8.5.8 S Waves and P Waves

Seismic waves are often categorized by whether they are waves of shearing (transverse) motion or compressional motion, because the properties of solids mean that solids deform to shearing motions differently than to compressional motions. Indeed, if you neglect viscosity, incompressible fluids (i.e., liquids) behave as though they do not compress at all and as though they do not resist shearing motions at all!

We define the displacement vector $\mathbf{u}(\mathbf{x}, t)$, which is a vector field that maps the initial position $\mathbf{x}$ of every piece of the solid to its time evolving location $\mathbf{y}$,

$$\mathbf{y} = \mathbf{x} + \mathbf{u}(\mathbf{x}, t), \tag{8.54}$$
$$\mathbf{u}(\mathbf{x}, t) = \mathbf{y} - \mathbf{x}. \tag{8.55}$$

For S waves, the displacements are purely solenoidal, and so could be related to a streamfunction,

$$\nabla \cdot \mathbf{u} = 0, \tag{8.56}$$
$$\mathbf{u} = \nabla \times \psi. \tag{8.57}$$

For P waves, the displacements are purely compressional, and so can be related to a potential,

$$\nabla \times \mathbf{u} = 0, \tag{8.58}$$
$$\mathbf{u} = \nabla \phi. \tag{8.59}$$

Thus, the separation of seismic waves into S waves and P waves is a Helmholtz decomposition.

Interestingly, near boundaries there are also waves that are neither divergent nor rotational– Rayleigh waves. Their displacement field is trapped near the boundary, and has the properties that

$$\nabla \cdot \mathbf{u} = 0, \text{and} \nabla \times \mathbf{u} = 0. \tag{8.60}$$

The surface water waves that we watch at the beach also are irrotational and nondivergent, and (Fox-Kemper et al., 2003) examines irrotational and nondivergent transport by turbulence.

The displacement gradient is a second-rank tensor (!) which relates the change of displacements from location to location, that is, how much the solid is deformed.

$$\frac{\partial u_i}{\partial x_j} = \nabla \mathbf{u}(\mathbf{x}, t). \tag{8.61}$$

8.5.9 Toroidal/Poloidal Decomposition

After the Helmholtz decomposition is the toroidal/poloidal decomposition, where a solenoidal vector field ($\mathbf{u}_{\text{rot}}$) can be further decomposed.

$$\mathbf{u}_{\text{rot}} = \mathbf{u}_{\text{tor}} + \mathbf{u}_{\text{pol}}, \tag{8.62}$$
$$\mathbf{u}_{\text{tor}} = \nabla \times (\psi_t \hat{\mathbf{r}}), \tag{8.63}$$
$$\mathbf{u}_{\text{pol}} = \nabla \times \left[\nabla \times (\psi_p \hat{\mathbf{r}})\right] = \nabla(\nabla \cdot \psi_p \hat{\mathbf{r}}) - \nabla^2 \psi_p \hat{\mathbf{r}}. \tag{8.64}$$

This decomposition helps to distinguish the symmetries of the vector fields around the origin, and is often used in seismology to decompose the patterns of deformation of earthquakes.

We will return to the Helmholtz and toroidal/poloidal decompositions when we have the tools to determine the streamfunctions and potentials.

8.6 The Fundamental Theorems: Integrating Vector Fields and Vector Derivatives

Reading: Boas (2006, 6.8-6.11)

In one variable, the fundamental theorem of calculus relates the integral along the independent variable of a differential to the value at the endpoints.

Theorem 8.1 *The Fundamental Theorem of Univariate Calculus:*

$$\int_u^v \left[\frac{\mathrm{d}}{\mathrm{d}t} f(t)\right] \mathrm{d}t = f(v) - f(u). \tag{8.65}$$

The definite integral of a (total) derivative with respect to the independent variable is the difference in values of the differentiated function at the endpoints.

This theorem comes up in rate and acceleration problems often. For example, the average velocity over a trip is the difference in distance between the starting location and the ending location over the total duration of the trip, which is an application of this theorem. Note that this result is independent of the detailed accelerations during the voyage.

8.7 Integrating Vectors

Since integration is a linear operation, we can indefinitely integrate vectors component by component and the result will still be a vector. Definite integrals of vectors require a bit more care. However, just as there were particular combinations of partial derivatives (div, grad, curl) that are important in differentiating vectors, there are particularly useful kinds of definite integrals that are useful.

8.8 The Fundamental Theorems

8.8.1 Gradient Theorem

Reading: Boas (2006, 6.8, 6.9)

There are vector field analogs of this relationship that are crucial in vector analysis. The first is the gradient theorem, which is closely related to the ordinary fundamental theorem, except *the path of the line integral through space is arbitrary.*

Theorem 8.2 *The Gradient Theorem:*

$$\int_{\mathbf{a}}^{\mathbf{b}} \nabla f(\mathbf{l}) \cdot \mathrm{d}\mathbf{l} = f(\mathbf{b}) - f(\mathbf{a}). \tag{8.66}$$

The path integral of the gradient of a function is the difference in values of the differentiated scalar function at the endpoints. Note that $\mathrm{d}\mathbf{l}$ *is a vector of length* $\mathrm{d}l$ *oriented tangent to the path of integration, and* $\mathbf{l}$ *represents each location to evaluate the gradient along the path. Also note that the result is the same over* any path *connecting the endpoints.*

Application: Average Slope

An application of the gradient theorem is to consider the average slope of topography (the gradient of elevation h) between two points. It is

$$\langle \nabla h \rangle = \frac{1}{|\mathbf{b} - \mathbf{a}|} \int_{\mathbf{a}}^{\mathbf{b}} \nabla h(\mathbf{l}) \cdot \mathrm{d}\mathbf{l} = \frac{1}{|\mathbf{b} - \mathbf{a}|} \left[h(\mathbf{b}) - h(\mathbf{a})\right]. \tag{8.67}$$

So, the average slope does not depend on the path taken. Some choices may have steeper bits sloping downward, but they are always compensated by upslopes elsewhere. Some paths may be longer, but with gentler slopes.

By contrast, the average velocity for a trip beginning at $\mathbf{a}$ at time t_a and ending at $\mathbf{b}$ at time t_b is *not* an application of the gradient theorem, but instead an application of a vector version of the fundamental theorem of calculus (8.65).

$$\langle \mathbf{u}(\mathbf{r},t) \rangle = \int_{t_a}^{t_b} \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\,\mathrm{d}t = \frac{\mathbf{r}(t_b)-\mathbf{r}(t_a)}{t_b-t_a}. \tag{8.68}$$

8.8.2 Divergence Theorem

Reading: Boas (2006, 6.9, 6.10)

The divergence theorem (a.k.a. Gauss's theorem or Ostrogradsky's theorem) relates the volume integral of a divergence to the surface area integral bounding the volume.

Theorem 8.3 *The Divergence Theorem:*

$$\iiint_V (\nabla\cdot\mathbf{v})\,\mathrm{d}V = \oiint_A \mathbf{v}\cdot\hat{\mathbf{n}}\,\mathrm{d}S. \tag{8.69}$$

The volume integral over V of a divergence of a vector is equal to the surface integral of the outward normal component of the vector over a bounding surface S that encloses the volume V. Note that the outward normal unit vector $\hat{\mathbf{n}}$ varies in direction.

Application: Budgets

The most common application of Gauss's theorem in geophysics is the conversion between a differential conservation equation and a budget equation over a volume. For example, the conservation of mass in a fluid, based on the density field $\rho(x,y,z,t)$ and fluid velocity $\mathbf{u}(x,y,z,t)$ is

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0. \tag{8.70}$$

The volume integral of this equation over any volume that is fixed in time (e.g., an estuary, a model gridcell, a melt inclusion) gives the rate of change of the mass inside the volume

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\iiint_V \rho\,\mathrm{d}V = \iiint_V \frac{\partial}{\partial t}\rho\,\mathrm{d}V = -\iiint_V \nabla\cdot(\rho\mathbf{u})\,\mathrm{d}V = -\oiint_A \rho\mathbf{u}\cdot\hat{\mathbf{n}}\,\mathrm{d}S \tag{8.71}$$

Thus, when the flow penetrates the surface bounding the volume, advecting dense or light fluid into or out of the volume, the mass inside the volume may change. The change is determined by the convergence (-1 times the divergence) of the density flux $\rho\mathbf{u}$.

8.8.3 Curl Theorem

Reading: Boas (2006, 6.9, 6.11)

The curl theorem (a.k.a. Stokes' theorem) relates the area integral of a curl to the line integral integral bounding the area.

Theorem 8.4 *The Curl Theorem:*

$$\iint_A (\nabla \times \mathbf{v}) \cdot \hat{\mathbf{n}} \, \mathrm{d}A = \oint_{\partial A} \mathbf{v}(\mathbf{l}) \cdot \mathrm{d}\mathbf{l}. \tag{8.72}$$

The surface integral over A of the normal component of a curl of a vector is equal to the line integral along the edge of that surface ∂A. Use the right hand rule, with thumb pointing in the direction of $\hat{\mathbf{n}}$ and fingers indicating the direction of the line integral element $\mathrm{d}\mathbf{l}$*. Note that the integral will be the same for* all surfaces *that meet at the same edge boundary.*

Application: Vorticity and Circulation

A common application of Stokes' theorem in geophysics is the conversion between vorticity ω (sometimes thought of as lines of vortex tubes penetrating a surface) and the circulation $\mathbf{\Gamma}$ around the edge of that surface, which is just the line integral of the velocity along the edge.

$$\omega = \nabla \times \mathbf{u}, \tag{8.73}$$

$$\iint_A \omega \cdot \hat{\mathbf{n}} \, \mathrm{d}A = \iint_A \nabla \times \mathbf{u} \cdot \hat{\mathbf{n}} \, \mathrm{d}A = \oint_{\partial A} \mathbf{u}(\mathbf{l}) \cdot \mathrm{d}\mathbf{l} = \Gamma. \tag{8.74}$$

A differential equation governing the evolution of vorticity ($\frac{\partial \omega}{\partial t} = \dots$) at every point in the fluid can be derived from the momentum equations, and integrating it over the volume gives the rate of change of circulation. This circulation constraint when applied to surfaces moving with the fluid is called Kelvin's Circulation Theorem.

8.9 Multivariate Taylor and Newton

Using the gradient, the linearized Taylor series for a function depending on a number of variables $(x_1, x_2, \dots, x_k)$ near $(a_1, a_2, \dots, a_n)$ can be written a little more simply.

$$f(x_1, x_2, \dots, x_k) = \sum_{n=0}^{\infty} \frac{1}{n!} \left([x_1 - a_1] \frac{\partial}{\partial x_1} + [x_2 - a_2] \frac{\partial}{\partial x_2} + \cdots + [x_k - a_k] \frac{\partial}{\partial x_k} \right)^n f(a_1, a_2, \dots a_k), \tag{8.75}$$

$$\approx f(\mathbf{a}) + [\mathbf{x} - \mathbf{a}] \cdot \nabla f(\mathbf{a}) \tag{8.76}$$

Where we consider the independent variables of the function as a vector: $\mathbf{x} = (x_1, x_2, \dots, x_k), \mathbf{a} = (a_1, a_2, \dots, a_k)$. The gradient of the function ∇ is taken with respect to each of these variables.

8.9.1 Multidimensional Newton-Raphson

Newton's method (or the Newton-Raphson method) is a method for improving guesses to the solution of a differentiable equation, which can be iterated (repeated) until a desired accuracy is reached (or no solution is found at all). Typically, the method results in a convergent sequence toward a solution of the equation. This is true in multidimensional cases as well as one-dimensional cases, but sometimes the multidimensional version is less likely to converge (see Section 10.7.2). The steps are as follows:

1. Arrange the equation so that it reads: $f(\mathbf{x}^*) = 0$. The * indicates a solution to the equation. There may be more than one...
2. Make a guess $\mathbf{x}_i$ of $\mathbf{x}^*$.
3. Find an improved version by evaluating the Taylor series of the function at $\mathbf{x}_i$, and choosing the next $\mathbf{x}_{i+1}$ where the Taylor series estimate is equal to zero,

$$f(\mathbf{x}_{i+1}) \approx f(\mathbf{x}_i) + \nabla f(x_i) \cdot (\mathbf{x}_{i+1} - \mathbf{x}_i) = 0, \tag{8.77}$$

$$\therefore \qquad \mathbf{x}_{i+1} = \mathbf{x}_i - \frac{f(\mathbf{x}_i)\nabla f(\mathbf{x}_i)}{\nabla f(\mathbf{x}_i) \cdot \nabla f(\mathbf{x}_i)}. \tag{8.78}$$

Step 3 can be repeated indefinitely. When you are near a solution, the steps will become increasingly small and you will converge. If the solution converged upon is not the desired one, or the iteration method diverges (which is rare), repeat the whole process from step 2 with a different initial guess. In a computer program, you can estimate the derivatives required by finite differences. We note that unlike the one-dimensional case, each Newton step may go in a different direction. It is therefore possible to have complicated, indirect paths to finding the solution, and it is much more likely that a solution will not converge.

8.10 Example Problems

8.10.1 Jargon to Argot

Example 8.1 *The symbol ∇ is described as a vector "operator." What is meant by operator?*

An operator is a mapping from one vector space to another or to a related scalar field. It is often used generically to describe the execution of an operation that takes a vector field and makes a scalar (called a functional in this case), or a vector field to a vector field (a vector function), or a scalar field to a vector field (a vector-valued function of a scalar).

Example 8.2 *Problem 6.4.17 Boas (2006).*

Here is the answer to 6.4.17, which wasn't assigned!
A force $\mathbf{F} = 2\hat{\mathbf{i}} - 3\hat{\mathbf{j}} + \hat{\mathbf{k}}$ acts at the point (1,5,2). Find the torque due to $\mathbf{F}$. (a) about the origin:

$$\mathbf{T} = \mathbf{r} \times \mathbf{F} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 1 & 5 & 2 \\ 2 & -3 & 1 \end{vmatrix} = \hat{\mathbf{i}}(5-(-6)) + \hat{\mathbf{j}}(4-1) + \hat{\mathbf{k}}(-3-10) = 11\hat{\mathbf{i}} + 3\hat{\mathbf{j}} - 13\hat{\mathbf{k}}.$$

(b) about the y-axis. We'll use (3.10) of Boas (2006), first we note that $\hat{\mathbf{n}} = \hat{\mathbf{j}}$, and thus

$$\hat{\mathbf{j}} \cdot \left(11\hat{\mathbf{i}} + 3\hat{\mathbf{j}} - 13\hat{\mathbf{k}}\right) = 3.$$

(c) about the line $x/2 = y/1 = z/(-2)$. We note that a point on this line is $x = 2, y = 1, z = -2$.

A unit vector in that direction is then,

$$\hat{\mathbf{n}} = \frac{2\hat{\mathbf{i}} + 1\hat{\mathbf{j}} - 2\hat{\mathbf{k}}}{\sqrt{2^2+1^2+2^2}} = \frac{2}{3}\hat{\mathbf{i}} + \frac{1}{3}\hat{\mathbf{j}} - \frac{2}{3}\hat{\mathbf{k}},$$
$$\hat{\mathbf{n}} \cdot \left(11\hat{\mathbf{i}} + 3\hat{\mathbf{j}} - 13\hat{\mathbf{k}}\right) = \left(\frac{2}{3}\hat{\mathbf{i}} + \frac{1}{3}\hat{\mathbf{j}} - \frac{2}{3}\hat{\mathbf{k}}\right) \cdot \left(11\hat{\mathbf{i}} + 3\hat{\mathbf{j}} - 13\hat{\mathbf{k}}\right) = \frac{22}{3} + 1 + \frac{26}{3} = \frac{51}{3} = 17.$$

Example 8.3 (Centripetal 2) *Problem 6.4.8 Boas (2006).*

In polar coordinates, the position vector of a particle is $\mathbf{r} = r\hat{\mathbf{r}}$. Using (4.13), find the velocity and acceleration of the particle.

$$\begin{aligned}
\mathbf{v} &= \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \hat{\mathbf{r}}\,\frac{\mathrm{d}r}{\mathrm{d}t} + r\,\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}t} = \hat{\mathbf{r}}\,\frac{\mathrm{d}r}{\mathrm{d}t} + r\hat{\phi}\,\frac{\mathrm{d}\phi}{\mathrm{d}t}\,, \\
\mathbf{a} &= \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}t}\,\frac{\mathrm{d}r}{\mathrm{d}t} + \hat{\mathbf{r}}\,\frac{\mathrm{d}^2r}{\mathrm{d}t^2} + \frac{\mathrm{d}r}{\mathrm{d}t}\,\hat{\phi}\,\frac{\mathrm{d}\phi}{\mathrm{d}t} + r\,\frac{\mathrm{d}\hat{\phi}}{\mathrm{d}t}\,\frac{\mathrm{d}\phi}{\mathrm{d}t} + r\hat{\phi}\,\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2}\,, \\
&= \hat{\phi}\,\frac{\mathrm{d}\phi}{\mathrm{d}t}\,\frac{\mathrm{d}r}{\mathrm{d}t} + \hat{\mathbf{r}}\,\frac{\mathrm{d}^2r}{\mathrm{d}t^2} + \frac{\mathrm{d}r}{\mathrm{d}t}\,\hat{\phi}\,\frac{\mathrm{d}\phi}{\mathrm{d}t} - r\hat{\mathbf{r}}\,\frac{\mathrm{d}\phi}{\mathrm{d}t}\,\frac{\mathrm{d}\phi}{\mathrm{d}t} + r\hat{\phi}\,\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2}\,, \\
&= \hat{\mathbf{r}}\left(\frac{\mathrm{d}^2r}{\mathrm{d}t^2} - r\left[\frac{\mathrm{d}\phi}{\mathrm{d}t}\right]^2\right) + \hat{\phi}\left(r\,\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} + 2\,\frac{\mathrm{d}\phi}{\mathrm{d}t}\,\frac{\mathrm{d}r}{\mathrm{d}t}\right).
\end{aligned}$$

Example 8.4 (Contrast to Exercise 8.5) *Problem 6.8.17 Boas (2006). Which, if either, of the two force fields* $\mathbf{F}_1 = ?y\mathbf{i} + x\mathbf{j} + z\mathbf{k}$, $\mathbf{F}_2 = y\mathbf{i} + x\mathbf{j} + z\mathbf{k}$ *from Chapter 6 is conservative? Calculate for each field the work done in moving a particle around the circle* $x = \cos t, y = \sin t$ *in the* (x, y) *plane. While similar to exercise 8.5, a shuffling of the values and directions leads to very different forces.*

$$\nabla \times \mathbf{F}_1 = -k\begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -y & x & z \end{vmatrix} = \hat{\mathbf{i}}(0) + \hat{\mathbf{j}}(0) + \hat{\mathbf{k}}(1-(-1)) = 2\hat{\mathbf{k}},$$

$$\nabla \times \mathbf{F}_2 = -k\begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ y & x & z \end{vmatrix} = \hat{\mathbf{i}}(0) + \hat{\mathbf{j}}(0) + \hat{\mathbf{k}}(1-1) = 0.$$

Thus, $\mathbf{F}_1$ is not conservative while $\mathbf{F}_2$ is. Therefore, no work is done by $\mathbf{F}_2$ in moving around a complete circle. $\mathbf{F}_1$ will take a little more work. We can calculate it in cylindrical coordinates, by first converting with some geometry from $\hat{\mathbf{i}}, \hat{\mathbf{j}}$ to $\hat{\mathbf{r}}, \hat{\phi}$

$$\begin{aligned}
\hat{\mathbf{i}} &= \hat{\mathbf{r}}\cos\phi - \hat{\phi}\sin\phi, \qquad \hat{\mathbf{j}} = \hat{\mathbf{r}}\sin\phi + \hat{\phi}\cos\phi, \\
x &= r\cos\phi, \qquad y = r\sin\phi, \\
\mathbf{F}_1 &= -y\hat{\mathbf{i}} + x\hat{\mathbf{j}} + z\hat{\mathbf{k}} = -r\sin\phi\left(\hat{\mathbf{r}}\cos\phi - \hat{\phi}\sin\phi\right) + r\cos\phi\left(\hat{\mathbf{r}}\sin\phi + \hat{\phi}\cos\phi\right) + z\hat{\mathbf{k}}, \\
\oint \mathbf{F}_1 \cdot \mathrm{d}\mathbf{l} &= \oint \mathbf{F}_1 \cdot \hat{\phi} r\,\mathrm{d}\phi = \oint \left(r\sin^2\phi + r\cos^2\phi\right) r\,\mathrm{d}\phi = 2\pi r^2 = 2\pi.
\end{aligned}$$

8.10.2 Scheming Schematics and Articulate Analysis

Example 8.5 (Gradient of r) *Problem 6.6.17 Boas (2006). Find ∇r, where $r = \sqrt{x^2+y^2}$, using (6.7) and also using (6.3). Show that your results are the same by using (4.11) and (4.12).*

$$6.3{:}\nabla\sqrt{x^2+y^2} = \hat{\mathbf{i}}\frac{\partial\sqrt{x^2+y^2}}{\partial x} + \hat{\mathbf{j}}\frac{\partial\sqrt{x^2+y^2}}{\partial y} = \frac{x\hat{\mathbf{i}}}{\sqrt{x^2+y^2}} + \frac{y\hat{\mathbf{j}}}{\sqrt{x^2+y^2}},$$

$$\hat{\mathbf{i}} = \hat{\mathbf{r}}\cos\phi - \hat{\phi}\sin\phi, \qquad \hat{\mathbf{j}} = \hat{\mathbf{r}}\sin\phi + \hat{\phi}\cos\phi, \qquad x = r\cos\phi, \qquad y = r\sin\phi,$$

$$\frac{x\hat{\mathbf{i}}}{\sqrt{x^2+y^2}} + \frac{y\hat{\mathbf{j}}}{\sqrt{x^2+y^2}} = \cos\phi\left(\hat{\mathbf{r}}\cos\phi - \hat{\phi}\sin\phi\right) + \sin\phi\left(\hat{\mathbf{r}}\sin\phi + \hat{\phi}\cos\phi\right) = \hat{\mathbf{r}},$$

$$6.7{:}\nabla r = \hat{\mathbf{r}}\frac{\partial r}{\partial r} = \hat{\mathbf{r}}$$

Quod erat demonstrandum.

Example 8.6 (Gravitational Potential of Thin Shell & Deep Domain) *Problem 6.8.20 Boas (2006). For motion near the surface of the earth, we usually assume that the gravitational force on a mass m is $\mathbf{F} = -mg\mathbf{k}$, but for motion involving an appreciable variation in distance r from the center of the earth, we must use $\mathbf{F} = -\frac{C}{r^2}\mathbf{e}_r = -\frac{C}{r^2}\frac{\mathbf{r}}{|\mathbf{r}|} = -\frac{C}{r^3}\mathbf{r}$ where C is a constant. Show that both these $\mathbf{F}$s are conservative, and find the potential for each.*

Thin, Cartesian system:

$$\mathbf{F} = -mg\hat{\mathbf{k}}, \qquad \nabla\times\mathbf{F} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & 0 & mg \end{vmatrix} = 0,$$

$$\int_0^z \mathbf{F}\cdot\mathrm{d}\mathbf{l} = \int_0^z \mathbf{F}\,\mathrm{d}z = \int_0^z -mg\,\mathrm{d}z = -mgz = \int_0^z -\nabla\phi\,\mathrm{d}z = \phi(0) - \phi(z).$$

Thus, $\mathbf{F} = -\nabla(mgz)$.
Thick, spherical system:

$$\mathbf{F} = -\frac{C\hat{\mathbf{r}}}{r^2} = -\frac{C\mathbf{r}}{r^3}, \qquad \nabla\times\mathbf{F} = -\frac{\hat{\theta}}{r}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\frac{C}{r^2}\right] + \frac{\hat{\phi}}{r}\left[\frac{\partial}{\partial\theta}\frac{C}{r^2}\right] = 0.$$

$$\int_{r_0}^r \mathbf{F}\cdot\mathrm{d}\mathbf{l} = \int_{r_0}^r \mathbf{F}\,\mathrm{d}r = -\int_{r_0}^r \frac{C}{r^2}\,\mathrm{d}r = \frac{C}{r} = -\int_{r_0}^r \nabla\phi\,\mathrm{d}r = -\phi(r) + \phi(r_0).$$

Note that we can't take $r = 0$ as the reference point (as the forces are not well-defined there), so instead we take an arbitrary constant radius r_0 as our reference. We choose $r_0 < r$, so that $\mathbf{F}\cdot\mathrm{d}\mathbf{l} = \mathbf{F}\,\mathrm{d}r$. The value of the potential there is not important, and our potential is (-C/r) so that $\mathbf{F} = \nabla\frac{C}{r}$.

8.11 Homework Problems

8.11.1 Manipulation

Exercise 8.1 (Centripetal) *Problem 6.3.17 Boas (2006).*

Exercise 8.2 (Direction of Decrease) *Problem 6.6.2 Boas (2006).*

Exercise 8.3 (Calculate Div, Grad, Curl) *Problem 6.7.1 Boas (2006).*

Exercise 8.4 (A Triple Product) *Problem 6.7.18 Boas (2006).*

Exercise 8.5 (Check for Conservative–Attraction to the origin) *Problem 6.8.10 Boas (2006). This problem was selected because it is very similar to Hooke's law, which is a fundamental of solid mechanics and gravitational attraction.*

8.11.2 Application

Exercise 8.6 (Gradient of r) *The combination of gravitation and the centrifugal force from the earth's rotation is a conservative force that can be expressed using the geopotential $\phi = mgz$, where z is distance from the surface and m and g are the constant mass and acceleration due to gravity. A motion that results in a change in geopotential indicates the possibility that energy can be extracted from the motion. Formulate a closed line integral (using Stokes theorem) for a route to school that proves that "in my day, we had to go to school uphill both ways!" cannot require a net expenditure of energy. Show that if there is a nonconservative force (e.g., viscosity of air, rusty bike wheels, etc.) there may be a nonzero expenditure of energy in the round trip.*

Exercise 8.7 (Stokes Theorem) *a) Calculate the curl of the vector $\mathbf{v} = (-y, x, 0)$. b) Take the area integral of $(\nabla \times \mathbf{v}) \cdot \hat{\mathbf{n}}$ over the surface of a disc bounded by $1 = x^2 + y^2$. c) Take the area integral of $(\nabla \times \mathbf{v}) \cdot \hat{\mathbf{n}}$ over the surface of the half-sphere bounded by $1 = x^2 + y^2 + z^2$ where $z \geq 0$. d) Use Stokes' theorem to find a line integral equal to both b) and c).*

Exercise 8.8 (Divergence Theorem) *a) Calculate the divergence of the vector $\mathbf{v} = (-y, x, 1)$. b) What is the area integral of $\oiint \mathbf{v} \cdot \hat{\mathbf{n}} \, \mathrm{d}S$, over any given closed area? c) Can you say anything about its integral over an area that is not closed?*

Table of Vector Identities Involving ∇

Note carefully that ϕ and ψ are scalar functions; $\mathbf{U}$ and $\mathbf{V}$ are vector functions. Formulas are given in rectangular coordinates; for other coordinate systems, see Chapter 10, Section 9.

(a) $\nabla \cdot \nabla\phi = \text{div grad}\,\phi = \nabla^2\phi = \text{Laplacian}\ \phi = \dfrac{\partial^2\phi}{\partial x^2} + \dfrac{\partial^2\phi}{\partial y^2} + \dfrac{\partial^2\phi}{\partial z^2}$

(b) $\nabla \times \nabla\phi = \text{curl grad}\,\phi = 0$

(c) $\nabla(\nabla \cdot \mathbf{V}) = \text{grad div}\,\mathbf{V}$
$$= \mathbf{i}\left(\frac{\partial^2 V_x}{\partial x^2} + \frac{\partial^2 V_y}{\partial x \partial y} + \frac{\partial^2 V_z}{\partial x \partial z}\right) + \mathbf{j}\left(\frac{\partial^2 V_x}{\partial x \partial y} + \frac{\partial^2 V_y}{\partial y^2} + \frac{\partial^2 V_z}{\partial y \partial z}\right) + \mathbf{k}\left(\frac{\partial^2 V_x}{\partial x \partial z} + \frac{\partial^2 V_y}{\partial y \partial z} + \frac{\partial^2 V_z}{\partial z^2}\right)$$

(d) $\nabla \cdot (\nabla \times \mathbf{V}) = \text{div curl}\,\mathbf{V} = 0$

(e) $\nabla \times (\nabla \times \mathbf{V}) = \text{curl curl}\,\mathbf{V} = \nabla(\nabla \cdot \mathbf{V}) - \nabla^2\mathbf{V} = \text{grad div}\,\mathbf{V} - \text{Laplacian}\ \mathbf{V}$

(f) $\nabla \cdot (\phi\mathbf{V}) = \phi(\nabla \cdot \mathbf{V}) + \mathbf{V} \cdot (\nabla\phi)$

(g) $\nabla \times (\phi\mathbf{V}) = \phi(\nabla \times \mathbf{V}) - \mathbf{V} \times (\nabla\phi)$

(h) $\nabla \cdot (\mathbf{U} \times \mathbf{V}) = \mathbf{V} \cdot (\nabla \times \mathbf{U}) - \mathbf{U} \cdot (\nabla \times \mathbf{V})$

(i) $\nabla \times (\mathbf{U} \times \mathbf{V}) = (\mathbf{V} \cdot \nabla)\mathbf{U} - (\mathbf{U} \cdot \nabla)\mathbf{V} - \mathbf{V}(\nabla \cdot \mathbf{U}) + \mathbf{U}(\nabla \cdot \mathbf{V})$

(j) $\nabla(\mathbf{U} \cdot \mathbf{V}) = \mathbf{U} \times (\nabla \times \mathbf{V}) + (\mathbf{U} \cdot \nabla)\mathbf{V} + \mathbf{V} \times (\nabla \times \mathbf{U}) + (\mathbf{V} \cdot \nabla)\mathbf{U}$

(k) $\nabla \cdot (\nabla\phi \times \nabla\psi) = 0$

Table 8.1: Reproduction of page 339 of Boas (2006). Note that all of these identities are true regardless of choice of coordinate system.

Chapter 9

Ordinary Differential Equations

9.1 Introduction–Let's Get Physical

We now have in place enough tools to talk about solving real problems–at least streamlined versions of real problems. So, we will start motivating our study of ordinary differential equations with real applications.

Even so, we will need to keep track of what kinds of equations are produced, because identifying the kind of equation is key to choosing a solution method. Are they:

- Algebraic Equations or Differential Equations?
- Ordinary or Partial Differential Equations?
- First-order, Second-order, Other? (order refers to the highest derivative in the differential equation)
- Linear or Nonlinear in the Unknowns?
- Constant Coefficient or Variable Coefficient?
- Inhomogeneous (Sources & Sinks) or Homogeneous (Just Free Modes and Responses to Boundary or Initial Conditions)

9.2 First-Order Equations

Reading: Boas (2006)[8.1-8.3].

First, we will begin with the classic linear, first order, ordinary differential equation. Boas discusses how to arrive at the general solution to such an equation on pg. 401. The equation takes the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x). \tag{9.1}$$

This equation is first order, because the only derivative in it is a first derivative. It is a linear equation, since every operation on y is a linear one–including the derivative. Note that y is the

unknown, not x, which might be called the independent variable or the *coordinate* variable. This equation could be nonlinear in x, e.g., $P(x) = x^2$, and it would still be called a linear differential equation because it would still be linear in y. If $Q(x) = 0$, the equation is *homogeneous*, or unforced.

There are many methods for solving differential equations, but here we will often take advantage of general solutions, or what I call "guess and check" where you have solved enough similar equations to know what the solution will be like. The general solution to this equation is

$$y = e^{-\int P(x)\,\mathrm{d}x} \int \left(Q(x) e^{\int P(x)\,\mathrm{d}x}\,\mathrm{d}x \right) + C e^{-\int P(x)\,\mathrm{d}x}. \tag{9.2}$$

You can verify that this is a solution to the equation by differentiation.

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{\mathrm{d}e^{-\int P(x)\,\mathrm{d}x}}{\mathrm{d}x} \int \left(Q(x) e^{\int P(x)\,\mathrm{d}x}\,\mathrm{d}x \right) + e^{-\int P(x)\,\mathrm{d}x} \frac{\mathrm{d}\int \left(Q(x) e^{\int P(x)\,\mathrm{d}x}\,\mathrm{d}x \right)}{\mathrm{d}x} + \frac{\mathrm{d}C e^{-\int P(x)\,\mathrm{d}x}}{\mathrm{d}x}, \\
&= P(x) e^{-\int P(x)\,\mathrm{d}x} \int \left(Q(x) e^{\int P(x)\,\mathrm{d}x}\,\mathrm{d}x \right) + e^{-\int P(x)\,\mathrm{d}x} \left(Q(x) e^{\int P(x)\,\mathrm{d}x} \right) + P(x) C e^{-\int P(x)\,\mathrm{d}x}, \\
&= P(x) y + e^{\int P(x)\,\mathrm{d}x - \int P(x)\,\mathrm{d}x} Q(x), \\
&= P(x) y + Q(x).
\end{aligned}$$

It is often the case that it is easier to check that you have a solution to a differential equation than it is to find the solution, here it certainly is! This solution can be used for all linear, first-order equations, although it might be hard to evaluate the integrals required. Even if you can't do the integration analytically, this form can still be useful in conjunction with computer approximations to the integrals.

9.2.1 Application: Hydrostatic Balance

In motionless fluids, the balance of vertical forces is one between gravity and the vertical pressure gradient force. Physically, we can think of this balance as "the pressure within the fluid is the weight per unit area above." For constant density fluids, this balance is pretty easy to think about, but for variable density fluids, a differential equation is the natural way to express the balance, it is

$$\frac{\partial p}{\partial z} = -\rho g. \tag{9.3}$$

The pressure is p, the distance in the vertical is z, ρ is density (which can be a function of z), and g is the gravitational acceleration (which also can be a function of z). Does this equation suit the form of (9.1)? Which is the variable, and which is the coordinate? Is it linear and first-order?

Let's consider first what happens if density is constant $\rho = \rho_0$. We can use our general solution (9.2), but since $P(z) = 0$, it's not hard to see that the general solution is just

$$p = -\int \rho g\,\mathrm{d}z + C. \tag{9.4}$$

C is a constant of integration. Alternatively, we could have performed a definite integral over a fixed range of z, for example,

$$p(z) - p(0) = -\int_0^z \rho g\,\mathrm{d}z. \tag{9.5}$$

Where does the value for $p(0)$ come from? It is from additional information, called boundary conditions or initial conditions, which must always be given along with a differential equation in order to find a solution that doesn't have integration constants lingering. Usually, there will be as many boundary conditions as the order of the differential equation.

So, let's examine a few specific examples: atmospheric pressure, ocean pressure, and Airy isostasy. Atmospheric pressure is the easiest one to begin with, as up in space, both the density of the atmosphere and the pressure are zero.

The pressure of a gas can be approximated with the ideal gas law, which can be written as

$$p = \rho R_s T. \tag{9.6}$$

Where the variables are pressure, density, specific ideal gas constant, and temperature of the gas. To a sufficient approximation for the purposes here, we can assume that the atmosphere is all the same temperature (isothermal at T_0) and constant composition, then $R_s \approx 287.058\,\mathrm{J\,kg^{-1}\,K^{-1}}$ and $p \propto \rho$. Plugging into the hydrostatic relation,

$$\frac{\partial p}{\partial z} = -p\frac{g}{R_s T_0}. \tag{9.7}$$

If we neglect variations in g (which is acceptable if, like on Earth, the atmosphere is thin when compared to the radius of the planet), then the hydrostatic relation is linear with constant coefficients. The solution is therefore given by (9.2) with $Q = 0, P = R_s T_0 g$, and the solution is

$$p = Ce^{-gz/(R_s T_0)}. \tag{9.8}$$

The value of C is set to be the pressure at $z = 0$, or p_0, which we can consider to be sea level pressure near 1 bar or $1 \times 10^5\,\mathrm{Pa}$. If there were no winds, then the pressure everywhere at sea level could take this value, but pressure varies as the atmosphere is moving and thus not actually hydrostatic. At very large altitudes, the pressure goes to zero. The rate at which it decays is $g/(R_s T_0)$, which is sometimes called the density lapse rate and is $0.11\,\mathrm{km^{-1}}$, or the pressure and density of air decrease by a factor of e with every 9 km of elevation. Returning to the notion that the hydrostatic pressure is the weight per unit area above, $1 \times 10^5\,\mathrm{Pa} = 1 \times 10^5\,\mathrm{N\,m^{-2}} = (9.81\,\mathrm{m\,s^{-2}})1 \times 10^4\,\mathrm{kg\,m^{-2}}$.

The ocean is a nearly incompressible liquid, so instead of a density that is proportional to pressure, it is very nearly constant at $\rho_0 = 1 \times 10^3\,\mathrm{kg\,m^{-3}}$. So, for the ocean, the hydrostatic relation is

$$\frac{\partial p}{\partial z} = -\rho_0 g. \tag{9.9}$$

The solution to this equation is easily gotten from (9.2) with $Q = -\rho_0 g, P = 0$, or

$$p = \int (-\rho_0 g\,\mathrm{d}z) + p_0 = -\rho_0 g z + p_0. \tag{9.10}$$

The pressure increases *linearly* instead of *exponentially*. However, both linear and exponential functions can be solutions to *linear differential equations!* If we continue to take z to be zero at sea level, then p_0 is still $1 \times 10^5\,\mathrm{Pa}$. As z becomes more negative, the oceanic pressure increases. However, since seawater is much heavier than air, the pressure increases at a rate of $\rho_0 g = 1 \times 10^3\,\mathrm{kg\,m^{-2}}(9.81\,\mathrm{m\,s^{-2}}) = 1\mathrm{bar}/10\,\mathrm{m}$. That is, every 10 meters of ocean water weighs the same as the whole atmosphere above it.

Note how the boundary condition in the oceanic case was the pressure matching at the air-sea interface. The idea of matching pressures, one at a time, is the basis of Airy isostasy, which is a simple theory to explain the thickness of continental crust versus seafloor versus crust under mountain ranges (11.54). In this model, consider beginning at the top of the domain and integrating downward. The pressure at isostatic compensation surface 2 is $p_0 + \rho_c g c$ under the plains, $p_0 e^{-g h_t/(R_s T_0)} + \rho_c g (h_t + c)$ under the mountains and $p_0 + \rho_w g h_2 + \rho_c g (c - h_2 - b_2) + \rho_m g b_2$. If the pressure is to be equal horizontally at the mountaintops (i.e., hydrostatic atmosphere) and the pressure is to be equal horizontally somewhere in the mantle (isostatic compensation surface 1), and $\rho_w \approx 1 \times 10^3 \,\mathrm{kg\,m^{-3}} < \rho_c \approx 2.8 \times 10^3 \,\mathrm{kg\,m^{-3}} < \rho_m \approx 3.3 \times 10^3 \,\mathrm{kg\,m^{-3}}$, then it is a straightforward exercise to integrate again to isostatic compensation surface 1 and see that the crustal thickness must be greater under mountains and less under the sea. If not, the crust would rise or sink to compensate (as it indeed does during ice ages, when water, or snow and ice, covers the mountains and plains, too).

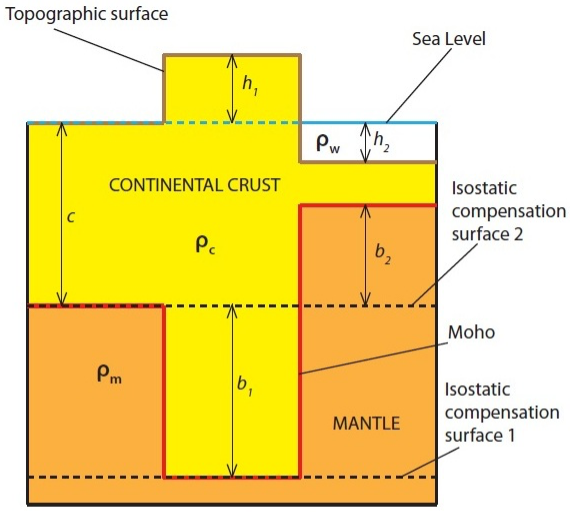


Figure 9.1: A schematic of some typical versions of Airy isostasy, (Wikimedia Commons)

9.2.2 Application: Estuarine Flow

The most essential use of the equations of motion of a fluid is to calculate budgets over a known volume or mass of fluid.

Perhaps the nicest example of this in oceanography is the flow into and out of estuaries. This estuarine circulation is a useful example of the use of volume and salt conservation attributed to Knudsen (1900).

Here, we will:

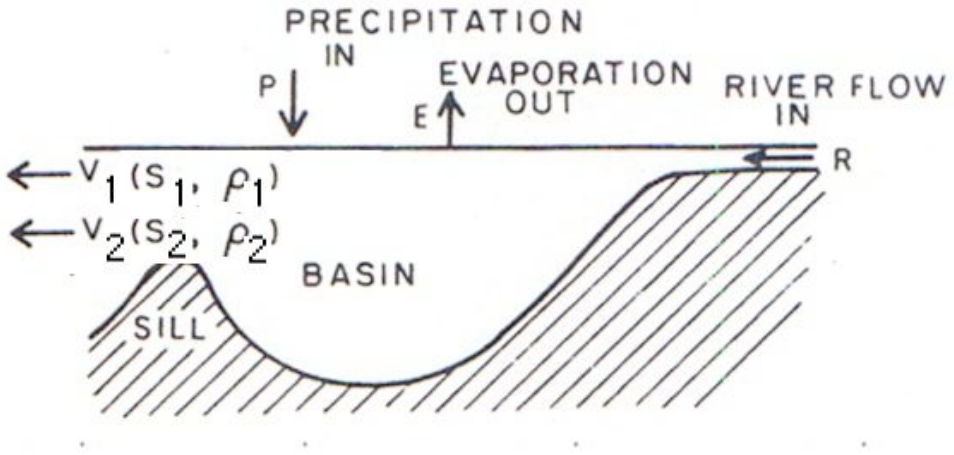


Figure 9.2: This figure schematizes the situation (Pickard and Emery, 1990). Generally, there is a two-layer flow over the sill, runoff, evaporation and precipitation.

1. derive the budgets for volume and salt conservation,
2. then do examples from the Black and Mediterranean Sea (from Pickard and Emery, 1990),
3. consider pollutants along with salt,
4. and finally solve a time-dependent differential equation problem to demonstrate the role of 'the flushing timescale'.

We begin with the conservation of mass

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{9.11}$$

Which is replaced by conservation of volume (for an incompressible fluid, where ρ can be approximated as a constant in space and time):

$$\nabla \cdot \mathbf{u} = 0. \tag{9.12}$$

Conservation of salt is (neglecting molecular diffusion):

$$\frac{\partial S}{\partial t} + \nabla \cdot (S \mathbf{u}) = 0 \tag{9.13}$$

We note that salinity can change (e.g., by evap./precip.), but it is the quantity of water that changes not the mass of salt.

We proceed by volume integrating the differential equations over the estuary.

$$\nabla \cdot \mathbf{u} = 0 \tag{9.14}$$

applies everywhere, so

$$0 = \iiint \nabla \cdot \mathbf{u} \, \mathrm{d}V = \oiint \mathbf{v} \cdot \hat{\mathbf{n}} \, \mathrm{d}A. \tag{9.15}$$

Where $\iiint \mathrm{d}V$ is by Gauss's divergence theorem equivalent to $\oiint \mathrm{d}A$ over the enclosed surrounding surface. We break up the surface integral, to find out how the total volume of fluid in the estuary is conserved:

$$\begin{aligned} 0 &= \oiint \mathbf{u} \cdot \hat{\mathbf{n}} \mathrm{d}A, \\ &= \iint_{A_1} \mathbf{u}_1 \,\mathrm{d}A + \iint_{A_2} \mathbf{u}_2 \,\mathrm{d}A + \iint_{A_E} E \,\mathrm{d}A - \iint_{A_P} P \,\mathrm{d}A + \iint_{A_R} \mathbf{u}_r \,\mathrm{d}A. \end{aligned} \tag{9.16}$$

Renaming using volume fluxes $[V] = L^3/T$:

$$0 = \int_{A_1} \mathbf{u}_1 \,\mathrm{d}A + \int_{A_2} \mathbf{u}_2 \,\mathrm{d}A + \int_{A_E} E \,\mathrm{d}A - \int_{A_P} P \,\mathrm{d}A + \int_{A_R} \mathbf{u}_r \,\mathrm{d}A.$$

$$0 = V_1 + V_2 + A_E E - A_P P - R, \tag{9.17}$$

$$\equiv V_1 + V_2 - F. \tag{9.18}$$

Where F is the freshwater supplied to the estuary by the combination of precipitation and runoff minus evaporation.

Recall: conservation of salt gives the equation for salinity. We can use it together with $\nabla \cdot \mathbf{u} = 0$.

$$0 = \frac{\partial S}{\partial t} + \nabla \cdot (\mathbf{u} S). \tag{9.19}$$

Integrate to find a steady state ($\frac{\partial S}{\partial t} \approx 0$) balance:

$$0 = \iiint \nabla \cdot \mathbf{u} S \;\mathrm{d}V = \oiint S \mathbf{u} \cdot \hat{\mathbf{n}} \,\mathrm{d}A. \tag{9.20}$$

Little salt in rivers, and evaporation & precipitation carry no salt, so

$$0 = \oiint S \mathbf{u} \cdot \hat{\mathbf{n}} \,\mathrm{d}A = \iint_{A_1} S \mathbf{u} \cdot \hat{\mathbf{n}} \,\mathrm{d}A + \iint_{A_2} S \mathbf{u} \cdot \hat{\mathbf{n}} \,\mathrm{d}A.$$

We want to relate the inflowing or outflowing salinities to the average salinities in the basin they come from, which we will do in a moment. For now, formally, we can define velocity-weighted average salinities S_1 and S_2, so

$$\begin{aligned} 0 &= \int_{A_1} S \mathbf{u} \cdot \hat{\mathbf{n}} \,\mathrm{d}A + \int_{A_2} S \mathbf{u} \cdot \hat{\mathbf{n}} \,\mathrm{d}A, \\ &= S_1 V_1 + S_2 V_2. \end{aligned}$$

Where we define

$$S_1 \equiv \frac{\int_{A_1} S \mathbf{u} \cdot \hat{\mathbf{n}} \,\mathrm{d}A}{\int_{A_1} \mathbf{u} \cdot \hat{\mathbf{n}} \,\mathrm{d}A}.$$

But, more loosely, we often assume that each basin is well-mixed in the interior, so there is a typical incoming salinity and a typical outgoing salinity, which makes approximate values for S_1 and S_2 clear.

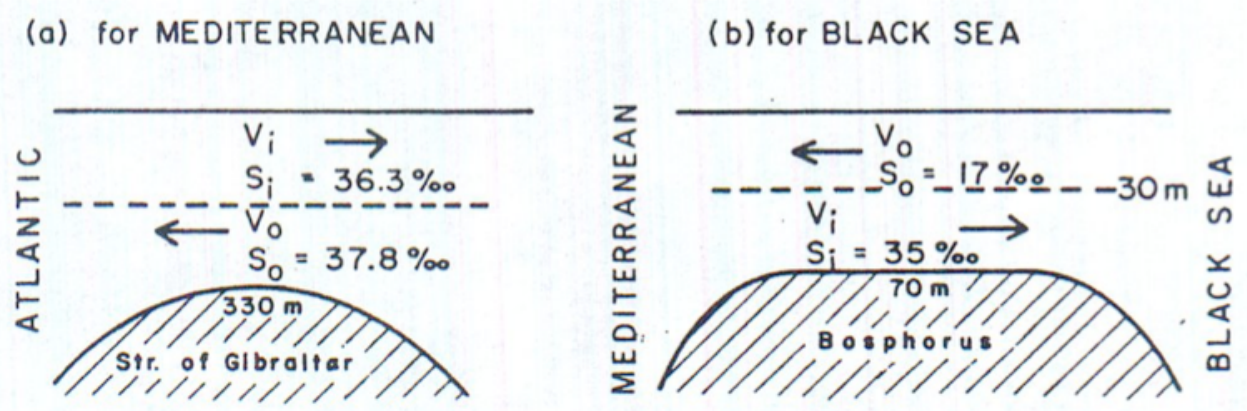


Figure 9.3: Classic exchange flows from (Pickard and Emery, 1990). The Mediterranean and the Black Sea exchange flows.

So, we have two equations so far

$$S_1 V_1 = -S_2 V_2, \qquad V_1 + V_2 = F.$$

We can eliminate either V_1 or V_2 using

$$V_1 = \frac{-S_2 V_2}{S_1}, \qquad V_2 = \frac{-S_1 V_1}{S_2}.$$

To give

$$F = V_2 \frac{S_1 - S_2}{S_1}, \qquad F = V_1 \frac{S_2 - S_1}{S_2}. \tag{9.21}$$

We will now do two classic examples from Pickard and Emery (1990): the Mediterranean and the Black Sea. The Mediterranean has a sill depth (at the Strait of Gibraltar) of 330m. It is observed that

$$\begin{aligned} S_1 &= 36.3‰, \\ S_2 &= 37.8‰, \\ V_1 &= -1.75 \times 10^6 \, \mathrm{m^3 \, s^{-1}} \equiv -1.75 \mathrm{Sv}. \end{aligned}$$

Where the Sverdrup, $1 \times 10^6 \, \mathrm{m^3 \, s^{-1}} \equiv 1 \mathrm{Sv}$, is a useful oceanographic unit.

So, we can infer from

$$\begin{aligned} &S_1 = 36.3‰, \qquad S_2 = 37.8‰, \\ &V_1 = -1.75 \mathrm{Sv}, \\ &V_2 = \frac{-S_1 V_1}{S_2}, \qquad F = V_1 \frac{S_2 - S_1}{S_2}, \end{aligned}$$

that

$$V_2 = 1.68 \mathrm{Sv}, \qquad F = -0.07 \mathrm{Sv}.$$

So, for a tiny amount of net freshwater loss (through evaporation exceeding precipitation and runoff), a huge exchange flow is required, with an outflow of salty water exiting at depth and fresher Atlantic water entering at the surface.

The Black Sea has a sill depth (at the Bosphorus) of 70m. It is observed that

$$\begin{aligned} S_1 &= 17‰, \\ S_2 &= 35‰, \\ V_1 &= 13 \times 10^3 \,\mathrm{m}^3\,\mathrm{s}^{-1}. \end{aligned}$$

So, we can infer that

$$V_2 = -6 \times 10^3 \,\mathrm{m}^3\,\mathrm{s}^{-1}, \qquad F = 7 \times 10^3 \,\mathrm{m}^3\,\mathrm{s}^{-1}.$$

Compare the two basins:
Mediterranean:

$$\begin{aligned} &S_1 = 36.3‰, S_2 = 37.8‰ \\ &V_1 = -1.75 \times 10^6 \,\mathrm{m}^3\,\mathrm{s}^{-1}, V_2 = 1.68 \times 10^6 \,\mathrm{m}^3\,\mathrm{s}^{-1}, \\ &F = -7 \times 10^4 \,\mathrm{m}^3\,\mathrm{s}^{-1}. \end{aligned}$$

Black:

$$\begin{aligned} &S_1 = 17‰, S_2 = 35‰ \\ &V_1 = 13 \times 10^3 \,\mathrm{m}^3\,\mathrm{s}^{-1}, V_2 = -6 \times 10^3 \,\mathrm{m}^3\,\mathrm{s}^{-1}, \\ &F = 7 \times 10^3 \,\mathrm{m}^3\,\mathrm{s}^{-1}. \end{aligned}$$

Mediterranean has outflow at depth, which is 25 times the volume of freshwater. Black has inflow at depth, which is nearly the same as the volume of freshwater.

Key differences: the amount of mixing in basin (Med. has $S_1 \approx S_2$, while Black has $S_1 \ll S_2$), and inflow/outflow at surface governed by freshwater deficit/supply (assuming $S_1 < S_2$).

We could also treat pollutants $\mathcal{P}$, not just salt. We do not expect (fear?) that there will be enough pollutants to appreciably change the density of the fluid or the flow rate, so we can incorporate the pollutants into the same steady state solutions we just investigated. Let's begin, for simplicity, with a steady source of polluted river input of a given concentration $\mathcal{P}_R$.

$$\frac{\partial \mathcal{P}}{\partial t} + \nabla \cdot (\mathcal{P}\mathbf{u}) \approx \nabla \cdot (\mathcal{P}\mathbf{u}) \approx 0. \tag{9.22}$$

The pollutants will have river sources, and potentially an exchange at sill, so the steady equation is

$$V_1 \mathcal{P}_1 + V_2 \mathcal{P}_2 = R\mathcal{P}_R. \tag{9.23}$$

Suppose we have constant pollutant concentration in river discharge in Black Sea, then

$$F \approx R = 7 \times 10^3 \,\mathrm{m}^3\,\mathrm{s}^{-1}, \qquad V_1 \approx 2R, \qquad \mathcal{P}_2 \approx 0.$$

So, the steady state result will be

$$V_1\mathcal{P}_1 + V_2\mathcal{P}_2 = R\mathcal{P}_R,$$
$$\mathcal{P}_1 \approx \frac{1}{2}\mathcal{P}_R.$$

Thus, the incoming pollution is diluted to only half the original riverine concentration in the Black Sea.

If we follow the pollutant through the Bosphorus ($\mathcal{P}_\mathcal{B}$) out into the Mediterranean (treat it as a river source there), then we have Black Sea sources:

$$\mathcal{P}_B \approx \frac{1}{2}\mathcal{P}_R, \qquad V_B = 13 \times 10^3\,\mathrm{m}^3\,\mathrm{s}^{-1}$$

And the Med. steady-state budget will be

$$V_1\mathcal{P}_1 + V_2\mathcal{P}_2 = V_B\mathcal{P}_B \approx V_B\frac{1}{2}\mathcal{P}_R.$$

Using our Med. numbers and a clean Atlantic,

$$V_1 = -1.75 \times 10^6\,\mathrm{m}^3\,\mathrm{s}^{-1}, V_2 = 1.68 \times 10^6\,\mathrm{m}^3\,\mathrm{s}^{-1},$$
$$\mathcal{P}_1 = 0.$$

We find the Med outflow is very dilute

$$\mathcal{P}_2 \approx \frac{1}{260}\mathcal{P}_R.$$

Reinforcing our notion that the Med. is better mixed than the Black Sea.

We have, however assumed a steady state pollution problem and consequence. While volume and salinity conservation didn't require strong time dependence (at least, when neglecting seasonality, etc.), pollutants do, because they often result from temporary spills or begin when a factory is built, etc.

Suppose we start polluting the rivers all of a sudden, when there is no pollution of that type in the Black Sea or the Mediterranean, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint \mathcal{P}\,\mathrm{d}V = -\oiint \mathcal{P}\mathbf{u}\cdot\hat{\mathbf{n}}\,\mathrm{d}S$$

We know the steady-state solution if this new pollution source stays constant $\mathcal{P}_{ss}$ (we just did it!), so let's see how long it takes to get there.

How long does it take to reach steady state? We can subtract the volume integral of the steady state solution from that of the time-dependent one, to find

$$\frac{d}{dt}\iiint(\mathcal{P}-\mathcal{P}_{ss})\,\mathrm{d}V = -\oiint(\mathcal{P}-\mathcal{P}_{ss})\mathbf{u}\cdot\hat{\mathbf{n}}\,\mathrm{d}S,$$
$$\mathrm{Vol.}\frac{d}{dt}\langle\mathcal{P}-\mathcal{P}_{ss}\rangle = V_R(\mathcal{P}_R-\mathcal{P}_{R;ss}) - V_{out}(\mathcal{P}_{out}-\mathcal{P}_{out;ss}),$$
$$\frac{d}{dt}\langle\mathcal{P}-\mathcal{P}_{ss}\rangle = -\frac{V_{out}}{\mathrm{Vol.}}(\mathcal{P}_{out}-\mathcal{P}_{out;ss}).$$

Angle brackets are volume averages, and Vol. is the volume. The last step followed from the steady state solution (9.23). If we assume a well-mixed basin, the outflow concentration, P_{out}, will be near the average $\langle P \rangle$, so

$$\frac{d}{dt}\langle \mathcal{P} - \mathcal{P}_{ss} \rangle + \frac{V_{out}}{\text{Vol.}} \langle \mathcal{P} - \mathcal{P}_{ss} \rangle \approx 0$$

This is our desired linear, homogeneous, constant coefficient differential equation. The (constant) coefficient that is outflow over volume is one over the *flushing time*, $\tau_f = Vol./V_{out}$. In the Mediterranean (Vol.=$3.8 \times 10^{15}\,\text{m}^3$) τ_f is about 70 yrs, for the Black (Vol.=$6 \times 10^{14}\,\text{m}^3\,\text{s}^{-1}$), it's about 1500 yrs.

How to solve it? We can guess that the solution looks like it might be an exponential, since the derivative of $\langle \mathcal{P} - \mathcal{P}_{ss} \rangle$

$$\frac{d}{dt}\langle P - P_{ss} \rangle \approx -\frac{V_{out}}{\text{Vol.}} \langle P - P_{ss} \rangle \rightarrow \qquad\qquad \langle P - P_{ss} \rangle \approx A e^{-t\frac{V_{out}}{\text{Vol.}}}$$

The only question remaining is how we find the constant A. It comes from the initial condition, which was to have no pollution in the estuary, which we can arrange for by a choice of $A = -\langle \mathcal{P}_{ss} \rangle$,

$$\langle \mathcal{P} \rangle \approx \langle \mathcal{P}_{ss} \rangle + A e^{-t\frac{V_{out}}{\text{Vol.}}} = \langle \mathcal{P}_{ss} \rangle \left(1 - e^{-t\frac{V_{out}}{\text{Vol.}}} \right) \tag{9.24}$$

This formula works for both the Black Sea (where $\langle \mathcal{P}_{ss} \rangle \approx \frac{1}{2}\langle \mathcal{P}_R \rangle$ and $\tau_f \approx 1500$ years) and the Mediterranean (where $\langle \mathcal{P}_{ss} \rangle \approx \frac{1}{260}\langle \mathcal{P}_R \rangle$ and $\tau_f \approx 70$ years). Alternatively, we could have used the formula on Boas (2006) pg. 401, which in this case is just $Ae^{-t\frac{V_{out}}{\text{Vol.}}}$, so it amounts to the same thing.

OK, the differential equation part of that went quickly, so let's review what happened. We boiled our partial differential equations for mass, salt, and pollution to ordinary differential equations for the same. We then set the time derivatives equal to zero and found steady-state solutions for the (algebraic) relations that resulted, which is what Knudsen (1900) proposed. Then, we decided to keep the time derivative on the pollution concentration equation. This resulted in a first-order, constant-coefficient, homogeneous differential equation. We solved it by guessing an exponential and checking the solution, which required setting one constant to fit the initial condition (it was only one constant because the equation was first order).

9.2.3 Application: Earth Energy Balance Model

Let's try a slightly more complicated problem, which still results in a first-order differential equation–albeit a nonlinear one (Sellers, 1969; Ghil and Childress, 1987; Zaliapin and Ghil, 2010). Like the salinity and pollution problem, it is an integration of a budget. In this case, it is an energy budget of the whole earth. At the top of the atmosphere, energy can only be transported into and out of the earth system by radiation. Thus, the two transport terms in this theory are the incoming and outgoing radiation from the earth. So, the following system of equations is used to predict the

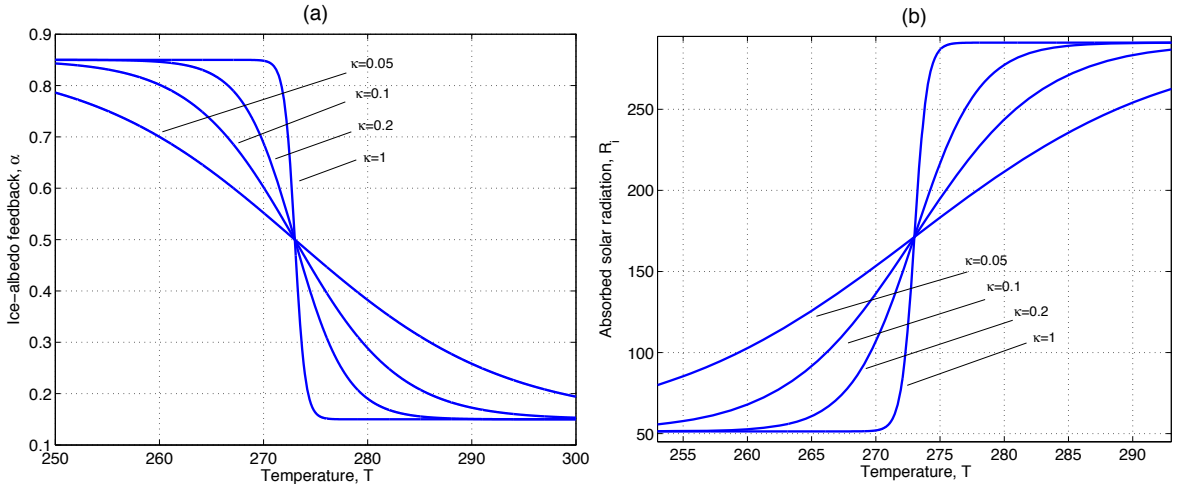


Figure 9.4: Figures from Zaliapin and Ghil (2010) showing the changes in albedo and absorbed radiation.

behavior of the mean temperature T.

$$c\,\frac{\mathrm{d}T}{\mathrm{d}t} = R_i(T) - R_o(T), \tag{9.25}$$

$$R_i = \mu Q_0(1-\alpha(T)) \approx \mu Q_0 \left[1 - \underbrace{\left(c_1 + \frac{c_2}{2}\left[1-\tanh(\kappa[T-T_c]\right]\right)}_{\text{ice albedo}}\right], \tag{9.26}$$

$$R_o = \sigma g(T) T^4 \approx \underbrace{\sigma T^4}_{\text{radiation}} \left[1 - \underbrace{m \tanh\left(\frac{T^6}{T_0^6}\right)}_{\text{water vapor}}\right]. \tag{9.27}$$

The heat capacity of a square meter of earth is c (about $1\times10^7\,\mathrm{J\,K^{-1}\,m^{-2}}$ for the average atmosphere, about $1\times10^{10}\,\mathrm{J\,K^{-1}\,m^{-2}}$ for the average ocean). R_i is radiation input of energy from the sun–minus the reflected part which is controlled by the albedo $\alpha(T)$. $Q_0 = 342.5\,\mathrm{W\,m^{-2}}$ is average solar energy input per square meter which is $\frac{1}{4}$ of the solar constant of $1370\,\mathrm{W\,m^{-2}}$. The albedo $\alpha(T)$ depends on temperature through the ice-albedo or other feedbacks, and T is the temperature of the earth in an appropriate mean). The parameter μ can be used to describe effects such as variations in where on earth the outgoing radiation is generated, i.e., high clouds (cold) or low clouds (warm). The outgoing radiation is taken to be $R_o = \sigma g(T)T^4$, which is taken to be nearly the blackbody radiation given by the Stefan-Boltzmann Law times a "grayness" parameter (Sellers, 1969, with $\sigma = 5.67\times10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}, T_0 = 284.146\,\mathrm{K}$). The form for $g(T)$ is a simple approximation to the greenhouse effect due to the water vapor feedback, it describes how much of a "graybody" the earth is as opposed to a blackbody where $g(T) = 1$. The other dimensionless parameters will be replaced by typical constant values from Zaliapin and Ghil (2010) (see Table 9.1).

A blackbody is a physical idealization of an object that absorbs all incident radiation, and then re-emits it according to the Stephan-Boltzmann law. That law states that the amount of energy

Description	Symbol	Value
Solar constant	$4Q_0$	$1370\,\mathrm{W\,m^{-2}}$
Average insolation	Q_0	$342.5\,\mathrm{W\,m^{-2}}$
mean albedo	c_1	0.15
T sensitivity of albedo	c_2	0.7
Albedo change rate	κ	$0.2\,\mathrm{K^{-1}}$
Temp. of central albedo	T_c	$273\,\mathrm{K}$
Stephan-Boltzmann constant	σ	$5.67\times10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$
Grayness	μ	1.0
Cloudiness sensitivity	m	0.4
Temp. of central cloudiness	T_0	$284.146\,\mathrm{K}$
Heat capacity per sq. meter	c	$1\times10^{10}\,\mathrm{J\,K^{-1}\,m^{-2}}$ for full-depth ocean $1\times10^{7}\,\mathrm{J\,K^{-1}\,m^{-2}}$ for full column of atmosphere

Table 9.1: Constants of the Zaliapin and Ghil (2010) model.

per unit time per unit area emitted by a blackbody is a function of temperature only. Real objects are not as efficient at emitting radiation as a blackbody, and so a greyness parameter, or emissivity, relates what fraction of the theoretical blackbody radiation is realized. On planets with a greenhouse gas atmosphere, the emissivity is a strong function of the concentration of these gasses, hence in our earth case, $g(T)$.

This set of equations is nonlinear, so an assumed exponential form will not work, nor will the solution from Boas (2006). What can we do? Well, we can restrict our attention to small deviations in T away from an equilibrium (steady state) solution. Let's first find the steady state solutions after plugging in most of our choices of constants.

$$c\frac{\mathrm{d}T}{\mathrm{d}t} = 0 \to 0 = R_i(T) - R_o(T), \tag{9.28}$$

$$R_i = \mu Q_0(1-\alpha(T)) \approx \mu Q_0\left[0.5 + 0.35\tanh\left(\frac{0.2}{\mathrm{K}}[T-T_c]\right)\right] \tag{9.29}$$

$$R_o = \sigma g(T)T^4 \approx \sigma T^4\left[1 - 0.4\tanh\left(\frac{T^6}{T_0^6}\right)\right], \tag{9.30}$$

$$0 = \mu Q_0\left[0.5 + 0.35\tanh\left(\frac{0.2}{\mathrm{K}}[T-T_c]\right)\right] - \sigma T^4\left[1 - 0.4\tanh\left(\frac{T^6}{T_0^6}\right)\right].$$

This equation is a bear to solve, but we can do it numerically with Mathematica using Newton's method. We can see from the graph that there will be 3 relevant solutions for $\mu = 1$. They are $T \approx 174.438\,\mathrm{K}, 276.217\,\mathrm{K}$, and $297.805\,\mathrm{K}$. The mean surface temperature of the earth from 1961 to 1990 was $287.2\,\mathrm{K}$, so these numbers compare favorably (as the choice of parameters is meant to do).

For reasons we will discuss later when we get to dynamical systems, the highest temperature solution is supposed to represent our climate. So, let's expand in a Taylor series around $T_{ss1} = 297.806\,\mathrm{K}$

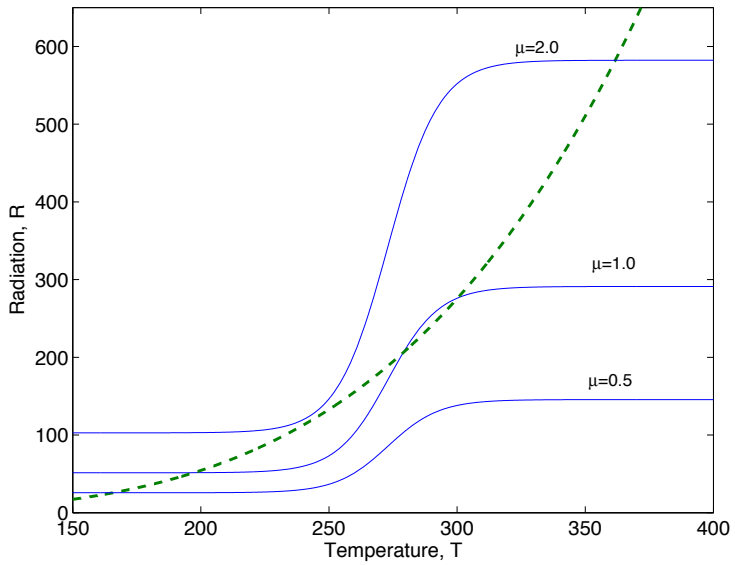


Figure 9.5: Figure from Zaliapin and Ghil (2010) showing the incoming radiation (solid lines) for some values of μ and outgoing radiation (dashed line). There are three possible steady state solutions (where $R_i = R_o$) for $\mu = 1$, and one possible for $\mu = 0.5$ or $\mu = 2.0$.

with $\mu = 1$.

$$c\frac{\mathrm{d}T}{\mathrm{d}t} = R_i(T) - R_o(T), \tag{9.31}$$

$$\begin{aligned} R_i(T) &\approx Q_0\left[0.5 + 0.35\tanh(\kappa[T - T_c])\right], \\ &\approx Q_0\left[0.5 + 0.35\tanh(\kappa[T_{ss1} - T_c])\right] + 0.35Q_0\kappa\cosh^{-2}\left(\kappa(T_{ss1} - T_c)\right)[T - T_{ss1}] + \ldots, \\ &\approx 291.113\,\mathrm{W\,m^{-2}}\left(1 \underbrace{+1.616\times10^{-5}\,\mathrm{K^{-1}}}_{\text{pos. feedback}}(T - T_{ss1}) - 3.23\times10^{-6}\,\mathrm{K^{-2}}(T - T_{ss1})^2 + \ldots\right), \end{aligned} \tag{9.32}$$

$$\begin{aligned} -R_o(T) &\approx -\sigma T^4\left[1 - 0.4\tanh\left(\frac{T^6}{T_0^6}\right)\right], \\ &\approx -\sigma T_{ss1}^4\left[1 - 0.4\tanh\left(\frac{T_{ss1}^6}{T_0^6}\right)\right] - 4T_{ss1}^3\sigma\left\{1 - 0.4\tanh\left[\frac{T_{ss1}^6}{T_0^6}\right] - \frac{0.6T_{ss1}^6}{T_0^6}\cosh^{-2}\left(\frac{T_{ss1}^6}{T_0^6}\right)\right\}(T - T_{ss1}) + \ldots, \\ &\approx 291.113\,\mathrm{W\,m^{-2}}\left(-1 \underbrace{-9.4\times10^{-3}\,\mathrm{K^{-1}}}_{\text{neg. feedback}}(T - T_{ss1}) - 7.3\times10^{-5}\,\mathrm{K^{-2}}(T - T_{ss1})^2 + \ldots\right), \end{aligned}$$

$$\frac{\mathrm{d}T}{\mathrm{d}t} \approx \underbrace{-\frac{2.73\,\mathrm{W\,m^{-2}\,K^{-1}}}{c}}_{\text{neg. feedback}}(T - T_{ss1}) + \ldots \tag{9.33}$$

For deviations in T that are less than 5 K from T_{ss1}, the first order term is the largest in R_i, so this is roughly the range where the linear equation is accurate.

OK! Now we have our linear, homogenous, constant coefficient, first-order equation (9.33). Let's

solve! Suppose the earth is warmed or cooled ΔT away from the steady state value by a temporary process (El Niño!), then

$$T - T_{ss1} \approx Ae^{-t/\tau} = \Delta T e^{-t/\tau},$$

$$\tau = \frac{c}{2.73\,\mathrm{W\,m^{-2}\,K^{-1}}} = \begin{cases} 3.6\times10^6\,\mathrm{s} = 0.12\mathrm{yr} & \text{atmosphere only} \\ 3.6\times10^9\,\mathrm{s} = 120\mathrm{yr} & \text{atmosphere and ocean} \end{cases}$$

So, we see that the anomalous flux of energy divided by the heat capacity provides the timescale for restoring back to the steady state value. Notice how the huge heat capacity of the ocean slows down the timescale over the atmosphere only. Also, notice that it doesn't matter whether the perturbation is a warming or a cooling. Finally, the imbalance per Kelvin is extremely small, under 1% of the incoming radiation–this explains why it is so hard to model global warming!

9.2.4 Feedbacks

Discussions of climate are often filled with assertions about positive and negative feedbacks. A negative feedback is one that tends to restore the system to the steady state value. In a linear, homogeneous differential equation, the feedback can be determined by the coefficient of the linear term. In the preceding example, we examined the solution near the $T_{ss1} = 297.806\,\mathrm{K}$ steady state solution. In this case, the radiation back to space had a negative feedback (the coefficient for the $(T - T_{ss1})$ term in the $\frac{\mathrm{d}T}{\mathrm{d}t}$ equation from R_o was negative). The ice-albedo feedback was positive, but it was too small to change the overall stability of T_{ss1}.

If we follow the same procedure, but for the $T_{ss2} = 276.217\,\mathrm{K}$ steady state solution, we find

$$c\frac{\mathrm{d}T}{\mathrm{d}t} = R_i(T) - R_o(T), \tag{9.34}$$

$$R_i(T) \approx 239.244\,\mathrm{W\,m^{-2}}\left(1\underbrace{+6.797\times10^{-2}\,\mathrm{K^{-1}}}_{\text{pos. feedback}}(T - T_{ss2}) + \dots\right),$$

$$-R_o(T) \approx 239.244\,\mathrm{W\,m^{-2}}\left(-1\underbrace{-9.152\times10^{-3}\,\mathrm{K^{-1}}}_{\text{neg. feedback}}(T - T_{ss2}) + \dots\right),$$

$$\frac{\mathrm{d}T}{\mathrm{d}t} \approx \underbrace{\frac{14.1\,\mathrm{W\,m^{-2}\,K^{-1}}}{c}}_{\text{pos. feedback}}(T - T_{ss2}) + \dots \tag{9.35}$$

$$T - T_{ss2} \approx \Delta T e^{t/\tau}, \qquad \tau = \frac{c}{14.1\,\mathrm{W\,m^{-2}\,K^{-1}}} \tag{9.36}$$

Thus, near this steady state, the positive ice-albedo feedback *wins* over the radiation to space negative feedback. Thus, this steady state is *unstable* to perturbations. When initially perturbed, the temperature moves away from the steady state exponentially fast, rather than exponentially decaying back to it.

For first-order, homogeneous, constant-coefficient, linear, ordinary differential equations, these are the only two possibilities–exponential decay or exponential growth.

It is also possible to have nonlinear feedback mechanisms. These may be negative or positive for small deviations and the opposite for large deviations. Indeed, the nonlinear ice-albedo feedback

here is a major reason for the number and complexity of the solutions in this model. However, in linear models (or after a nonlinear model is linearized by a Taylor series) feedbacks are simply positive or negative.

9.3 Second-Order Equations and Coupled First Order Equations

Reading: Boas (2006)[8.5].

A second-order differential equation involves more derivatives of the unknown function. We saw that a linear first-order equation with steady forcing and constant coefficients can either converge to a steady-state exponentially or diverge from it exponentially quickly. Solutions to a second-order system can also converge or diverge exponentially, or they can oscillate, or they can converge or diverge while oscillating.

9.3.1 Application: Buoyancy or Brunt-Väisälä Frequency

The vertical momentum equation for a parcel of fluid, following the motion of the parcel, in a density stratified fluid is

$$\rho \frac{\mathrm{d}w}{\mathrm{d}t} = -\frac{\partial p}{\partial z} - \rho g.$$

Where ρ is density, w is vertical velocity, p is pressure, and g is gravitational acceleration. The hydrostatic balance provides the pressure that results in a motionless fluid, which is just the steady-state balance of this equation,

$$\frac{\partial p}{\partial z} = -\rho g.$$

What if we displace the fluid parcel a distance Δz away from the location where it would be motionless? Then, we expect it to experience buoyant forces–it it is more dense than its surroundings then it will sink, and if less dense then it will rise. This is most easily studied if the pressure before displacement is taken to be hydrostatic and unaffected by the displacement, and if the displaced parcel maintains its original density $\rho(z)$. Then, the parcel of density $\rho(z)$ finds itself located at $z + \Delta z$ where the pressure is based on a different density,

$$\frac{\partial p(z+\Delta z)}{\partial z} = -\rho(z+\Delta z)g,$$
$$\rho \frac{\mathrm{d}w}{\mathrm{d}t} = -\frac{\partial p(z+\Delta z)}{\partial z} - \rho(z)g = \rho(z+\Delta z)g - \rho(z)g.$$

If the displacement of the parcel is Δz, then the vertical velocity will be related to the variations in time of this quantity, or $\frac{\mathrm{d}^2\Delta z}{\mathrm{d}t^2} = \frac{\mathrm{d}w}{\mathrm{d}t}$. The density difference can be related to the displacement by a Taylor series: $\rho(z+\Delta z) - \rho(z) = \dfrac{\partial \rho}{\partial z}\Delta z + \ldots$. Thus, for small displacements,

$$\frac{\mathrm{d}^2\Delta z}{\mathrm{d}t^2} \approx \frac{g}{\rho}\frac{\partial \rho}{\partial z}\Delta z \equiv -N^2 \Delta z. \tag{9.37}$$

Where N, the buoyancy or Brunt-Väisälä frequency, is defined to be $\sqrt{\frac{-g}{\rho_0}\frac{\partial \rho}{\partial z}}$. Let us suppose that this quantity is constant.

Then (9.37) is a second-order, linear, homogeneous differential equation with constant coefficients. We again assume a solution of the form,

$$\Delta z = Ae^{t/\tau}.$$

But, when we plug this in, we find 2 possible general solutions!

$$\Delta z = Ae^{iNt}, Be^{-iNt}.$$

Or equivalently,

$$\begin{aligned} \Delta z &= (A+B)\cos(Nt) + i(A-B)\sin(Nt), \\ \Delta z\Big|_{t=0} &= A+B. \end{aligned}$$

If we take $2H$ as the initial displacement height at $t=0$, we see that $A+B=2H$, but what about $A-B$? How do we find the value of the difference? We need to find another initial condition.

If we assume that the displacement occurs such that the velocity is 0 at $t=0$ when $\Delta z = 2H$, then

$$\begin{aligned} \Delta z &= 2H\cos(Nt) + i(A-B)\sin(Nt), \\ \tfrac{\mathrm{d}\Delta z}{\mathrm{d}t}\Big|_{t=0} &= w\Big|_{t=0} = 0 = i(A-B), \\ \Delta z &= 2H\cos(Nt) = He^{iNt} + He^{-iNt}. \end{aligned}$$

Note that you could go wrong by forgetting to look for the second general solution, and it would only be at the end of the problem when matching initial conditions that the error would be revealed.

Second order differential equations can indeed result in oscillations (describable by either sines and cosines or a pair of complex exponentials). This may remind you of Fourier series, which as we'll see is not coincidental!

An Alternative–Two Coupled First-Order Equations

What if we had stuck with w instead of replacing it with $\frac{\mathrm{d}^2\Delta z}{\mathrm{d}t^2}$? Then we could have included the equation for w, as in

$$\begin{aligned} \frac{\mathrm{d}w}{\mathrm{d}t} &\approx -N^2\Delta z, \\ \frac{\mathrm{d}\Delta z}{\mathrm{d}t} &= w. \end{aligned} \tag{9.38}$$

We know these two first-order equations are exactly equivalent to the one second-order equation in (9.37), but don't we already know that first-order equations only decay or grow exponentially? This

set of first-order equations and the second-order equation (9.37) *oscillate*. It is true that first-order equations standing alone with constant coefficients do not oscillate. However, here we have *coupled* first order equations.

Coupling is the term for when the variable that appears inside the derivative (here w and Δz, respectively) appear as a coefficient or parameter in the equation for the other variable. So, the w equation here is not an exponential decay or growth because the Δz appears on the right-hand side and it is *not a constant*! Likewise, w is not a constant and it appears in the z equation. In fact, in general any higher-order equation can always be written as a set of coupled first-order equations, with new variables being added (as w was added here).

9.3.2 Application: Chemical Reaction Kinetics

In a closed system at constant volume, the chemical reaction

$$xX + yY \to pP + qQ,$$

occurs at a rate determined by the concentration of each of the reactants ($[X], [Y]$) and products ($[P], [Q]$) based on their stoichiometric coefficients x, y, p, q. The reaction rate is

$$r = \frac{-1}{x}\frac{\mathrm{d}[X]}{\mathrm{d}t} = \frac{-1}{y}\frac{\mathrm{d}[Y]}{\mathrm{d}t} = \frac{1}{p}\frac{\mathrm{d}[P]}{\mathrm{d}t} = \frac{1}{q}\frac{\mathrm{d}[Q]}{\mathrm{d}t} \tag{9.39}$$

Often a reaction rate depends on temperature T, or other factors, as well as the concentration of the reactants, according to

$$r = k(T)[X]^m[Y]^n, \tag{9.40}$$

where m, n are the order of reaction for X and Y. If the reaction is simple (single-stage), then $m = x$ and $n = y$, but often reactions involve intermediate products that limit the rate of the overall reaction, so equality between the stoichiometry coefficients and the reaction orders needs to be determined experimentally.

In a closed container, the amount of a reactant may be limited. Suppose reactant X is much less abundant than Y, then we can model the concentration change in $[X]$ by

$$\frac{-1}{x}\frac{\mathrm{d}[X]}{\mathrm{d}t} = r = k(T)[X]^m[Y]^n. \tag{9.41}$$

If $m = 1$, and the concentration $[Y]$ is constant (because Y is so much more abundant than X), and the temperature is also nearly constant, then we expect

$$\frac{-1}{x}\frac{\mathrm{d}[X]}{\mathrm{d}t} = \frac{1}{\tau}[X] \to [X] = [X]_0 e^{-t/\tau}, \tau = \left(k(T)[Y]^m\right)^{-1}. \tag{9.42}$$

Thus, we have another example involving first-order, constant-coefficient, linear ordinary differential equations.

9.4 Inhomogeneous, Constant, Linear Second-Order Equations

Now we turn to a problem where the second-order differential equation is not homogeneous.

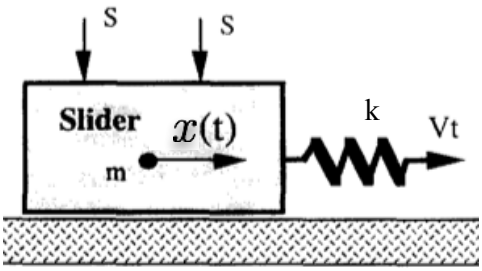


Figure 9.6: This figure schematizes the stick-slip situation (Campillo et al., 1996), the right side of the spring is drawn at constant velocity V, and the mass m responds via stretching of the spring under a static or dynamic friction proportional to the (constants) normal force S and friction coefficient μ.

9.4.1 Application: Stick-Slip

Reading: Boas (2006)[8.6].

A similar set of equations to those for buoyancy oscillations occurs if we take a solid and deformed it according to Hooke's law for displacements in an elastic medium (i.e., masses on springs). Let's consider an interesting case of Hooke's Law, the equations of stick-slip motion used to model earthquakes. The equation of motion for the stick-slip situation is

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = \left(k\left[Vt - x - D\right] - \mu S\right)\mathcal{H}\left[k\left[Vt - x - D\right] - \mu S\right]$$

Where $\mathcal{H}$ is the Heaviside function (zero when it has a negative argument and 1 when it has a positive argument). If we restrict ourselves to times when the mass is always slipping, then the Heaviside function is just 1, and

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = k\left[Vt - x - D\right] - \mu S = -kx + k(Vt - D) - \mu S.$$

Now we have two parts of the solution for x, one part that is the wavelike part (satisfying $m\frac{\mathrm{d}^2x_c}{\mathrm{d}t^2} = -kx_c$) which is the solution to the homogeneous equation (and Boas calls the complementary solution) and another part describing the dragging flow $x = (Vt - D) - \mu S/k$ which Boas calls the particular solution. We can add them together to find a general solution,

$$x = Ae^{it\sqrt{k/m}} + Be^{-it\sqrt{k/m}} + (Vt - D) - \mu S/k.$$

Note that in this case, it was fairly easy to find the particular solution, since $\frac{\mathrm{d}^2x_p}{\mathrm{d}t^2} = 0$. However, unlike in the first-order equation set, where we set the derivative to zero and looked for steady state solutions to eliminate the inhomogeneous terms, in second-order equations you can't assume that setting $\frac{\mathrm{d}^2x_p}{\mathrm{d}t^2} = 0$ is allowed to find the particular solution. Sometimes finding the particular solution (including $\frac{\mathrm{d}^2x_p}{\mathrm{d}t^2} \neq 0$) might be the hardest part!

9.5 How Many Solutions? Constants? Boundary and Initial Conditions?

For first-order equations, our general solution had one unknown and we used one initial condition (for time-based problems) or one boundary condition (if the derivative had been with respect to space). For second-order equations (or two coupled first-order equations), the general solution has two constants and two initial conditions are needed. Let's consider one final problem requiring four unknowns and a bit of linear algebra to boot!

Suppose instead of one mass on a spring dragged along in the stick-slip problem, we had two of equal mass and the second mass was dragged by a spring attached to the first (which was still dragged by a constant velocity pulling its spring). The equations of motion are now,

$$m\frac{\mathrm{d}^2x_1}{\mathrm{d}t^2} = \left(k\left[Vt - x_1 - D\right] - k\left[x_1 - x_2 - D\right] - \mu S\right)\mathcal{H}\left[k\left[Vt - x_1 - D\right] - k\left[x_1 - x_2 - D\right] - \mu S\right],$$
$$m\frac{\mathrm{d}^2x_2}{\mathrm{d}t^2} = \left(k\left[x_1 - x_2 - D\right] - \mu S\right)\mathcal{H}\left[k\left[x_1 - x_2 - D\right] - \mu S\right]$$

Again assuming that we are always sliding, not sticking,

$$m\frac{\mathrm{d}^2x_1}{\mathrm{d}t^2} = \left(k\left[Vt - x_1 - D\right] - k\left[x_1 - x_2 - D\right] - \mu S\right), \tag{9.43}$$
$$m\frac{\mathrm{d}^2x_2}{\mathrm{d}t^2} = \left(k\left[x_1 - x_2 - D\right] - \mu S\right)$$

We can write these two coupled second-order equations as a matrix equation! We can immediately split into the complementary and particular solution, just as above.

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = k\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} kVt - \mu S \\ -kD - \mu S \end{bmatrix},$$
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_{p1} \\ x_{p2} \end{bmatrix} + \begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix},$$
$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix} = k\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix}, \quad 0 = k\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_{p1} \\ x_{p2} \end{bmatrix} + \begin{bmatrix} kVt - \mu S \\ -kD - \mu S \end{bmatrix}.$$

We can solve for the particular solution by matrix inverse.

$$\begin{bmatrix} x_{p1} \\ x_{p2} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}^{-1}\begin{bmatrix} Vt - \mu S/k \\ -D - \mu S/k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} Vt - \mu S/k \\ -D - \mu S/k \end{bmatrix} = \begin{bmatrix} Vt - D - 2\mu S/k \\ Vt - 2D - 3\mu S/k \end{bmatrix}. \tag{9.44}$$

But what about the coupled mess of an equation for the complementary solution?

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix} = \frac{k}{m}\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix}.$$

Well, we've seen that similar equations can be solved by assuming a form $Ae^{i\omega t}$, so lets try that here. We will assume that *both* parts, x_{c1}, x_{c2}, of the solution have the same frequency.

$$-\omega^2 \begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix} = \frac{k}{m} \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix}, \tag{9.45}$$

$$\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix} = -\frac{\omega^2 m}{k} \begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix}. \tag{9.46}$$

We've seen an equation like that before–*it is an eigenvalue-eigenvector equation!* The complementary solution will be an eigenvector, and the eigenvalue will give us the allowable choices of the frequency!! Or, completely equivalently, this is a homogenous equation and we are seeking nontrivial solutions, thus the determinant of the coefficient matrix must vanish.

$$\begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix} = \lambda \begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix} = -\frac{\omega^2 m}{k} \begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix}, \tag{9.47}$$

$$\lambda_1 = -\frac{3}{2} - \frac{\sqrt{5}}{2}, \qquad \lambda_2 = -\frac{3}{2} + \frac{\sqrt{5}}{2}. \tag{9.48}$$

Working through the linear algebra, the solution is

$$\begin{bmatrix} x_{c1} \\ x_{c2} \end{bmatrix} = (Ae^{i\omega_1 t} + Be^{-i\omega_1 t}) \begin{bmatrix} 1+\sqrt{5} \\ -2 \end{bmatrix} + \left(Ce^{i\omega_2 t} + De^{-i\omega_2 t}\right) \begin{bmatrix} 1-\sqrt{5} \\ -2 \end{bmatrix}, \tag{9.49}$$

$$\omega_1 = \sqrt{\frac{k}{m}}\sqrt{\frac{3+\sqrt{5}}{2}}, \qquad \omega_2 = \sqrt{\frac{k}{m}}\sqrt{\frac{3-\sqrt{5}}{2}}$$

So, in this coupled second-order set, there are four constants to determine, and four frequencies (two pairs of opposites), consistent with solving the fourth-order characteristic equation for ω.

In general, this will be the case for coupled sets of equations. The sum of the orders of the equations will govern the number of constants (aside from linearly dependent redundancies), and there will often be a role for an eigenvector complementary solution with eigenvalues as the frequency. Thus, the idea that eigenvalues "characterize" the modes is borne out. Here we had a second order differential equation, which featured the frequency (eigenvalue) squared, and setting the determinant to zero for nontrivial solutions results in a fourth-order polynomial for ω. If we had written this equation as four first order equations (as in 9.38), we would have had 4×4 coefficient matrix, and thus a fourth-order eigenvalue equation.

9.6 Example Problems

9.6.1 Jargon to Argot

Example 9.1 *We are now fairly used to "linear" functions and equations. Are the solutions to linear equations always straight lines?*

No. The typical solution to a linear differential equation is an exponential or an oscillation, not a straight line at all!

Example 9.2 *Problem 8.3.1 of Boas (2006). Using (3.9), find the general solution of the following differential equation.*

$$y' + y = e^x \tag{9.50}$$

Hint: See comments just after (3.9), and Example 1.

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} + y &= e^x, \\
I &= \int 1 \,\mathrm{d}x = x, \\
ye^x &= \int e^x e^x \,\mathrm{d}x + c, \\
ye^x &= \frac{1}{2}e^{2x} + c, \\
y &= \frac{1}{2}e^x + ce^{-x}.
\end{aligned}$$

Note that no constant of integration is needed in I.

Example 9.3 *Problem 8.5.3 of Boas (2006). Solve the following differential equation by the methods discussed in Boas. You do not need to compare to a computer solution.*

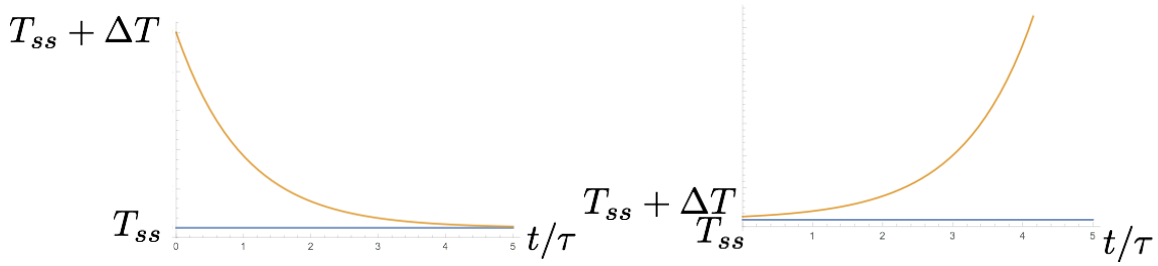
$$y'' + 9y = 0 \tag{9.51}$$

$$\begin{aligned}
\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 9y &= 0, \\
y &= Ae^{-3ix} + Be^{3ix}.
\end{aligned}$$

9.6.2 Scheming Schematics and Articulate Analysis

Example 9.4 *a) Plot trajectories for T that converge toward or diverge away from the three steady solutions of the Energy Balance Model of Section 9.2.3. b) Plot a timeseries of x_1 and x_2 from two different initial conditions, with each chosen to exemplify one of the two frequency pairs inherent in (9.49). That is, use $A = B; C = D = 0$ for one time series set, and $A = B = 0; C = D$ for the other.*

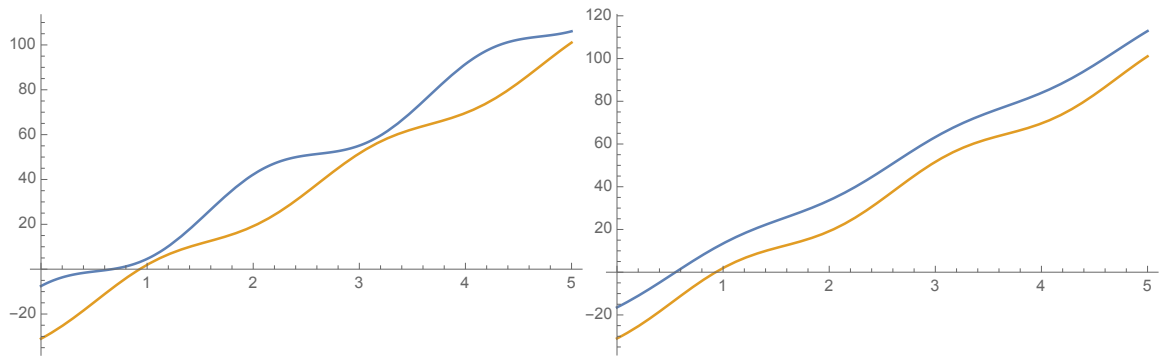
In the EBM, the solutions all either exponentially converge or exponentially diverge. Here are examples,



The solutions are

$$x_1 = Vt - D - 2\mu S/k + \left[1+\sqrt{5}\right](Ae^{i\omega_1 t} + Be^{-i\omega_1 t}) + \left[1-\sqrt{5}\right]\left(Ce^{i\omega_2 t} + De^{-i\omega_2 t}\right),$$
$$x_2 = Vt - 2D - 3\mu S/k - 2(Ae^{i\omega_1 t} + Be^{-i\omega_1 t}) - 2\left(Ce^{i\omega_2 t} + De^{-i\omega_2 t}\right).$$

Choosing arbitrary values for separation (D), rate of pulling (V), and k, m, we find the two example trajectories for the $A = B, \omega_1$ mode (left) and the $C = D, \omega_2$ mode (right).



A few things to note. 1) Both positions are constantly moving, to avoid the sticking in the stick-slip law. 2) The frequencies are different, with $\omega_1 > \omega_2$. 3) The oscillations in the ω_1 mode are in opposition, while those in the ω_2 mode covary (up together, down together). The average velocity of both masses is the same and is just V.

9.7 Homework Problems

9.7.1 Manipulation

Exercise 9.1 *Problem 8.2.1 of Boas (2006). For the following differential equation, separate variables and find a solution containing one arbitrary constant. Then find the value of the constant to give a particular solution satisfying the given boundary condition.*

$$xy' = y \tag{9.52}$$

You do not need to compare to a computer solution.

Exercise 9.2 *Problem 8.5.1 of Boas (2006). Solve the following differential equation by the methods discussed in Boas. You do not need to compare to a computer solution.*

$$y'' + y' - 2y = 0 \tag{9.53}$$

9.7.2 Application

Exercise 9.3 *Show that the general solution to the first-order, linear differential equation on pg. 401 of (Boas, 2006) is the same as the guess & check solution in (9.24) of these notes.*

Exercise 9.4 *Find the leading order differential equation, by Taylor series expansion of the radiation forcing, to the Energy Balance Model of Section 9.2.3 for the remaining steady state solution (near $T_{ss3} = 174.438\,\mathrm{K}$). Solve this differential equation, and decide if this third steady state is a stable or unstable steady state solution.*

9.7.3 Evaluate & Create

Exercise 9.5 *Verify that the sum of the characteristic solution (9.49) and the particular solution (9.44) constitute a solution to the original equations for the two sliding masses (9.43).*

Chapter 10

Dynamical Systems and Chaos

10.1 Introduction–Describing, Not Solving

A dynamical system involves three parts: 1) a set of variables describing the system state, 2) a set of times when the states can be observed, and 3) a set of rules for evolving the state at one time until the next. Newton's laws are an example of a dynamical system, but more modern applications include population mathematics, biology, and many geophysical processes.

As computers have become more powerful and software has become easier to use, the modeling of dynamical systems has gone from a niche activity of applied mathematics and science to a common method to study simplified models of complex natural systems. The approach of many scientists when confronting complex dynamical systems has been to categorize the solutions as is possible, but to rely on computers to provide the solutions to the equations. Thus, the theory of dynamical systems often goes toward categorizing, sorting, and understanding rather than toward solving.

Throughout this chapter, I will assign optional reading from Strogatz (1994) for those of you who would like to learn more.

10.2 Basic Notions–Chaos, Fractals, Dynamics

Optional Reading: (Strogatz, 1994, 1)

This categorization approach is particularly important when considering systems that are potentially chaotic. Strogatz (1994) defines chaos as the exhibition of aperiodic behavior of a deterministic system that depends sensitively on the initial conditions, thereby rendering long-term prediction impossible. The "discovery" of chaos is often attributed to Lorenz (1963), because of his clear exposition of the issues at hand and the implications, particularly of the sensitivity to initial conditions, for dynamical problems such as weather forecasts. Earlier work by Poincaré set the stage for descriptions of behavior, rather than solution techniques. We now believe that chaos fundamentally limits our ability to predict the weather beyond about two weeks, even if we had much more powerful computers and much more accurate and complete measurements of the present weather state.

While our agenda here is to move enough in the direction of chaotic dynamics to understand the implications for geophysics, we will also revisit the dynamical systems interpretations of our now-familiar solutions to linear ODEs, to help learn the tools.

10.3 1d Flows

The first kind of system we will study is a one-dimensional unbounded ODE without time-dependence in any parameters or forcing, which Strogatz (1994) calls flows on the line. The system can be written generally as

$$\dot{x} = f(x; r) \tag{10.1}$$

The dot is a shorthand for ordinary time derivative ($\frac{\mathrm{d}x}{\mathrm{d}t} = \dot{x}$), and x is the state of our one-dimensional system. A parameter r (or more generally a list of parameters) is given to allow a tunable family of related functions f, but it is given after the semicolon to make it clear that these parameters are not part of the state, but rather part of the function f.

We will consider *trajectories* through *phase space*. A trajectory is a time history from a chosen initial condition forward in time. Phase space is a plot of the state (and possibly also time derivatives of the state), where a point represents an instant in time and a line represents a trajectory.

In 1d, we can plot both $\dot{x}$ and x on the same graph. Here are some examples. In Fig. 10.1 two one-

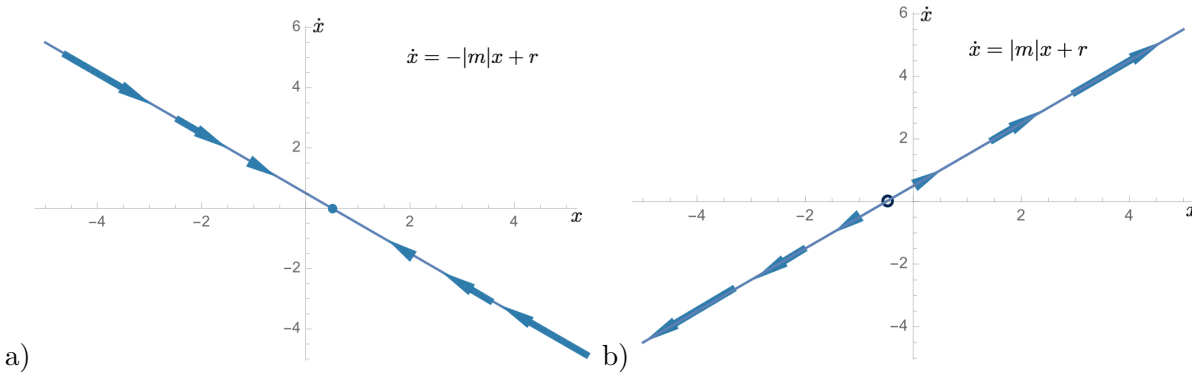


Figure 10.1: The two types of linear, one-dimensional dynamical systems. Note that the steady state solutions (or fixed points) are stable (closed circle) or unstable (open circle) depending on the sign of the slope.

dimensional linear dynamical systems are plotted. The vertical axis reflects $\dot{x}$ while the horizontal axis reflects x. All trajectories must lie on the line of the function if they are to obey the dynamics of the system. A negative slope leads to a stable fixed point, or negative feedback. A positive slope leads to an unstable fixed point or positive feedback. Arrows indicate the direction of typical trajectories, and the size of the arrows indicates rate of motion along the trajectory.

In this case, we can easily solve the set of equations as we have done many times for linear

ODEs.

$$x = (x_0 - r/|m|)e^{-|m|t} + r/|m|, \tag{10.2}$$

$$x = (x_0 + r/|m|)e^{|m|t} - r/|m|, \tag{10.3}$$

$$\tag{10.4}$$

The initial value of x is x_0, and the steady-state solutions (also known as fixed points) are $x^* = r/|m|$ for Fig. 10.1a and $x^* = -r/|m|$ for Fig. 10.1b. As we saw in the solutions to the energy balance model in Section 9.2.3, positive and negative feedbacks (and thus unstable and stable fixed steady solutions) result from the slope of the function near the steady-state solutions. Even if the function is not linear, as long as it is continuous and differentiable, we can perform a Taylor series near the steady-state solution and determine stability based on the first-order coefficients in the series.

Fig. 10.2 shows some nonlinear dynamical system trajectory examples. In Fig. 10.2a, we see that

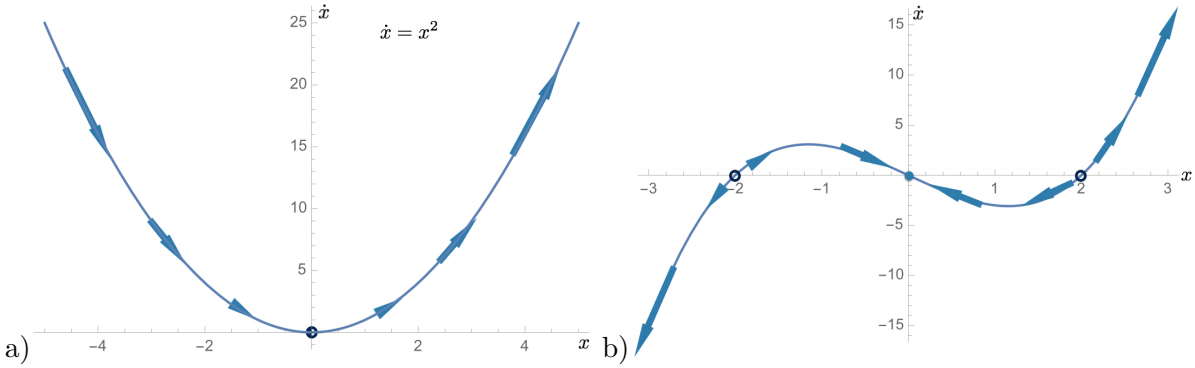


Figure 10.2: Examples of nonlinear, one-dimensional dynamical systems.

there is another type of steady solution–a half-stable one. It attracts trajectories from the left, but repels ones from the right. In Fig. 10.2b, we see that a nonlinear system can have more than one steady state solution, and that they may vary in stability. This should remind you of the solutions to the energy balance model in Section 9.2.3.

Example: Logistic Growth

In population dynamics, a simple model for the population of organisms is

$$\dot{N} = rN. \tag{10.5}$$

We know the solution to this equation is just exponential growth at the rate r. However, in more realistic situations, there is a limit for how many organisms can coexist in the habitat, called the *carrying capacity*, K. The *logistic growth* model captures this effect.

$$\dot{N} = rN\left(1 - \frac{N}{K}\right). \tag{10.6}$$

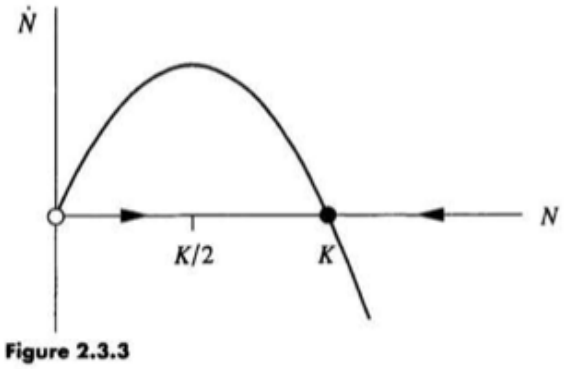


Figure 10.3: Phase space plot of logistic growth (10.6). From Strogatz (1994).

10.3.1 Bifurcations

In Fig. 10.1, the parameters carried along with the system (m, r) did not play much of a role. However, if we studied instead $\dot{x} = mx + b$, then the sign of m would determine the stability! Often is it useful to study a range of the parameters in the system, so that their effects on the kinds of solutions that are possible can be better understood. Here are some examples.

Saddle-point or Blue-Sky Bifurcation

A classic bifurcation occurs in the model

$$\dot{x} = r + x^2, \tag{10.7}$$

when the parameter r is varied. Fig. 10.4 shows the bifurcation that occurs as r passes through zero. For negative r, there are two steady state solutions (one stable, one unstable), then at $r = 0$ there is a single half-stable steady state, and finally for $r > 0$ there are no steady solutions. This

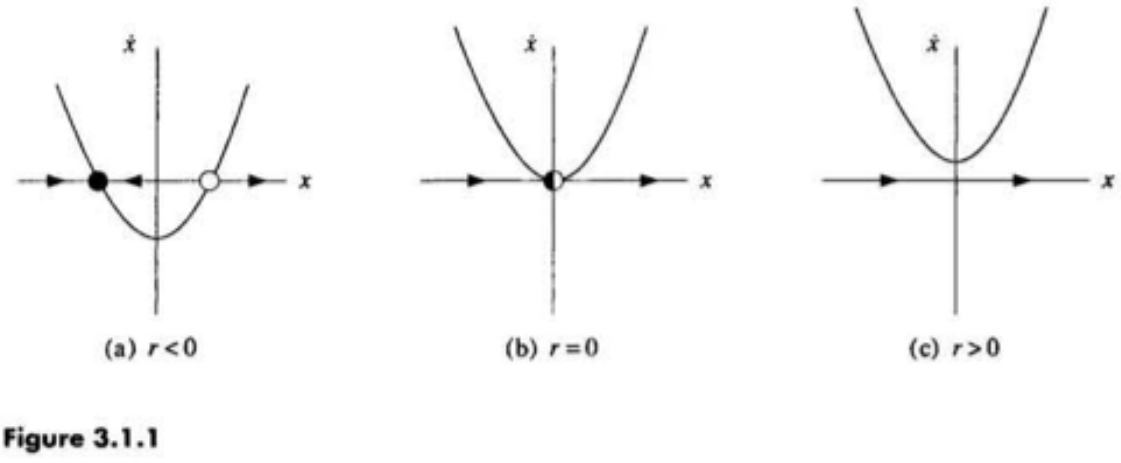


Figure 10.4: Phase space plots showing a saddle node bifurcation under variation of the parameter r in (10.7). From Strogatz (1994).

situation can also be summarized with a *bifurcation diagram*, where now the steady state solutions of x are plotted on a figure where the parameter r is varied (Fig. 10.7). This situation explains

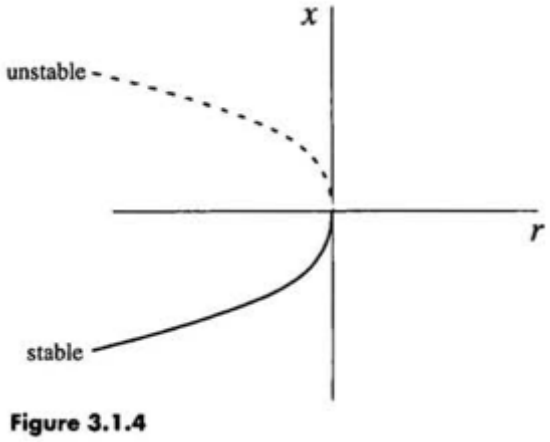


Figure 10.5: Bifurcation diagram versus r in (10.7). From Strogatz (1994).

why the word "bifurcation" is used, as in this case, the solution splits into two branches.

Transcritical Bifurcation

A different kind of bifurcation occurs in the system

$$\dot{x} = rx - x^2, \tag{10.8}$$

Fig. 10.6 shows the bifurcation that occurs as r passes through zero. For negative r, there are two steady state solutions (one stable, one unstable), then at $r = 0$ there is a single half-stable steady state, and finally for $r > 0$ there are two steady solutions–one stable and one unstable, but they have exchanged stability. Now both steady state solutions of x persist as r varies, but the stability

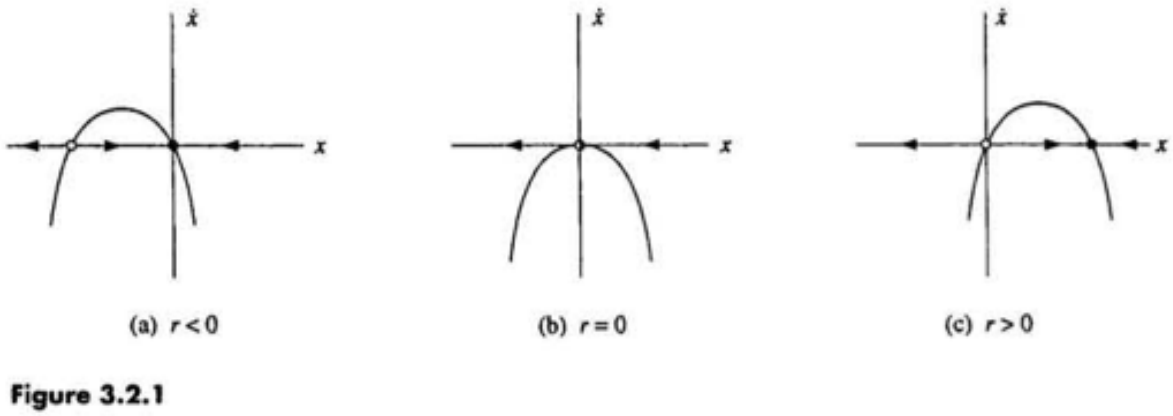


Figure 10.6: Phase space plots showing a saddle node bifurcation under variation of the parameter r in (10.8). From Strogatz (1994).

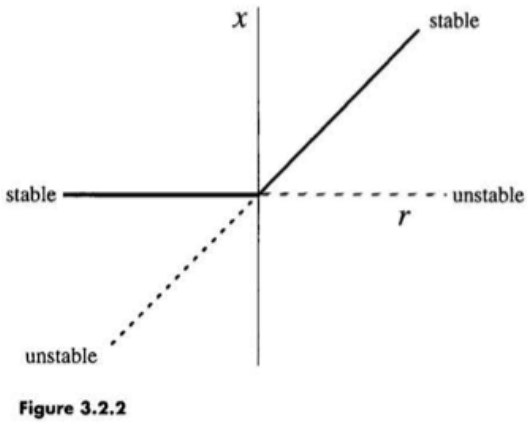


Figure 10.7: Bifurcation diagram versus r in (10.8). From Strogatz (1994).

of the two changes as r changes sign (Fig. 10.6). Now the word "bifurcation" is somewhat of a misnomer, as in this case no solutions split in two. However, the change in the qualitative behavior of the solutions is still referred to as a bifurcation.

Pitchfork Bifurcation

A different kind of bifurcation occurs in the system

$$\dot{x} = rx - x^3, \tag{10.9}$$

Fig. 10.6 shows the bifurcation that occurs as r passes through zero. For negative r, there is one stable steady state solution, then at $r = 0$ there is a single stable steady state, and finally for $r > 0$ there are three steady solutions–two stable, one unstable. Now the one steady state solution of

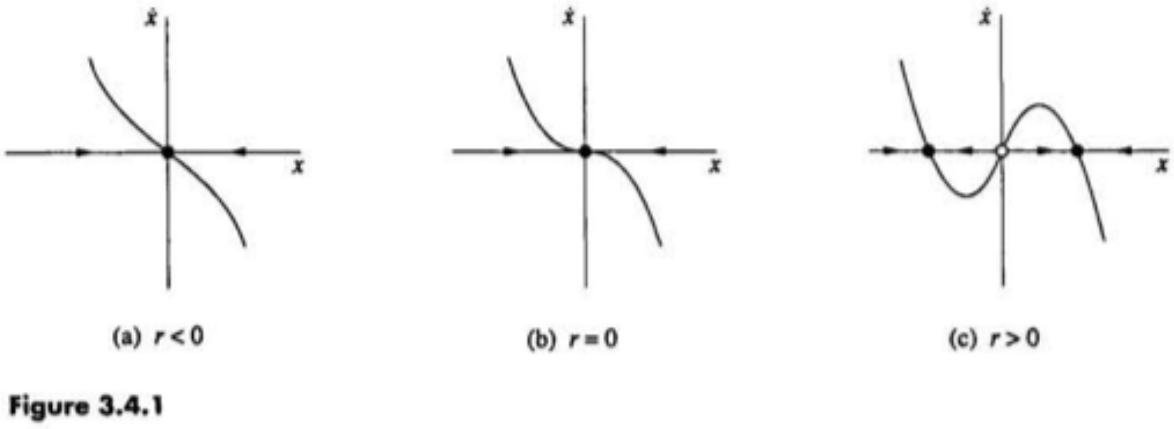


Figure 10.8: Phase space plots showing a pitchfork bifurcation under variation of the parameter r in (10.9). From Strogatz (1994).

x when $r < 0$ transforms into an unstable solution, while two new stable solutions appear at the

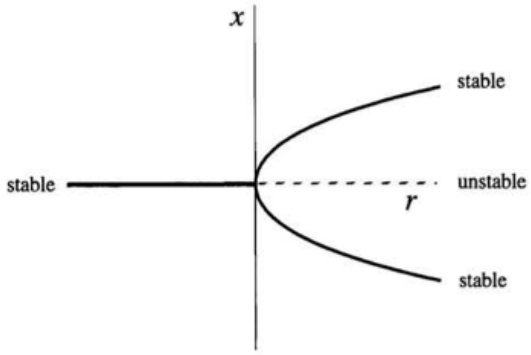


Figure 3.4.2

Figure 10.9: Bifurcation diagram versus r in (10.9). From Strogatz (1994).

same time as r changes sign (Fig. 10.8). Again the word "bifurcation" is somewhat of a misnomer, as in this case the solutions split in three instead of in two. The reason for the name pitchfork is clear from Fig. 10.9.

Other

We note that all of these systems resulted from low-order polynomials whose roots changed character as the parameters were varied. Because we can also perform a Taylor series of any continuous function near the steady states, and retain just enough polynomial order to keep the basic bifurcation structures, studying the bifurcations of polynomials serves as a cookbook for more complex functions. Here are a few more interesting examples to consider.

A different kind of subcritical pitchfork bifurcation occurs in the system

$$\dot{x} = rx + x^3 - x^5. \tag{10.10}$$

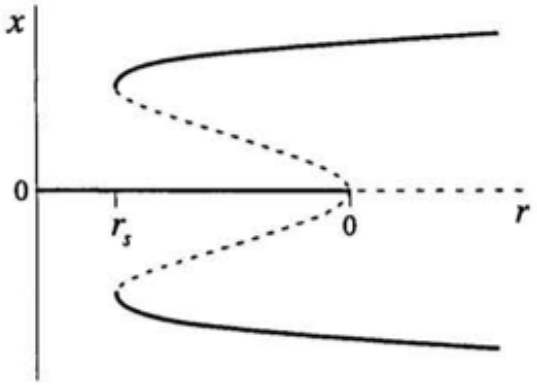


Figure 3.4.7

Figure 10.10: Subcritical pitchfork bifurcation diagram versus r in (10.10). From Strogatz (1994).

$$\dot{x} = h + rx - x^3. \tag{10.11}$$

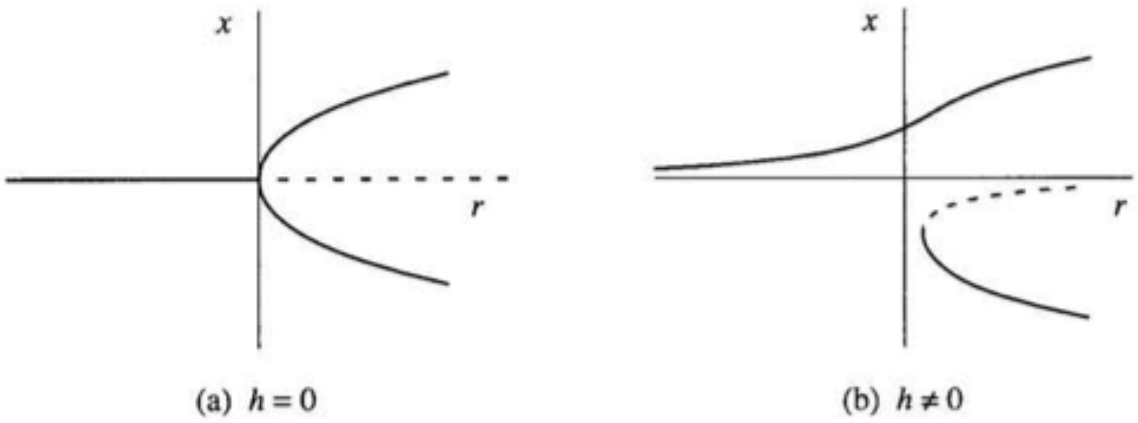


Figure 3.6.3

Figure 10.11: Imperfect pitchfork bifurcation diagram versus r for two values of h in (10.11). From Strogatz (1994).

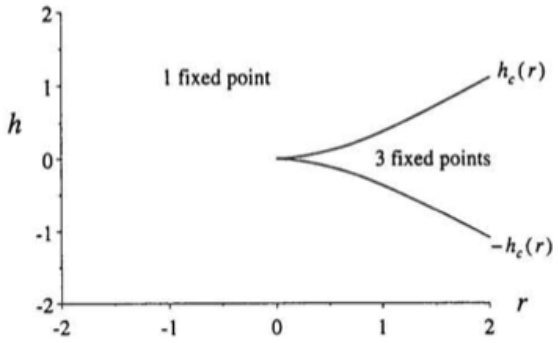


Figure 3.6.2

Figure 10.12: Imperfect pitchfork regime diagram with solution types r, h plane for (10.11). From Strogatz (1994).

Despite all of the complexity inherent in these one-dimensional systems, there are still relevant geophysical behaviors missing. For example, near the steady-state solutions, the approach or departure from those solutions was always exponential. We know that in the real world there are many oscillations about a steady solution (waves on the ocean, Brunt-Väisälä oscillations (Section 9.3.1), seismic waves, stick-slip oscillations, etc.).

10.4 2d Flows

In one-dimensional systems, there was no way to overshoot a steady-state solution, but in two-dimensional systems, you can just "go around."

One good geophysical linear example are the Brunt-Väisälä oscillations of a stratified fluid. We found the governing equation (9.37), which is

$$\frac{\mathrm{d}^2\Delta z}{\mathrm{d}t^2} \approx -N^2\Delta z. \tag{10.12}$$

Rewriting this system in our dynamical systems version, we can express it as,

$$\dot{v} = -N^2 z, \tag{10.13}$$

$$\dot{z} = v. \tag{10.14}$$

Note that we have changed the second derivatives into two first derivatives. The nature of the oscillation can be seen by considering local deviations, or the *phase portrait* of the closed orbits exhibited by this system (Fig. 10.13).

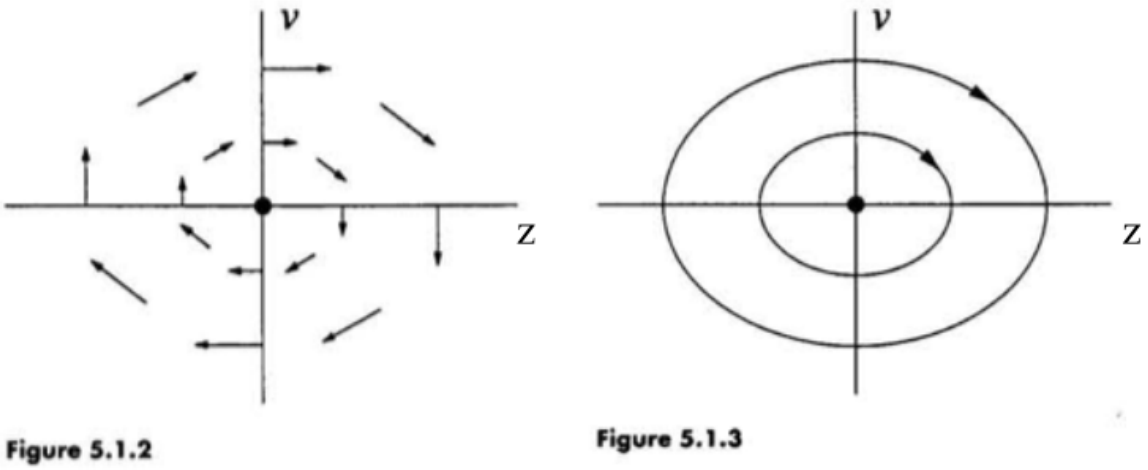


Figure 10.13: Local trajectories (left) and phase portrait for the Brunt-Väisälä oscillation system (10.13). Adapted from Strogatz (1994).

In Section 9.3.1, we noted that (10.12) can be solved by a guess and check solution with an exponential

$$\Delta z = Ae^{t/\tau}.$$

But, when we plugged this in, we found 2 possible general solutions!

$$\Delta z = Ae^{iNt}, Be^{-iNt}.$$

Alternatively, if we plug into (10.13) with the assumed form,

$$z = Ze^{t/\tau},$$

$$v = Ve^{t/\tau}.$$

We find (after dividing through by $e^{t/\tau}$),

$$\frac{V}{\tau} = -N^2 Z,$$
$$\frac{Z}{\tau} = V.$$

This is a set of linear, homogeneous equations. If we want there to be oscillations, then we want a nontrivial solution, which means we can rewrite as a matrix equation and set the determinant to zero.

$$\begin{bmatrix} N^2 & \frac{1}{\tau} \\ \frac{1}{\tau} & -1 \end{bmatrix} \begin{bmatrix} Z \\ V \end{bmatrix} = 0,$$
$$N^2 = -\frac{1}{\tau^2} \to \frac{1}{\tau} = \pm iN.$$

Which is the same result.

10.4.1 Linear Systems

So then, let's consider the general linear, two-dimensional system.

$$\dot{x} = a(x - x^*) + b(y - y^*), \tag{10.15}$$
$$\dot{y} = c(x - x^*) + d(y - y^*), \tag{10.16}$$

where x^*, y^* gives the one possible steady state solution and a, b, c, d are constants. We can write this system as a matrix equation.

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix}. \tag{10.17}$$

If determine the eigenvalues and eigenvectors of this coefficient matrix (Section 3.8), then we will be able to solve this problem generally. The eigenvalues of the matrix can be real or complex, and their values will determine the type of oscillation or exponential variations that occur near the steady solution at $x = x^*, y = y^*$.

Let us proceed by diagonalizing the matrix following (3.44). The eigenvectors $\mathbf{v}_i$–scaled to have length 1–are the columns of the matrix $\mathbf{C}$ (and therefore the rows of the matrix $\mathbf{C}^{-1} = \mathbf{C}^\intercal$), and we take m_i as the eigenvalues which are the diagonal elements of the matrix $\mathbf{D}$. Then,

$$\mathbf{M} \equiv \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{10.18}$$
$$\mathbf{M} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1}, \tag{10.19}$$
$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix}, \tag{10.20}$$
$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{C}^{-1} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix} = \mathbf{D}\mathbf{C}^{-1} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix}. \tag{10.21}$$

This last equation is actually two separate, uncoupled equations, if we take

$$\frac{\mathrm{d}}{\mathrm{d}t}\, v_1 \cdot (\mathbf{x} - \mathbf{x}^*) = m_1 \mathbf{v}_1 \cdot (\mathbf{x} - \mathbf{x}^*), \tag{10.22}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\, v_2 \cdot (\mathbf{x} - \mathbf{x}^*) = m_2 \mathbf{v}_2 \cdot (\mathbf{x} - \mathbf{x}^*), \tag{10.23}$$

where $\mathbf{x} - \mathbf{x}^* = (x - x^*, y - y^*)$. Since the equations are uncoupled, we can solve them each separately (as our favorites, exponentials), and then combine them back together to find:

$$(\mathbf{x} - \mathbf{x}^*) = c_1 \mathbf{v}_1 e^{m_1 t} + c_2 \mathbf{v}_2 e^{m_2 t}. \tag{10.24}$$

For the vast majority of cases, this will be the solution to the linear system near the steady state solution $(\mathbf{x} - \mathbf{x}^*)$. All that is needed is to determine the constants c_1, c_2 based on the initial conditions (and the eigenvectors and eigenvalues if they are not given). Note that the "trivial solution" $c_1 = c_2 = 0$ is actually meaningful in this case, it just means that you are already located at the steady solution.

Let us see some examples of this method in action.

Two Real Eigenvalues

Consider the following system

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix}.$$

All of the coefficients, variables, and derivatives are real. Let's see about the eigenvalues! We have a linear, constant coefficient, homogeneous set of equations, so let's assume an exponential form.

$$\begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix} e^{mt},$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix} = m \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix},$$

$$m \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}.$$

If we are not exactly at the steady state solution, then we seek a nontrivial solution $(A \neq 0, B \neq 0)$. In order to find such a solution, we combine the two sides of the equation.

$$m \begin{bmatrix} A \\ B \end{bmatrix} = m \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix},$$

$$0 = \begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} - \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 1 - m & 1 \\ 4 & -2 - m \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}.$$

To avoid Cramer's rule, we set the determinant of the matrix to zero, which gives the characteristic or eigenequation,

$$(1 - m)(-2 - m) - 4 = m^2 + m + 6 = 0, \rightarrow m_1 = -3, m_2 = 2.$$

To find the corresponding eigenvectors, we solve

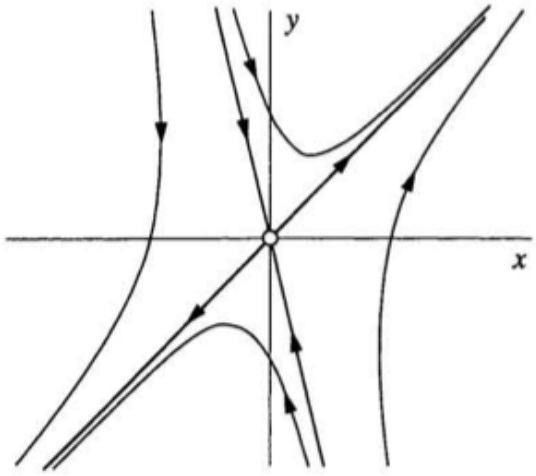
$$\begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix} \mathbf{v}_1 = m_1 \mathbf{v}_1 \rightarrow \mathbf{v}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix},$$
$$\begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix} \mathbf{v}_2 = m_2 \mathbf{v}_2 \rightarrow \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Thus, the solution is,

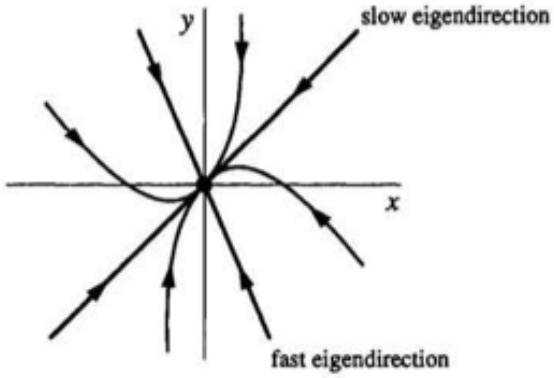
$$(\mathbf{x} - \mathbf{x}^*) = c_1 \begin{bmatrix} 1 \\ -4 \end{bmatrix} e^{-3t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{2t}. \tag{10.25}$$

The constants c_1, c_2 depend on the initial conditions.

Examine the nature of this solution. Any projection of the initial conditions along the first eigenvector will exponentially decay with time. Any projection of the initial conditions on the second eigenvector will grow exponentially in time. So, we can make a map of trajectories in phase space like Fig. 10.14a. Fig. 10.14b shows a different system for comparison, with two negative eigenvalues, one faster than the other.



a) Figure 5.2.2



b) Figure 5.2.3

Figure 10.14: a) Trajectory in phase space for (10.25). b) Trajectory in phase space for a different system. From Strogatz (1994).

Two Imaginary Eigenvalues

Consider the following system

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix}.$$

We follow the same routine as above, and find that assuming an exponential solution leads to:

$$m^2 + 1 = 0, \rightarrow m_1 = i, m_2 = -i. \tag{10.26}$$

So, this time we find that even though all of the coefficients, variables, and derivatives are real, the *eigenvalues* are not. This result can only occur in sets of two or more first-order ODEs. If the system is defined by only one ODE, then the eigenvalues must be real if all of the variables are real. We have already seen one example of complex eigenvalues in the Brunt-Väisälä oscillations studied above. Let's examine this case.

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{v}_1 = m_1 \mathbf{v}_1 \rightarrow \mathbf{v}_1 = \begin{bmatrix} -i \\ 1 \end{bmatrix},$$
$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{v}_2 = m_2 \mathbf{v}_2 \rightarrow \mathbf{v}_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Thus, the solution is of the form

$$(\mathbf{x} - \mathbf{x}^*) = c_1 \begin{bmatrix} -i \\ 1 \end{bmatrix} e^{it} + c_2 \begin{bmatrix} i \\ 1 \end{bmatrix} e^{-it}, \tag{10.27}$$
$$= c_1 \begin{bmatrix} -i \\ 1 \end{bmatrix} (\cos t + i \sin t) + c_2 \begin{bmatrix} i \\ 1 \end{bmatrix} (\cos t - i \sin t). \tag{10.28}$$

We can guarantee that $(\mathbf{x} - \mathbf{x}^*)$ will be real by setting $c_1 - c_2 = ia$ and $c_1 + c_2 = b$ where a, b are real. This makes the solution,

$$(\mathbf{x} - \mathbf{x}^*) = \begin{bmatrix} a \cos t + b \sin t \\ b \cos t - a \sin t \end{bmatrix}. \tag{10.29}$$

Despite the fact that these are real, the eigenvalues remain imaginary. This system has the same trajectories as in Fig. 10.13, but with $x - x^*, y - y^*$ as the coordinates.

Other Combinations

The form of the linear solution (10.15) admits many flavors of solutions in phase space (Fig. 10.16). The kinds of patterns can be assessed by examining the trace trace τ and determinant Δ of the coefficient matrix, because these quantities are so closely related to the eigenvalues.

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} (x - x^*) \\ (y - y^*) \end{bmatrix}, \tag{10.30}$$
$$\tau = a + d, \qquad \Delta = ad - bc, \tag{10.31}$$
$$\tau = m_1 + m_2, \qquad \Delta = m_1 m_2, \tag{10.32}$$
$$m_{1,2} = \frac{\tau \pm \sqrt{\tau^2 - 4\Delta}}{2}. \tag{10.33}$$

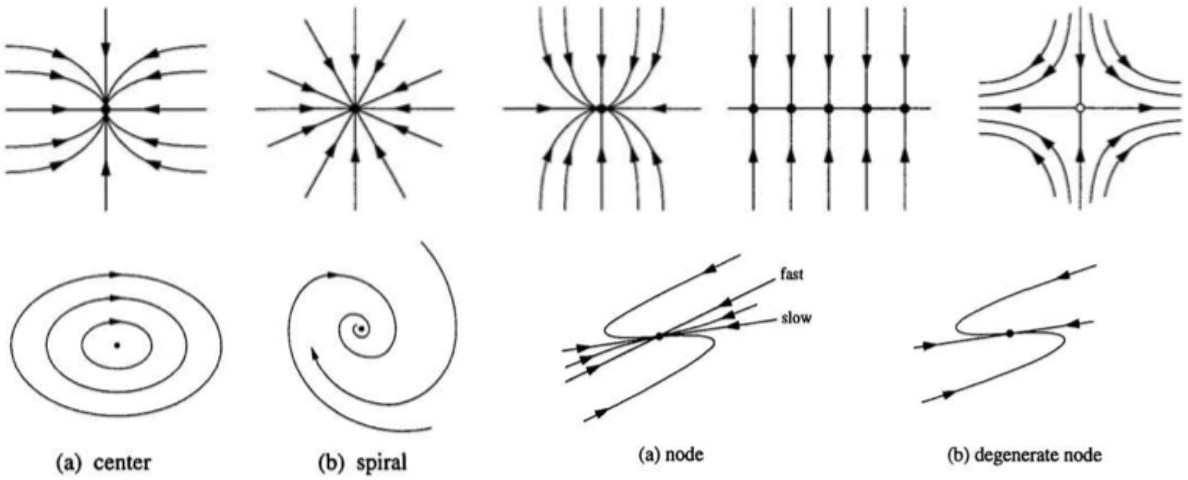


Figure 10.15: Example trajectories in phase space for linear two-dimensional systems. From Strogatz (1994).

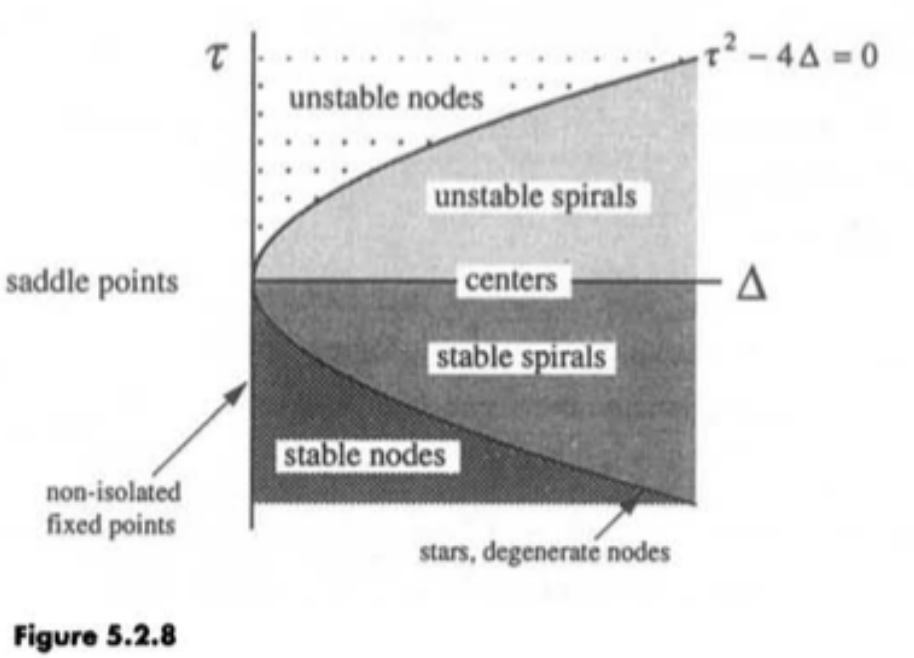


Figure 10.16: Regime diagram of the kinds of steady-state solutions based on the trace ($\tau = m_1 + m_2$) and determinant ($\Delta = m_1 m_2$) of the coefficient matrix, which are closely related to the characteristic equation and its eigenvalues (m_1, m_2). From Strogatz (1994).

10.4.2 Taylor Series Near a Steady Solution

Let us consider the Taylor series of a two-dimensional nonlinear system near a steady solution truncated to the linear terms only. It can be written as a matrix equation with a bit of shorthand,

$$\mathcal{F}_x \equiv \left.\frac{\partial \mathcal{F}}{\partial x}\right|_{\mathbf{x}=\mathbf{x}^*}, \mathcal{F}_y \equiv \left.\frac{\partial \mathcal{F}}{\partial y}\right|_{\mathbf{x}=\mathbf{x}^*}, \mathcal{G}_x \equiv \left.\frac{\partial \mathcal{G}}{\partial x}\right|_{\mathbf{x}=\mathbf{x}^*}, \mathcal{G}_y \equiv \left.\frac{\partial \mathcal{G}}{\partial y}\right|_{\mathbf{x}=\mathbf{x}^*}.$$

Then, the nonlinear system may be approximated as,

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}(x-x^*)\\(y-y^*)\end{bmatrix} = \begin{bmatrix}\mathcal{F}(x,y;r)\\\mathcal{G}(x,y;r)\end{bmatrix}, \tag{10.34}$$

$$\approx \begin{bmatrix}\mathcal{F}_x & \mathcal{F}_y\\\mathcal{G}_x & \mathcal{G}_y\end{bmatrix}\begin{bmatrix}(x-x^*)\\(y-y^*)\end{bmatrix}. \tag{10.35}$$

The characteristic equation and the eigenvalues are

$$(\mathcal{F}_x - m)\left(\mathcal{G}_y - m\right) - \mathcal{G}_x\mathcal{F}_y = 0, \tag{10.36}$$

$$m^2 - m\mathcal{F}_x - m\mathcal{G}_y + \mathcal{F}_x\mathcal{G}_y - \mathcal{G}_x\mathcal{F}_y = 0, \tag{10.37}$$

$$m = \frac{\mathcal{F}_x + \mathcal{G}_y \pm \sqrt{\left(\mathcal{F}_x + \mathcal{G}_y\right)^2 - 4\left(\mathcal{F}_x\mathcal{G}_y - \mathcal{G}_x\mathcal{F}_y\right)}}{2}. \tag{10.38}$$

One can assess the kind of steady state rapidly using (10.38) alone.

10.4.3 Nonlinear Systems

We have already seen in one dimension that a nonlinear system, unlike a linear one, can have multiple solutions distinct from one another. The same in true in two dimensions. One example is shown in Fig. 10.17, which is a phase portrait of the system,

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}(x-x^*)\\(y-y^*)\end{bmatrix} = \nabla\left(e^{-(x-2)^2-(y+1)^2} + e^{-(x+1)^2-(y-2)^2}\right). \tag{10.39}$$

This system is generated from a potential function, which is contoured in Fig. 10.17b. It is clear that this potential function has two different extrema, near $x = 2, y = -1$ and $x = -1, y = 2$. Also plotted in Fig. 10.17a are the *nullclines*. These lines indicate where $\dot{x} = 0$ (blue) or $\dot{y} = 0$ (orange). Everywhere above the orange line, the y component of trajectories is downward and everywhere below the orange line, the y component of trajectories is upward. Everywhere left of the blue line, the x component of trajectories is rightward and everywhere right of the blue line, the x component of trajectories is leftward.

Another key concept in higher dimensions is the idea of a *basin of attraction*. In this system nearly all trajectories end up at one of the two stable steady state solutions. The basin of attraction of each solution is the region where all trajectories fall toward that solution rather than the other one. Right along the line leading into the half-stable solution is also an (infinitely thin) basin of attraction for that solution. Most points nearby a steady solution lie in its basin of attraction, but sometimes very remote points can as well. In more complicated flows with stable, unstable, and

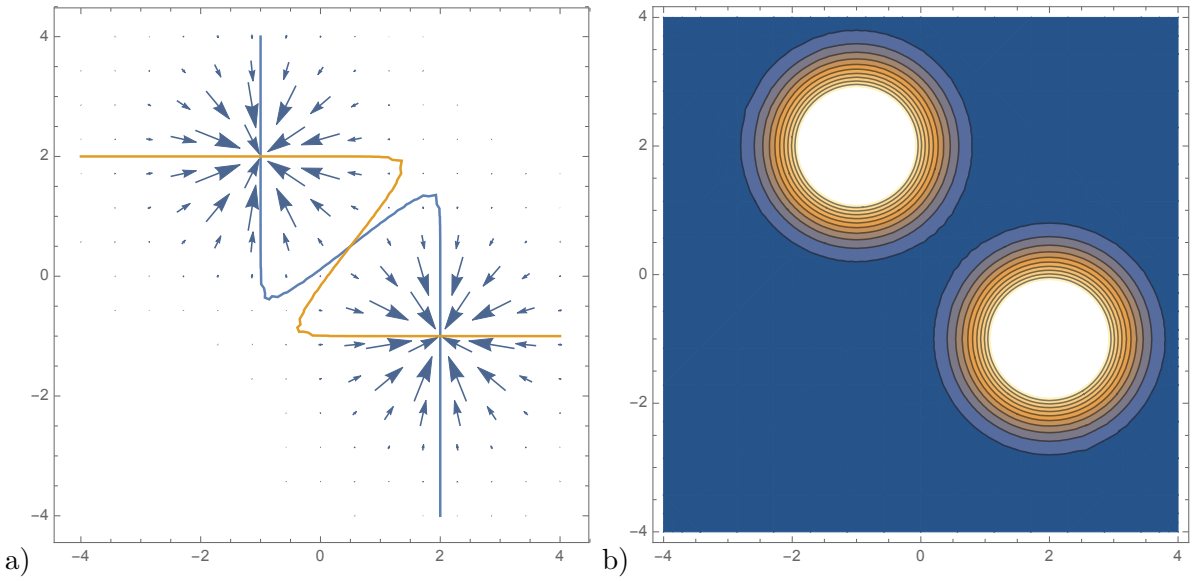


Figure 10.17: a) Phase portrait and b) potential function of (10.39) which has two stable steady state solutions and one half-stable steady state solution (at the intersection of the nullclines).

saddle-point steady states, it can be very difficult to locate the basins of attraction. If these points of attraction and repulsion move in time (as they do in celestial mechanics), then the basins of attraction are extremely complex. The line between basins of attraction is called a *separatrix*, and if it exists this "continental divide" separates the basins of attraction. Sometimes, this line is easy to infer, but for complicated systems it can be very hard as we will see.

10.4.4 Limit Cycles

In the linear two-dimensional system, we saw that some features, those with pure imaginary eigenvalues called centers, featured oscillations around a fixed point rather than spiraling inward or outward. Centers are independent of amplitude–every orbit around the fixed point is oscillatory. This is a characteristic of a linear oscillators, in fact, that the frequency and existence of oscillations is independent of the amplitude of the oscillation. This results from the fact that the eigenvalues are independent of amplitude in (10.33).

Nonlinear oscillations, on the other hand, typically *do* have frequency, stability, and other aspects of the oscillation dependent on amplitude. A simple example to consider is constructed by examining a system in polar coordinates from Strogatz (1994).

$$\dot{r} = r(1 - r^2) + \mu r \cos\theta, \tag{10.40}$$

$$\dot{\theta} = 1. \tag{10.41}$$

Here r, θ are the polar coordinates. You can convert these equations into x, y coordinates if you like, but the results are messy and not informative. The $\dot{\theta}$ equation tells us that the rate of rotation is constant, but the r equation is very interesting. The nullclines of $\dot{r} = 0$ is fairly easy to understand when $\mu = 0$, see Fig. 10.18a. Fig. 10.18b shows the limit cycle that results from this system.

Just as there are different types of steady solutions, there are different types of limit cycles.

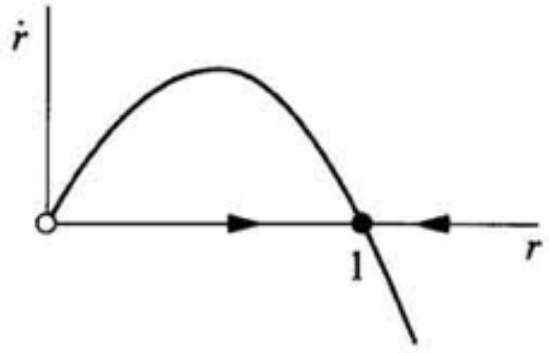


Figure 7.1.1

a)

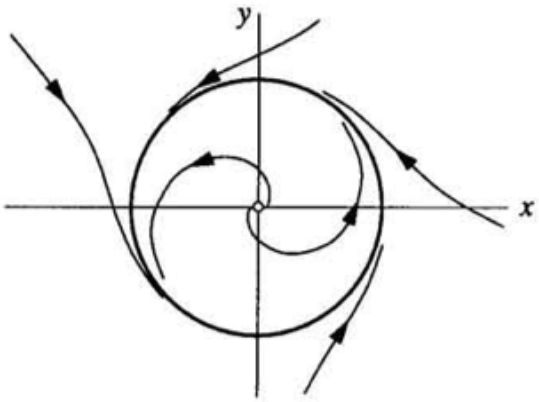


Figure 7.1.2

b)

Figure 10.18: a) and b) potential function of (10.39) which has two steady state solutions.

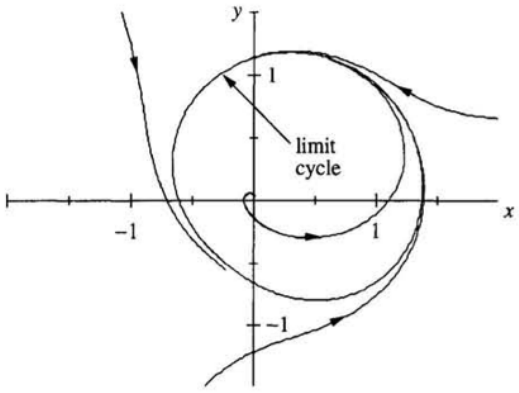


Figure 7.3.3

Figure 10.19: A weakly nonlinear oscillator, from Strogatz (1994).

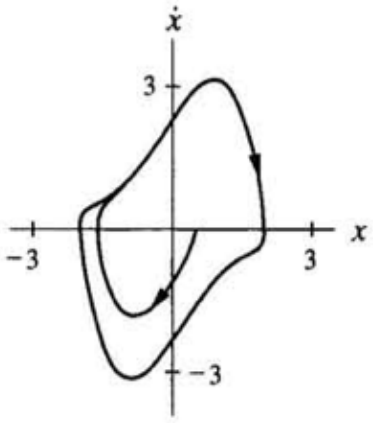


Figure 7.1.4

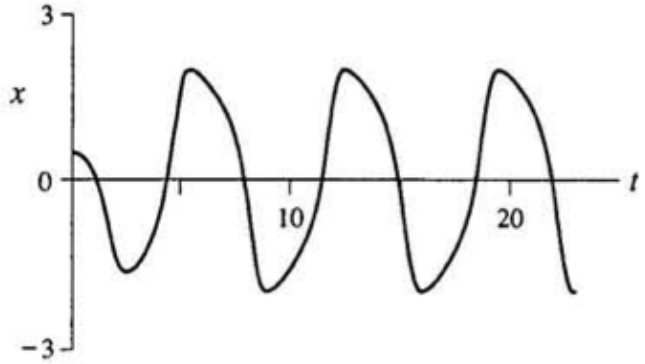


Figure 7.1.5

Figure 10.20: A more obviously nonlinear oscillator, the Van Der Pol oscillator, from Strogatz (1994).

Fig. 10.21 schematizes some of them. They can be attractive or repulsive, stable or unstable. There are basins of attraction for limit cycles just as there are for steady state solutions.

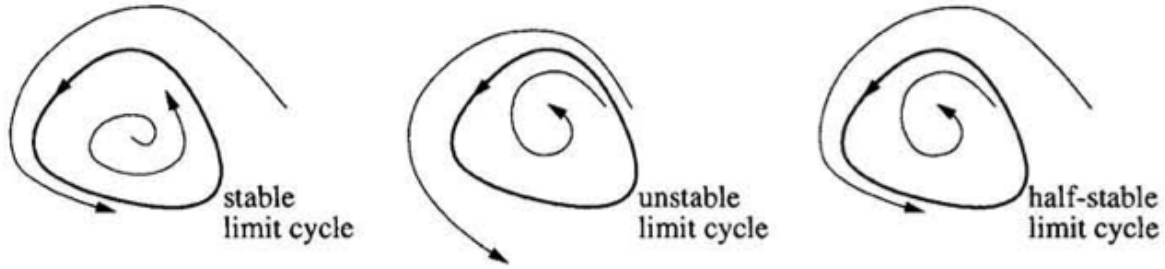


Figure 7.0.1

Figure 10.21: Kinds of limit cycles, from Strogatz (1994).

10.4.5 More Bifurcations

All of the bifurcation types we studied in one dimension carry into two-dimensional analogs, plus there are others! A Hopf bifurcation is where a stable steady state fixed point transforms into an unstable fixed point surrounded by a limit cycle. A homoclinic bifurcation is when a saddle point merges into a limit cycle. A large fraction of effort into studying nonlinear systems goes into mapping and understanding such bifurcations.

10.4.6 Application: Stommel Meridional Overturning Box Model

The Stommel (1961) box model is a simple representation of the buoyancy-driven meridional overturning circulation. It consists of two well-mixed containers of fluid (one represents the tropical ocean and one represents the polar ocean). These containers are allowed to exchange energy and salinity with reservoirs of fixed temperature and salinity. The two containers are allowed to exchange fluid through a capillary tube and an overflow. If there is no overflow, then the temperature in each container will just converge to the reservoir temperatures (which is why I use our notation for steady state solutions * on the reservoir temperatures and salinities).

The Stommel model is a four-dimensional nonlinear system, but two of the dimensions are not interesting. See Vallis (2006, 15.3) for an introduction. Figure 10.22 shows a two box version of

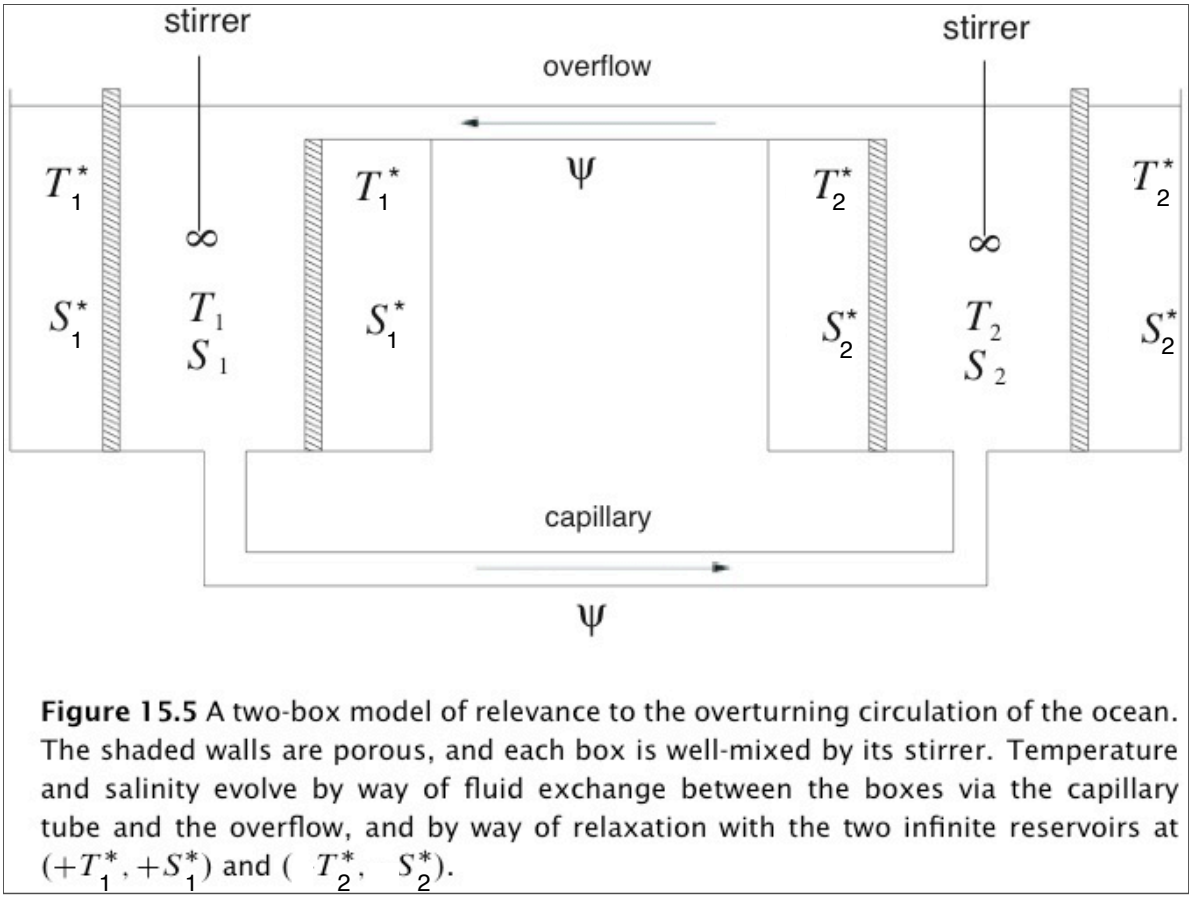


Figure 10.22: Figure defining variables adapted from Vallis's Fig. 15.5.

the Stommel model. The variables to be determined are the temperature and salinity in each box. These are most easily derived by considering the budget of a single box. Consider the amount of salt in one box, it is given by

$$\text{salt} = \iiint \rho S dV. \tag{10.42}$$

Where S is the salinity, or grams of salt per kilogram of salty water. The rate of change of salt is

thus

$$\frac{\mathrm{dsalt}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\iiint \rho S dV, \tag{10.43}$$

$$= \iiint \frac{\partial \rho S}{\partial t} + \nabla\cdot(\rho \mathbf{v} S) dV. \tag{10.44}$$

The stirrers keep the fluid in each box well mixed, so

$$\frac{\mathrm{dsalt}}{\mathrm{d}t} = \frac{\partial \rho S}{\partial t} V + \oiint \rho S \mathbf{v}\cdot\hat{n} dA. \tag{10.45}$$

Water is nearly incompressible, and thus ρ and volume are constant to first order, thus a change in average salinity in a box of changing volume is composed of the change in salinity and the inflow and outflow of salt,

$$\frac{1}{\rho_0 V_0}\frac{d\mathrm{salt}}{dt} = \frac{\partial S}{\partial t} + \frac{\oiint S\mathbf{v}\cdot\hat{n} dA}{V_0}. \tag{10.46}$$

Likewise, the energy budget is related to the transport of (potential) temperature. Interpreting Figure 10.22, we find that the temperature in the boxes obey

$$\frac{dT_1}{dt} = \frac{F_r A_r}{V_0}(T_1^* - T_1) + \frac{|\Psi|}{V_0}(T_2 - T_1), \tag{10.47}$$

$$\frac{dT_2}{dt} = \frac{F_r A_r}{V_0}(T_2^* - T_2) + \frac{|\Psi|}{V_0}(T_1 - T_2). \tag{10.48}$$

Where F_r is a flux rate over the area exposed to restoring A_r. Typical ocean surface values are $30-60 W m^{-2} K^{-1}$ (Haney, 1971), and a typical ocean value of V_0/A_r is the ocean depth, $4km$. One converts temperature to heat in water using $\rho c_p \approx 4\cdot 10^6 J K^{-1} m^{-3}$. We combine these coefficients into a (constant) restoring timescale

$$\tau_r = \frac{V}{F_r A_r} \approx 10 \to 25 yr. \tag{10.49}$$

The volumetric flow rate Ψ in the oceanographically relevant regime is roughly $10-30 Sv$, depending on whether North Atlantic Deep Water and or Antarctic Bottom Water or both are considered. Sv stands for Sverdrups and $1Sv = 10^6 m^3/s$. The absolute value of Ψ appears because the flow into each box in the last term of (10.45) always brings the water type of the upstream box regardless of flow direction. A typical oceanic volume might be $4km(5000km)^2 = 10^{17} m^3$. Thus, we rewrite in terms of the flushing timescale

$$\tau_f(\Psi) = \frac{V_0}{|\Psi|} \approx 200 \to 3000 yr. \tag{10.50}$$

The equations of motion are

$$\frac{dT_1}{dt} = \frac{T_1^* - T_1}{\tau_r} + (T_2 - T_1)\frac{|\Psi|}{V_0}, \tag{10.51}$$

$$\frac{dT_2}{dt} = \frac{T_2^* - T_2}{\tau_r} + (T_1 - T_2)\frac{|\Psi|}{V_0}. \tag{10.52}$$

Salinity might obey similar equations, although we allow for a slightly different restoring time with a constant δ, which is the ratio of temperature restoring timescale to salinity restoring timescale,

$$\frac{dS_1}{dt} = \delta\frac{S_1^* - S_1}{\tau_r} + (S_2 - S_1)\frac{|\Psi|}{V_0}, \tag{10.53}$$

$$\frac{dS_2}{dt} = \delta\frac{S_2^* - S_2}{\tau_r} + (S_1 - S_2)\frac{|\Psi|}{V_0}. \tag{10.54}$$

Using restoring of salinity is specious. More physical results occur if one adds a salt flux where evaporation is to occur, or most realistically add a freshwater flux, making the flow rate out of the rainy box greater than the flow out of the evaporating box. The freshwater flux case will be treated below.

At this stage only the well-mixed property of each box and the artifical salinity restoring are patently unrealistic. Adding more boxes would relieve this problem, and the salinity flux issue is addressed below. However, now we move on to solve a model for τ_f that is unrealistic.

What sets the flushing time τ_f? Realistically, the dynamical mechanisms underlying the strength of the meridional overturning are very complex. At the very least, a momentum equation–and in particular an ageostrophic momentum equation due to the direction of flow–should be solved for τ_f. Stommel makes an assumption, however, that eases the subsequent analysis: that the pressure gradient force is proportional to the flow rate. This balance occurs in very viscous flow, such as driven capillary tubes. Given that the overflow will tend to balance, one expects the pressure to work out so that light water is driven over dense water and the flow and return rates will be equal. By hydrostasy, the driving pressure gradient force should be proportional to the density difference between the boxes. Thus,

$$\frac{1}{\tau_f} = \frac{|\Psi|}{V_0} = \frac{1}{\tau_0}\frac{|\rho_2 - \rho_1|}{\rho_0}. \tag{10.55}$$

The value of $\Delta\rho$ can be set to a typical modern equator to pole density difference, roughly $5kg\ m^{-3}$. Assuming a linear equation of state,

$$\rho = \rho_0\left[1 - \beta_T(T - T_{ref}) + \beta_S(S - S_{ref})\right], \tag{10.56}$$

$$\frac{1}{\tau_f} = \frac{1}{\tau_0}|\beta_T(T_1 - T_2) + \beta_S(S_2 - S_1)| \tag{10.57}$$

What direction is the resulting circulation? We suppose that if box 1 is denser, then it will sink and flow through the bottom pipe. If box 2 is denser, then the opposite will occur.

Combining this equation for τ_f with (10.51-10.54) yields a set of four closed nonlinear first-order evolution equations.

$$\begin{aligned}
\frac{\partial t}{\partial T_1} &= \frac{T_1^* - T_1}{\tau_r} + \frac{(T_2 - T_1)|\beta_T(T_1 - T_2) + \beta_S(S_2 - S_1)|}{\tau_0},\\
\frac{\partial t}{\partial T_2} &= \frac{T_2^* - T_2}{\tau_r} + \frac{(T_1 - T_2)|\beta_T(T_1 - T_2) + \beta_S(S_2 - S_1)|}{\tau_0},\\
\frac{\partial t}{\partial S_1} &= \delta\frac{S_1^* - S_1}{\tau_r} + \frac{(S_2 - S_1)|\beta_T(T_1 - T_2) + \beta_S(S_2 - S_1)|}{\tau_0},\\
\frac{\partial t}{\partial S_2} &= \delta\frac{S_2^* - S_2}{\tau_r} + \frac{(S_1 - S_2)|\beta_T(T_1 - T_2) + \beta_S(S_2 - S_1)|}{\tau_0}.
\end{aligned}$$

We add and subtract these equations to yield,

$$\begin{aligned}
\frac{\partial t}{\partial \Sigma T} &= \frac{\Sigma T^* - \Sigma T}{\tau_r}, \\
\frac{\partial t}{\partial \Sigma S} &= \delta \frac{\Sigma S^* - \Sigma S}{\tau_r}, \\
\frac{\partial t}{\partial \Delta T} &= \frac{\Delta T^* - \Delta T}{\tau_r} - \frac{2\Delta T|\beta_T \Delta T - \beta_S \Delta S|}{\tau_0}, \\
\frac{\partial t}{\partial \Delta S} &= \delta \frac{\Delta S^* - \Delta S}{\tau_r} - \frac{2\Delta S|\beta_T \Delta T - \beta_S \Delta S|}{\tau_0}.
\end{aligned}$$

The first two equations describe the approach of the summed temperatures and salinities $(\Sigma T, \Sigma S)$ to the summed forcing temperature and salinity $(\Sigma T^*, \Sigma S^*)$ beginning from their initial values $(\Sigma T_0, \Sigma S_0)$. The solutions are

$$\begin{aligned}
\Sigma T &= \Sigma T^* + \left[\Sigma T_0 - \Sigma T^*\right] e^{-\frac{t}{\tau_r}}, \\
\Sigma S &= \Sigma S^* + \left[\Sigma S_0 - \Sigma S^*\right] e^{-\frac{t\delta}{\tau_r}}.
\end{aligned}$$

The second two equations are more interesting and nonlinear. If we factor out all of the dimensions based on parameter settings (that is, we use the reservoir temperatures and salinities to eliminate the units), they are

$$\breve{T}_\tau = 1 - \breve{T} - \gamma \breve{T}|\breve{T} - \mu \breve{S}|, \tag{10.58}$$

$$\breve{S}_\tau = \delta(1 - \breve{S}) - \gamma \breve{S}|\breve{T} - \mu \breve{S}|. \tag{10.59}$$

I use a breve over a variable $(\breve{\cdot})$ to denote that its dimensions have been factored out using the reservoir temperature and salinity. The removal of the dimensions is performed by defining $\tau = t/\tau_r$, $\breve{T} = \Delta T/\Delta T^*$, $\breve{S} = \Delta S/\Delta S^*$, $\gamma = 2\beta_T \Delta T^* \tau_r/\tau_0$, $\mu = \beta_S \Delta S^*/(\beta_T \Delta T^*)$. Recall that δ is the ratio of temperature restoring timescale to salinity restoring timescale. Note that $\Delta T^* > 0$ is assumed. It is natural to consider the evolution equation for the nondimensional density difference $\breve{\rho} = \mu \breve{S} - \breve{T}$ as well.

$$\breve{\rho}_\tau = \left[(\delta\mu - 1) - \breve{\rho}\right] - \gamma \breve{\rho}|\breve{\rho}| + \mu \breve{S}(1 - \delta). \tag{10.60}$$

The steady state solutions of (10.58-10.60) occur when

$$\breve{T} = \frac{1}{1 + \gamma|\breve{\rho}|}, \tag{10.61}$$

$$\breve{S} = \frac{\delta}{\delta + \gamma|\breve{\rho}|}, \tag{10.62}$$

$$\breve{\rho} = \frac{\mu\delta}{\delta + \gamma|\breve{\rho}|} - \frac{1}{1 + \gamma|\breve{\rho}|}. \tag{10.63}$$

The steady state solutions are often studied by comparing the left of (10.63) to the right, see Figure 10.23. It is clear from the figure that occasionally multiple steady solutions exist when the straight line of the left-hand-side intersects the curvy line of the right-hand-side in two (10.23b) or three (10.23a) locations.

However, we can get a clearer approximate picture of the trajectories during equilibration If we approximate the last term in (10.60) to make one decoupled equation for $\breve{\rho}$. One could use the

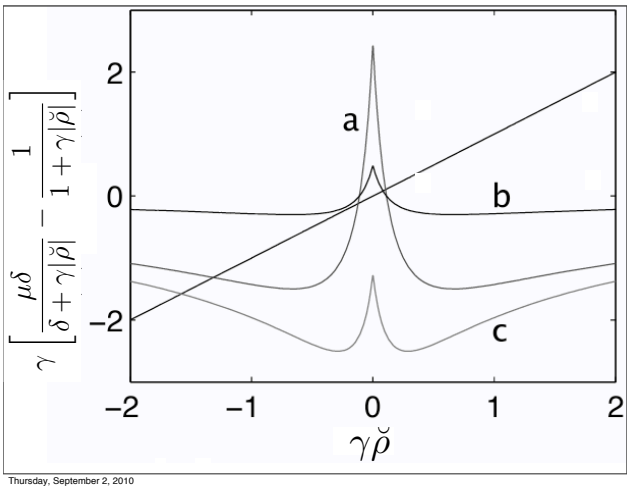


Figure 10.23: The left side of (10.63) versus different right sides with varying parameters: a) $\gamma = 5, \mu = 1.5$, b) $\gamma = 1, \mu = 1.5$, and c) $\gamma = 5, \mu = 0.75$. In all plots, $\delta = 1/6$. From Vallis Fig. 15.7.

steady solution (10.62), but since $\delta < 1$ usually, $\breve{T}$ usually spins up to near-equilibrium faster than $\breve{S}$. Thus, we approximate $\mu S = \breve{\rho} + \breve{T} \approx \breve{\rho} + \frac{1}{1+\gamma|\breve{\rho}|}$, so

$$\breve{\rho}_\tau \approx \delta\mu - 1 - \breve{\rho}(\delta + \gamma|\breve{\rho}|) + \tfrac{1-\delta}{1+\gamma|\breve{\rho}|}.. \tag{10.64}$$

This equation allows us to draw a phase space diagram, as shown in Fig. 10.24

Hysteresis in the Stommel Model

The Stommel model has become famous as an example of a plausibly climate-like system with an accessible 'tipping point'. This tipping point has led to speculation that there might be a similar one in the real ocean, which might be triggered by climate forcing. If so, the nature of the meridional overturning circulation might change dramatically and quickly, and from that state be very difficult to restore. We will illustrate the 'tipping points' in the Stommel model along with the more general concept of *hysteresis*. Hysteresis is defined as a lack of reversibility in behavior under a change of forcing or parameter.

Consider forcing the Stommel model by altering the value of $\Delta T^* \to \Delta T^*(1+\alpha)$, a crude representation of warming the poles under climate change. So long as α varies slowly, we expect the model to approach the steady state solutions (10.61-10.64), and then track along with the slow changes to α. That is,

$$\begin{aligned}
\breve{T} &= \frac{1+\alpha}{1+\gamma|\breve{\rho}|}, \\
\breve{S} &= \frac{\delta}{\delta+\gamma|\breve{\rho}|}, \\
\breve{\rho} &= \frac{\mu\delta}{\delta+\gamma|\breve{\rho}|} - \frac{1+\alpha}{1+\gamma|\breve{\rho}|}, \\
\breve{\rho}_\tau &\approx \delta\mu - (1+\alpha) - \breve{\rho}(\delta+\gamma|\breve{\rho}|) + \frac{(1-\delta)(1+\alpha)}{1+\gamma|\breve{\rho}|}.
\end{aligned}$$

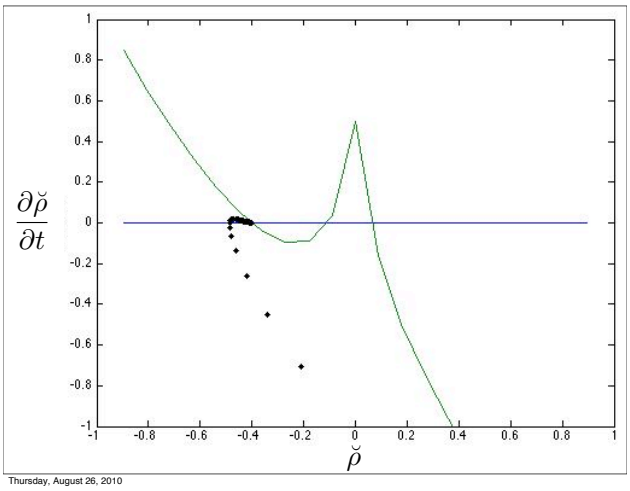


Figure 10.24: An approximate phase space diagram from (10.64). Assuming $\breve{S}$ is near its steady state solution, for all $\breve{\rho}$ where the green line is positive, $\frac{d\breve{\rho}}{d\tau} > 0$ so the trajectory will proceed to the right (increasing $\breve{\rho}$). For all $\breve{\rho}$ where the green line is negative, $\frac{d\breve{\rho}}{d\tau} < 0$. Intersections of the green and blue lines are fixed points. Can you figure out which fixed points are stable? How does one change the initial conditions to reach the other fixed points? A sample trajectory is shown.

Based on these equations, we can plot the slowly-changing equilibrium values as a function of α. The solid line in figure 10.25 shows the equilibrium values for one example. For this example, three branches are available for α roughly between 0 and 0.5. The upper branch is salt-driven (i.e., sinking of salty equatorial water), the lower branch is temperature-driven (i.e., sinking of cold water at the pole), and the middle branch is the unstable equilibrium. Superimposed on the equilibrium values are the trajectories for three different simulations begun with the same initial conditions subjected to sinusoidal variations of α. In all cases, the model initially converges to the temperature-driven equilibrium. For the smallest amplitude variations of α (black), the model oscillates about this value, but no obvious hysteresis is observed–a local linearization about the initial equilibrium would explain this behavior well. For the next highest amplitude of α forcing, the forcing triggers a transition from the temperature-driven mode to the salt-driven mode the first time α reduces the temperature gradient forcing. After this transition, the forcing is not strong enough to trigger a second transition back to the temperature-driven mode. Thus, the meridional overturning circulation collapses, and even though the forcing is sufficient to restore *beyond* the initial temperature gradient forcing, the model stays in the salt-driven mode. For the strongest forcing shown (red), the model switches back and forth between the temperature-driven and salt-driven modes. Note that the transitions to and from the temperature-driven mode do not occur at the same value of $(\alpha, \breve{\rho})$, there is obvious hysteresis in the system. Oscillators of this type are sometimes called *relaxation oscillators*. While α forces the temperature difference, a similar effect occurs when the salinity difference is forced (as might occur were the Greenland ice cap to melt).

Note that the bifurcation points (where the upper and middle branches appear near $\alpha = 0.5$ and where the lower and middle branches appear near $\alpha = 0$) are not sharp transition points, but approximate regions of transition. This effect is more or less pronounced if the frequency of variation of α is higher or lower, respectively.

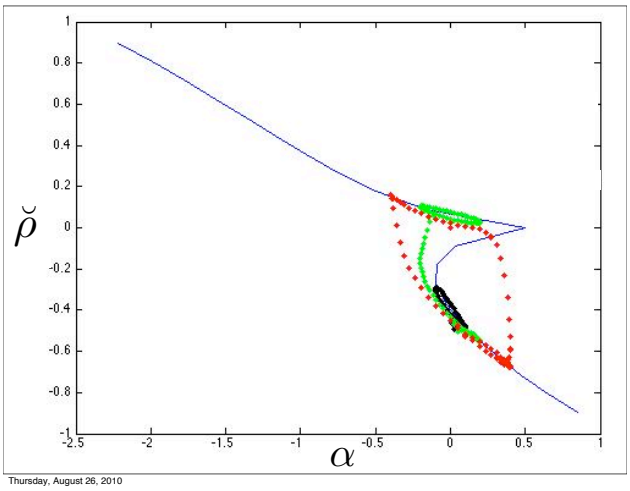


Figure 10.25: The possible equilibrium values of $\breve{\rho}$ for different values of α (blue solid), and three trajectories where α is varied sinusoidally in time: (black) $-0.1 \leq \alpha \leq 0.1$, (green) $-0.2 \leq \alpha \leq 0.2$, and (red) $-0.4 \leq \alpha \leq 0.4$.

10.5 Linearization Near a Fixed Point: General Treatment

We have been interested in systems of low-order ordinary differential equations. By replacing the higher-order derivatives with extra equations introducing new intermediate variables (e.g., acceleration & position to acceleration, velocity, & position) these can all be written as sets of first-order equations. Generically, we write:

$$x_{j,t} = \mathcal{N}_j(x_i; \beta) \tag{10.65}$$

for some nonlinear operator $\mathcal{N}$, set of variables x_i, and set of parameters β. There must be as many equations as unknowns, so we expect j and i to have the same number of values.

We then found fixed points of the model, solved by

$$0 = \mathcal{N}_j(x_i^*; \beta). \tag{10.66}$$

In the models discussed above, finding the fixed points x_i^* involved inverting trigonometric functions or solving coupled polynomial equations, but it is possible that numerical determination of these points (e.g., by a Newton-Raphson method, Section 8.9.1) may be required.

Then, we linearized the system about each fixed point, by Taylor series expansion

$$\begin{aligned} x_{j,t} &\approx \mathcal{N}_j(x_i^*; \beta) + \sum_{All\ i} \left(\left. \frac{\partial \mathcal{N}_j}{\partial x_i} \right|_{x_i = x_i^*} \cdot \left(x_i - x_i^*\right) \right), \\ &= \sum_{All\ i} \left(\left. \frac{\partial \mathcal{N}_j}{\partial x_i} \right|_{x_i = x_i^*} \cdot \left(x_i - x_i^*\right) \right) + O(\left(x_i - x_i^*\right)^2). \end{aligned} \tag{10.67}$$

The second equation follows from x_i^* being fixed points of the solution in (10.66). Again, if closed analytic form is not possible, numerical estimation of the derivative is straightforward. For exam-

ple,

$$\left.\frac{\partial \mathcal{N}_j}{\partial x_i}\right|_{x_i=x_i^*} \approx \lim_{x_i \to x_i^*} \frac{\mathcal{N}_j(x_i;\beta) - \mathcal{N}_j(x_i^*;\beta)}{x_i - x_i^*} \tag{10.68}$$

Since we must have as many equations as unknowns, the Jacobean matrix–the set of the partial derivative of all j values of $\mathcal{N}$ with respect to all directions $x_i - x_i^*$ away from the fixed point in (10.68)–is a square matrix. All square matrices obey an eigenrelationship, for scalar eigenvalues λ_n and eigenvectors with components v_{jn}.

$$\sum_{All\ i} \left(\left.\frac{\partial \mathcal{N}_j}{\partial x_i}\right|_{x_i=x_i^*} \cdot v_{in} \right) = \lambda_n v_{jn}. \tag{10.69}$$

If the eigenvalues are distinct, the eigenvectors are unique. Even if they are not distinct, then there is always a basis set of eigenvectors for any duplicated eigenvalues. The eigenvalues may be complex: real parts will be the exponential decay or growth rates of each eigenmode and the imaginary parts will be the oscillation frequencies.

If the eigenvectors are written as the columns of a matrix, and the eigenvalues are written as the diagonal elements of a matrix, then this equation is equivalent to the matrix equation (repeated indices indicate summation via matrix multiplication):

$$\left[\left.\frac{\partial \mathcal{N}_j}{\partial x_i}\right|_{x_i=x_i^*} \right] [v_{in}] = [v_{jm}]\,[\lambda_{mn}]\,. \tag{10.70}$$

Section 3.8 shows how to calculate the eigenvectors and eigenvalues of a matrix. The eigenvector matrix $[v_{in}]$ can be inverted to yield:

$$\left[\left.\frac{\partial \mathcal{N}_j}{\partial x_i}\right|_{x_i=x_i^*} \right] = [v_{jm}]\,[\lambda_{mn}]\,[v_{in}]^{-1} \tag{10.71}$$

When this result is plugged into the linearized equations, the resulting decoupling of modes is acheived.

$$\begin{aligned} x_{j,t} &\approx \sum_{All\ i} \left(\left.\frac{\partial \mathcal{N}_j}{\partial x_i}\right|_{x_i=x_i^*} \cdot \left(x_i - x_i^*\right) \right), && (10.72) \\ &\approx [v_{jm}]\,[\lambda_{mn}]\,[v_{in}]^{-1} \left(x_i - x_i^*\right) && (10.73) \end{aligned}$$

Using the inverse eigenvector matrix one more time yields the desired decoupling for the 'linear eigenmodes' near the fixed point are $\phi_m = \left[v_{jm}\right]^{-1} \left[x_j - x_j^*\right]$.

$$\begin{aligned} \left[v_{jm}\right]^{-1} \left[x_j - x_j^*\right]_t &\approx [\lambda_{mn}]\,[v_{in}]^{-1} \left(x_i - x_i^*\right), \\ \phi_{m,t} &\approx \lambda_m \phi_m. \end{aligned} \tag{10.74}$$

It is clear that the solution for each linear eigenmode is just

$$\phi_m = \phi_m|_{t=0}\, e^{\lambda_m t} \tag{10.75}$$

Clearly, the eigenmodes may grow ($\mathrm{Re}(\lambda_m) > 0$), decay ($\mathrm{Re}(\lambda_m) < 0$), and may oscillate, too, if $\mathrm{Im}(\lambda_m) \neq 0$.

The evolution of the original variables x_i near the fixed point will also grow, decay, and/or oscillate *along the phase space direction aligned with each eigenvector*. That is, if $\lambda_1 > 0$, then the fixed point will be unstable to perturbations in the direction of the column vector $\left[v_{j1}\right]$. Likewise, if $\lambda_2 < 0$, then the fixed point will be stable to perturbations in the direction of the column vector $\left[v_{j2}\right]$. If the real part of all eigenvalues are negative/positive, then the fixed point is stable/unstable. If the imaginary part of any eigenvalue is nonzero, then there can be oscillations near or around the fixed point.

10.6 Chaos

When a system becomes sufficiently complex, the basins of attraction can be very complicated. This is our first window into *deterministic chaos*, where a simple deterministic question, "If I start at this point in phase space, where do I end up?" has a very complicated answer.

Consider the motion of a particle in a potential with friction,

$$\dot{v} = -\frac{\partial V(x)}{\partial x} - rv, \tag{10.76}$$

$$\dot{x} = v. \tag{10.77}$$

This second-order system requires two initial conditions, one on velocity and one on position. Even for simple potentials, the basin of attraction can be complicated (Fig. 10.26a). For a more complicated system, such as

$$\dot{v} = -0.1\dot{\theta} + \sin\theta = 2.1\cos t, \tag{10.78}$$

$$\dot{\theta} = v. \tag{10.79}$$

the basin of attraction is a fractal. Using this system here is cheating a bit, because adding the time-dependent forcing raises the dimensionality of the system above two dimensions, but it is a fun example!

10.6.1 Lorenz Equations

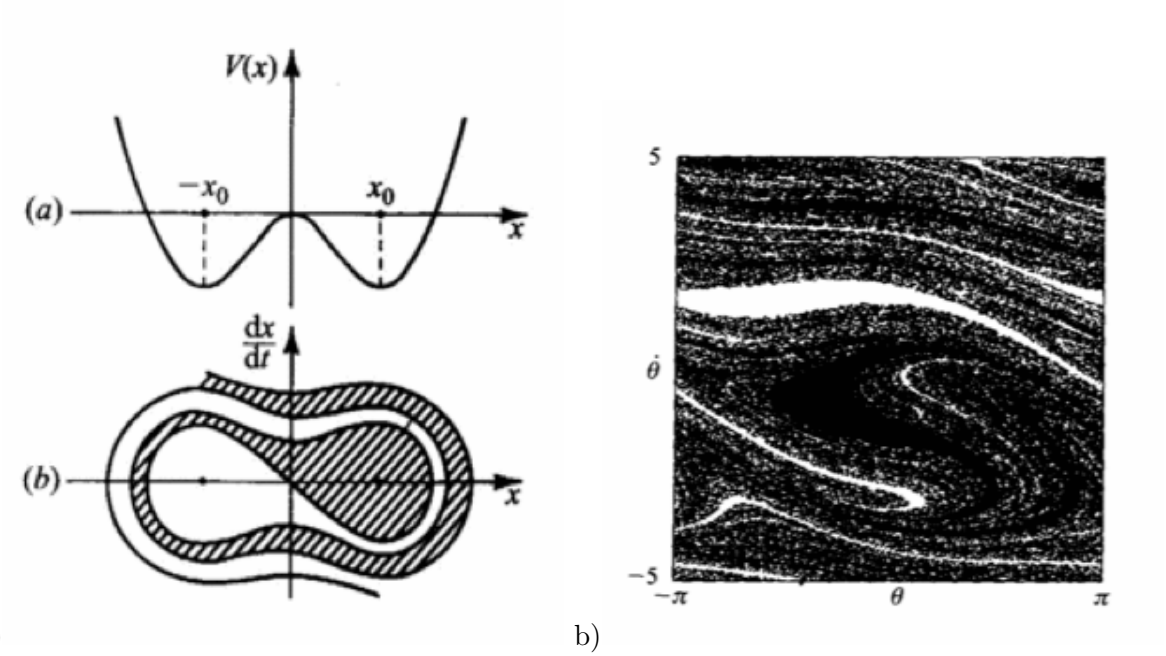
The study of chaos in some sense originated with Lorenz (1963). He was interested in numerical forecasts of the weather (and indeed was one of the best atmospheric scientists of the 20th century, in addition to being the father of chaos). He boiled down a basic model for convection to the equations,

$$\dot{x} = \sigma(y - x), \tag{10.80}$$

$$\dot{y} = rx - y - xz, \tag{10.81}$$

$$\dot{z} = xy - bz. \tag{10.82}$$

At first glance, these equations are no more crazy than any we have studied. Lorenz coded them up and found that they were integrable and smooth (Fig. 10.28). However, it is very difficult



a) b)

Figure 10.26: a) The double well potential and its basin of attraction (10.76), and b) the basin of attraction for a forced, damped pendulum (10.78). From Ott (2006).

to anticipate when a trajectory will jump from oscillating about one of the fixed points to the other. Before Lorenz, it was well-known that nonlinear systems exhibit nonperiodic or quasiperiodic behavior. However, in Lorenz's model, even neighboring initial conditions, which one would expect to remain neighbors, soon become separated onto different sides of the attractor. Lorenz describes the implications,

> When our results concerning the instability of nonperiodic flow are applied to the atmosphere, which is ostensibly nonperiodic, they indicate that predication of the sufficiently distant future is impossible by any method, unless the present conditions are known exactly. In view of the inevitable inaccuracy and incompleteness of weather observations, precise very-long-range forecasting would seem to be non-existent.

The Liapunov exponent describes how quickly neighboring trajectories diverge. If the distance between the points is initially, δ_0, then once the trajectories begin to diverge, they do so according to

$$|\delta(t)| \sim |\delta_0| e^{\lambda t}. \tag{10.83}$$

The largest such separation rate λ over different ways of measuring the separation distance is called the Liapunov exponent. The time beyond which forecasts become pointless scales as

$$T_{\text{chaos}} \sim O\left(\frac{1}{\lambda} \ln \frac{a}{|\delta_0|}\right). \tag{10.84}$$

Chaotic and nonperiodic solutions often arise from a series of bifurcations that occur as a parameter varies. This is called the "period doubling" route to chaos. First, a trajectory stays near a fixed point or limit cycle. As the parameter is varied, a bifurcation occurs leading to two fixed points

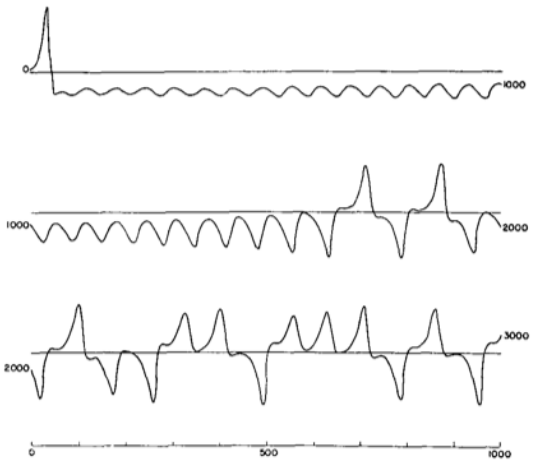


FIG. 1. Numerical solution of the convection equations. Graph of Y as a function of time for the first 1000 iterations (upper curve), second 1000 iterations (middle curve), and third 1000 iterations (lower curve).

a)

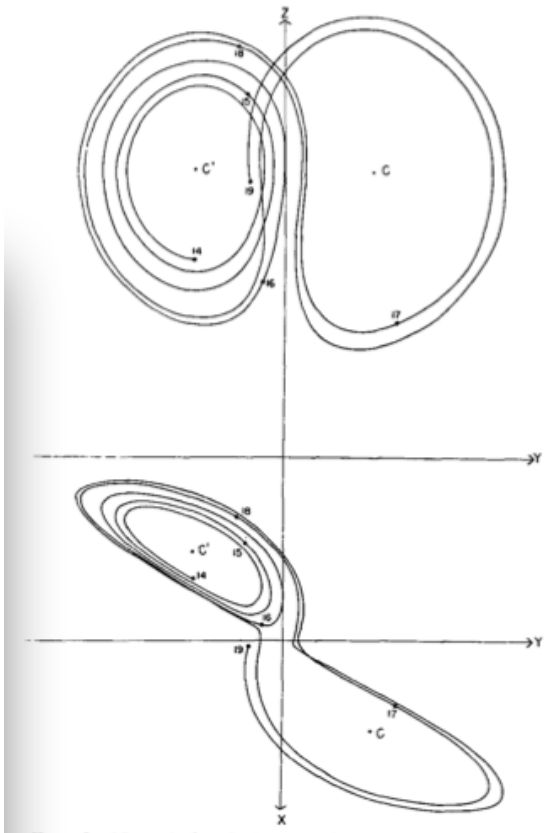


FIG. 2. Numerical solution of the convection equations. Projections on the X-Y-plane and the Y-Z-plane in phase space of the segment of the trajectory extending from iteration 1400 to iteration 1900. Numerals "14," "15," etc., denote positions at iterations 1400, 1500, etc. States of steady convection are denoted by C and C'.

b)

Figure 10.27: a) Time series of y from (10.80) and b) a three-dimensional visualization of trajectories in (10.80).

or limit cycles, then four, then eight. Often, not only do the number of attractors change as the parameter varies, but the rate of bifurcation with change of parameter also increases. Before, long, the system fills into a strange attractor instead of a large number of fixed points.

Some movies help with visualization: http://www.youtube.com/watch?v=FYE4JKAXSfY.

10.6.2 Strange Attractors

The Lorenz trajectories hug an object similar to a limit cycle, but which is much more complex. It is called a "strange attractor" because it attracts trajectories and because it is strange. It is strange because it is fractal in space. The Lorenz attractor is a fractal, because although it appears to be two dimensional, it has in fact a fractal dimension of 2.06 (the meaning of this will be revealed below). Thus, the closer one examines it, the more structure it has.

10.7 Fractals and Self-Similarity

Mandelbrot (1967) is a nice example of fractals being applied to a geophysical problem. Here is the abstract,

> Geographical curves are so involved in their detail that their lengths are often infinite or, rather, undefinable. However, many are statistically "self-similar," meaning that each portion can be considered a reduced-scale image of the whole. In that case, the degree of complication can be described by a quantity D that has many properties of a "dimension," though it is fractional; that is, it exceeds the value unity associated with the ordinary, rectifiable, curves.

This insight was inspired partly by Lewis Richardson, a mathematician and physicist deeply interested in geophysical problems. He performed the first (unsuccessful) numerical weather forecast. Richardson also authored the following poem about turbulence,

> Big whirls have little whirls that feed on their velocity, and little whirls have lesser whirls and so on to viscosity.

The fractal dimension D chosen by Mandelbrot (there are many flavors, with the Hausdorff one typically being preferred) is now called the box-counting dimension, and is defined as

$$D = \lim_{\epsilon \to 0} \frac{\ln N(\epsilon)}{\ln \frac{1}{\epsilon}}. \tag{10.85}$$

Where $N(\epsilon)$ is the number of boxes of side length ϵ required to cover the shape being observed. Mandelbrot and Richardson also considered a version of this formula using lines instead of boxes.

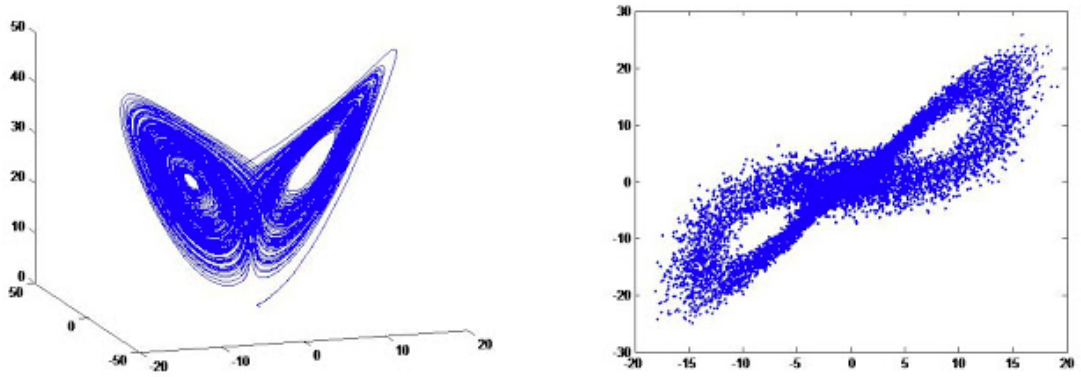
This Nova special is a nice description: http://www.youtube.com/watch?v=s65DSz78jW4.

10.7.1 Poincaré Maps

One reason why the appreciation of chaos in dynamical systems was slow to come is because before computers, we could not easily produce trajectories like in Fig. 10.28. In the real world, we might

be severely limited to what we can observe. Poincaré invented a mathematical analog of our limited capability to observe that is useful in understanding chaos.

Suppose we are studying a limit cycle or an attractor, but we can only observe it at a few locations in phase space. For example, if we have a pendulum swinging, and monitor only its maximum height, not its velocity. Or, in our estuary example, we monitor only the inflow and outflow, not the flows within the estuary, etc. We can think about such a limited observation as being tracking subsequent punctures through a plane in phase space that transects the attractor. Fig. 10.28 shows a Poincaré map for the Lorenz model. Limit cycles and periodic trajectories will have simple



a) b)

Figure 10.28: a) Trajectory from (10.80) and a matching b) Poincaré map from a plane near the center of the attractor. From Batterson (2008).

Poincaré maps, where the punctures through the plane will repeat. Strange attractors tend to have dense Poincaré maps, sometimes they are called *space-filling*, because eventually every point in a region of the plane will have a trajectory pass closely by it. No matter how close you consider close (pick a distance, any distance), there will eventually be a trajectory or point on the map within that distance eventually.

Recall our definition of a dynamical system: 1) a set of variables describing the system state, 2) a set of times when the states can be observed, and 3) a set of rules for evolving the state at one time until the next. Now, notice that *we don't even need the trajectories to form a dynamical system, the Poincaré map itself is a dynamical system!* We only need the trajectories to generate the map, but if somebody told us a short cut to the map, we could use that as a dynamical system. This broadening of our conception of what constitutes a dynamical system leads to many fun and (maybe) useful fractal patterns, as well as the concept of sets.

10.7.2 Fractal Sets and Basins of Convergence

A set is a pattern that maps onto itself when operated on by a map. Many sets are chaotic or fractal, and the classic examples of such are the Julia and Mandelbrot sets. These sets of points are taken as beautiful examples of approximately self-similar fractals. The Julia sets are all of the sets that are invariant under repeated application of a complex function (often a simple polynomial like $z_{n+1} = z_n^2 + C$), while the colors around the Julia set denote the rate of escape from a bounded region. The Mandelbrot is a similar complex parameter plot, except the shading is determined by the behavior of a Julia set (is it made from connected or unconnected points). Explore!

Another fun example is the map of basins of attraction in Newton's method. Newton's method is a way to locate roots of a nonlinear equation or extrema of a nonlinear function, for example finding the maximum or minimum of a nonlinear differentiable function. Newton's method is iterative: you guess an initial location and the repeat over and over again. Interestingly, the possible subsequent iterations form a dynamical system, as Newton's method is a map (a deterministic method) going from point to point in a space of possible solutions. When there is more that one root for Newton's method to find, which one do you end up at and how quickly? Fig. 10.29

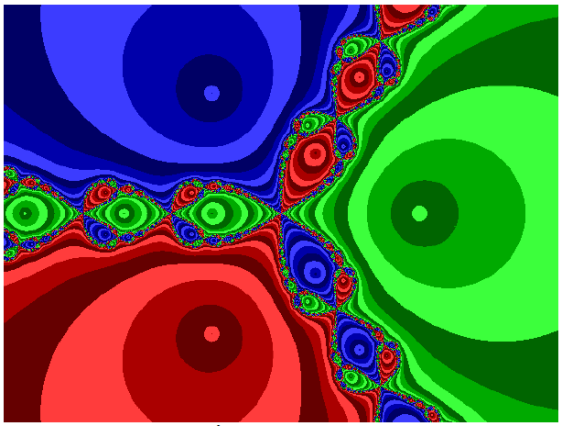


Figure 10.29: Basins of attraction in a Newton's method with colors indicating the convergence to the three complex solutions of $z^3 - 1 = 0$. From Lahr and Reiter (2014).

10.7.3 Scaling Laws

Because of the self-similar nature of fractals, they often obey simple scaling laws on different length scales of consideration. The fractal dimension is one such example. However, other relationships are also self-similar, and geophysics is rife with such rules. The Kolmogorov (1941) turbulence scaling gives a relationship between the energy in turbulence and the lengthscale. The Richter (1935) and Gutenberg and Richter (1942) earthquake magnitude scaling laws are also consistent with self-similarlity (Rundle, 1989).

10.8 Example Problems

10.8.1 Jargon to Argot

Example 10.1 *Look these up online if need be. a) What is the difference between a "fixed point" and a "limit cycle" and an "strange attractor"? b) What is a "trajectory"? c) What is a "map"?*

a) A fixed point is a steady state solution at a single location in phase space. A limit cycle is a simple, repeating trajectory through phase space that attracts or repels neighboring trajectories. A strange attractor is a complex, chaotic, structure often of fractal dimension that behaves like a limit cycle, but is not a simple trajectory itself. b) A trajectory is a time history through phase space of an integration of the field equations beginning with a defined initial condition and over a finite duration. c) A map is a function that relates every point in phase space to another point in

another phase space. Many maps point from and to the same phase space. Maps may be one-to-one (every point maps to a unique point and every point is mapped to by only one point) or multivalued (more that one location maps to the same point).

Example 10.2 *Complex Eigenvalues (Strogatz, 1994, problem 5.2.2).*
This example leads you through the solution of a linear system where the eigenvalues are complex. The system is $\dot{x} = x - y, \dot{y} = x + y$.
a) Write the equations in matrix form and determine the trace, determinant, and eigenvalues of the coefficient matrix. Show that the eigenvalues and eigenvectors are: $m_1 = 1 + i$ *and* $\mathbf{v}_1 = (i, 1)$, $m_2 = 1 - i$ *and* $\mathbf{v}_2 = (-i, 1)$. *b) The general solution is* $\mathbf{x}(t) = c_1 e^{m_1 t}\mathbf{v}_1 + c_2 e^{m_2 t}\mathbf{v}_2$. *So in one sense we're done! But, this way of writing* $\mathbf{x}(t)$ *purely involves complex coefficients and looks unfamiliar. Express* $\mathbf{x}(t)$ *purely in terms of real-valued functions. (Hint: Use Euler's relation to rewrite in terms of sines and cosines, and the separate the terms that have a prefactor of i from those that don't). c) Given the determinant and trace of the coefficient matrix, as well as the eigenvalues, what kind of trajectories are the result of this system? Where are the fixed points?*

a)

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}(x-x^*)\\(y-y^*)\end{bmatrix} = \begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}(x-x^*)\\(y-y^*)\end{bmatrix}. \tag{10.86}$$

$$\tau = 2 \tag{10.87}$$

$$\Delta = 2 \tag{10.88}$$

$$m_{1,2} = \frac{2 \pm \sqrt{4-8}}{2} = 1 \pm i, \tag{10.89}$$

$$\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}i\\1\end{bmatrix} = \begin{bmatrix}i-1\\i+1\end{bmatrix} = (1+i)\begin{bmatrix}i\\1\end{bmatrix} \tag{10.90}$$

$$\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}-i\\1\end{bmatrix} = \begin{bmatrix}-i-1\\-i+1\end{bmatrix} = (1-i)\begin{bmatrix}-i\\1\end{bmatrix} \tag{10.91}$$

b)

$$\begin{bmatrix}x\\y\end{bmatrix} = c_1 e^{(1+i)t}\begin{bmatrix}i\\1\end{bmatrix} + c_2 e^{(1-i)t}\begin{bmatrix}-i\\1\end{bmatrix}, \tag{10.92}$$

$$\tag{10.93}$$

b)

$$\begin{bmatrix} x \\ y \end{bmatrix} = c_1 e^{(1+i)t} \begin{bmatrix} i \\ 1 \end{bmatrix} + c_2 e^{(1-i)t} \begin{bmatrix} -i \\ 1 \end{bmatrix}, \tag{12.92}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = c_1 e^{t}(\cos t + i \sin t) \begin{bmatrix} i \\ 1 \end{bmatrix} + c_2 e^{t}(\cos t - i \sin t) \begin{bmatrix} -i \\ 1 \end{bmatrix}, \tag{12.93}$$

$$c_1 = c_{1r} + i c_{1i}, c_2 = c_{2r} + i c_{2i} \tag{12.94}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = e^{t} \begin{bmatrix} -c_{1i}\cos t - c_{1r}\sin t + i c_{1r}\cos t - i c_{1i}\sin t \\ c_{1r}\cos t - c_{1i}\sin t + i c_{1i}\cos t + i c_{1r}\sin t \end{bmatrix} + e^{t} \begin{bmatrix} c_{2i}\cos t - c_{2r}\sin t - i c_{2r}\cos t - i c_{2i}\sin t \\ c_{2r}\cos t + c_{2i}\sin t + i c_{2i}\cos t - i c_{2r}\sin t \end{bmatrix}, \tag{12.95}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = e^{t} \begin{bmatrix} (c_{2i} - c_{1i})\cos t - (c_{1r} + c_{2r})\sin t \\ (c_{1r} + c_{2r})\cos t + (c_{2i} - c_{1i})\sin t \end{bmatrix} + i e^{t} \begin{bmatrix} (c_{1r} - c_{2r})\cos t - (c_{1i} + c_{2i})\sin t \\ (c_{1i} + c_{2i})\cos t + (c_{1r} - c_{2r})\sin t \end{bmatrix}. \tag{12.96}$$

This final equation can be made pure real (or imaginary) by choosing the sums and differences of the coefficients correctly. If real parts are equal ($c_{1r} = c_{2r}$) and the imaginary parts are opposite ($c_{1i} = -c_{2i}$), a.k.a., c_1 and c_2 are complex conjugates, then the solution is pure real.
c) These are clockwise and counterclockwise outward spirals from a single (unstable) fixed point at $x = 0, y = 0$.

10.9 Homework Problems

10.9.1 Manipulation

Exercise 10.1 *Locate the fixed points of the following one-dimensional functions and use linear stability analysis and phase planes ($\dot{x}$ vs. x) to categorize them. Consider only real values of a and x. a) $\dot{x} = a - e^{-x^2}$, b) $\dot{x} = ax - x^5$ (note that a can be of positive, zero, or negative–consider all 3).*

Exercise 10.2 *Draw the bifurcation diagrams (x^* vs. a) for the following functions as a varies: a) $\dot{x} = a - e^{-x^2}$, b) $\dot{x} = ax - x^5$.*

10.9.2 Application

Exercise 10.3 *The Stommel (1961) model exhibits hysteresis (Section 10.4.6) when the forcing is varied. The Zaliapin and Ghil (2010) model (Section 9.2.3) also exhibits hysteresis when the incoming solar radiation is varied. Which of the the following systems has hysteresis as the parameter a is varied between -1 and 1? Why or why not?*

$$a) \quad \dot{x} = ax - x^3, \tag{10.94}$$

$$b) \quad \dot{x} = a + 2x - x^3, \tag{10.95}$$

$$c) \quad \dot{x} = a - x^2 \tag{10.96}$$

10.9.3 Evaluate & Create

Exercise 10.4 *Go to* http://www.easyfractalgenerator.com *and make some pretty Julia and Mandelbrot sets! Print out your favorites. Advanced: Try out the design a custom fractal!*

Chapter 11

Partial Differential Equations

11.1 Introduction–The Bedrock of Geophysics

Reading: (Boas, 2006, 13.1-13.4)

We now have the tools in place to begin examining partial differential equations. These equations may include both temporal and spatial derivatives *at the same time*. We know now how to think about the different flavors of derivatives that are important in three dimensions–div, grad, and curl–as well as their multi-dimensional antiderivatives–the fundamental theorems: Gauss's divergence theorem, the line integral, and the Stokes theorem. Using such equations to describe the motions and dynamics of fields in space and time was the transformation that allowed physics to evolve over the twentieth century. Longair (2015) notes that the origin of these developments goes back to the (Maxwell, 1865) paper on electricity and magnetism. Einstein (1931) sums it up:

> We may say that, before Maxwell, Physical Reality, in so far as it was to represent the process of nature, was thought of as consisting in material particles, whose variations consist only in movements governed by partial differential equations. Since Maxwell's time, Physical Reality has been thought of as represented by continuous fields, governed by partial differential equations, and not capable of any mechanical interpretation. This change in the conception of Reality is the most profound and the most fruitful that physics has experienced since the time of Newton.

We will study many of the most common partial differential equations in a variety of geometries. To keep things simple, we will often use rectangular geometry, even though realistic problems are more often not rectangular (e.g., the earth). These partial differential equations are the fundamentals of describing the relationships between environmental variables, and they come up again and again.

11.2 Laplace's Equation

Laplace's equation is the simplest one involving a Laplacian operator. It is

$$\nabla^2 \phi = 0. \tag{11.1}$$

This equation occurs frequently in time-independent problems (e.g., what is the steady state temperature distribution, what is the shape of an elastic surface under tension, etc.). The Laplacian itself is non-zero when the variable it operates on is anomalous. Thus, setting the Laplacian to zero implies that there are no anomalies within the field away from the boundaries.

The standard solution method for Laplace's equation is separation of variables. Here is the fast outline of the process, in a two dimensional example. We begin with a very general assumed form for ϕ, then proceed to solve the equation for every term in the sum, then recombine to satisfy the boundary conditions.

$$\phi = \sum_k X_k(x)Y_k(y), \tag{11.2}$$

$$\nabla^2 \sum_k X_k(x)Y_k(y) = 0 \leftrightarrow \nabla^2 X_k(x)Y_k(y) = 0, \tag{11.3}$$

$$\nabla^2 X_k(x)Y_k(y) = \frac{\partial^2 X_k(x)Y_k(y)}{\partial x^2} + \frac{\partial^2 X_k(x)Y_k(y)}{\partial y^2} = Y_k(y)\frac{\partial^2 X_k(x)}{\partial x^2} + X_k(x)\frac{\partial^2 Y_k(y)}{\partial y^2} = 0, \tag{11.4}$$

$$\frac{1}{X_k(x)Y_k(y)}\left[Y_k(y)\frac{\partial^2 X_k(x)}{\partial x^2} + X_k(x)\frac{\partial^2 Y_k(y)}{\partial y^2}\right] = \frac{0}{X_k(x)Y_k(y)}, \tag{11.5}$$

$$\frac{\dfrac{\partial^2 X_k(x)}{\partial x^2}}{X_k(x)} + \frac{\dfrac{\partial^2 Y_k(y)}{\partial y^2}}{Y_k(y)} = 0, \tag{11.6}$$

$$\frac{\dfrac{\partial^2 X_k(x)}{\partial x^2}}{X_k(x)} = -\frac{\dfrac{\partial^2 Y_k(y)}{\partial y^2}}{Y_k(y)} = -k^2. \tag{11.7}$$

The last line ends in the statement that since the left quantity depends only on x and the middle depends only on y, they can only be equal for all x and y if they are both constant. The separation constant k is chosen, and it is written as k^2 for convenience in the next step. The next step is to solve both of the ordinary differential equations in (11.7).

$$X_k(x) = A_{k+}e^{ikx} + A_{k-}e^{-ikx}, \tag{11.8}$$

$$Y_k(y) = B_{k+}e^{ky} + B_{k-}e^{-ky}, \tag{11.9}$$

$$\phi = \sum_k \left[A_{k+}e^{ikx} + A_{k-}e^{-ikx}\right]\left[B_{k+}e^{ky} + B_{k-}e^{-ky}\right]. \tag{11.10}$$

Summing over all of the possible k solutions yields the general solution, so long as the functions that occur in the sum are complete. Complete means that any meaningful function can be approximated as an infinite sum of them. Sines and cosines are complete, which is why Fourier series and transforms work. Exponentials are complete as well which is why Laplace transforms work.

In the example problems, we solved a boundary value problem for Laplace's equation with the boundary conditions: $\phi(L/2, y) = 0$, $\phi(-L/2, y) = 0$, $\phi(x, L) = 0$, $\phi(x, 0) = A\cos(\pi x/L)$. We can

evaluate all of the boundaries like this:

$$\sum_k \left[A_{k+}e^{ikL/2} + A_{k-}e^{-ikL/2}\right]\left[B_{k+}e^{ky} + B_{k-}e^{-ky}\right] = 0, \tag{11.11}$$

$$\sum_k \left[A_{k+}e^{-ikL/2} + A_{k-}e^{ikL/2}\right]\left[B_{k+}e^{ky} + B_{k-}e^{-ky}\right] = 0, \tag{11.12}$$

$$\sum_k \left[A_{k+}e^{ikx} + A_{k-}e^{-ikx}\right]\left[B_{k+}e^{kL} + B_{k-}e^{-kL}\right] = 0, \tag{11.13}$$

$$\sum_k \left[A_{k+}e^{ikx} + A_{k-}e^{-ikx}\right][B_{k+} + B_{k-}] = A\cos(\pi x/L). \tag{11.14}$$

Considering (11.14) first, our understanding of Fourier series leads us to conclude that the only nonzero value of k will be π/L, and since all other boundary conditions (11.11-11.12) are even in x, then $A_{k+} = A_{k-} = A/2$. Including these facts and (11.13-11.14) become

$$A\frac{e^{ikx} + e^{-ikx}}{2}\left[B_{k+}e^{kL} + B_{k-}e^{-kL}\right] = 0 \leftrightarrow B_{k+}e^{kL} + B_{k-}e^{-kL} = 0$$
$$A\frac{e^{ikx} + e^{-ikx}}{2}[B_{k+} + B_{k-}] = A\cos(\pi x/L) \leftrightarrow B_{k+} + B_{k-} = 1$$

Some algebra solves these two equations for the remaining coefficients as

$$\phi = A\cos\left(\frac{\pi x}{L}\right)\frac{e^{\pi(1-y/L)} - e^{\pi(y/L-1)}}{e^{\pi} - e^{-\pi}}. \tag{11.15}$$

We can also use a guess and check method to solve Laplace's equation. The guess and check solution derives from simply noting that since Laplace's equation is linear and homogeneous with constant coefficients, we expect to arrive after separating variables in an equation something like (11.7). That whole class of equations can be solved by assuming solutions of the form

$$\phi = \sum_{m,n} A_{mn}e^{mx}e^{ny}. \tag{11.16}$$

This approach is sometimes called a *normal mode* solution or a *plane wave solution*. What happens when we plug such a guess into Laplace's equation?

$$0 = \nabla^2\phi = \sum_{m,n} A_{mn}\left[\frac{\partial^2 e^{mx}e^{ny}}{\partial x^2} + \frac{\partial^2 e^{mx}e^{ny}}{\partial y^2}\right], \tag{11.17}$$

$$= \sum_{m,n} A_{mn}\left[m^2e^{mx}e^{ny} + n^2e^{mx}e^{ny}\right], \tag{11.18}$$

$$= m^2\left[\sum_{m,n} A_{mn}e^{mx}e^{ny}\right] + n^2\left[\sum_{m,n} A_{mn}e^{mx}e^{ny}\right], \tag{11.19}$$

$$= m^2\phi + n^2\phi. \tag{11.20}$$

$$\therefore 0 = m^2 + n^2. \tag{11.21}$$

Which has precisely the same implications as the single separation constant in (11.7).

11.3 Helmholtz Equation

The Helmholtz equation is closely related to Laplace's equation, but it adds a new constant k^2. It is

$$\nabla^2\phi + k^2\phi = 0. \tag{11.22}$$

It is also a homogeneous, linear, constant coefficient equation, so assuming again the same solution form as in (11.16), we find

$$\phi = \sum_{m,n} A_{mn} e^{mx} e^{ny}, \tag{11.23}$$

$$0 = m^2\phi + n^2\phi + k^2\phi, \tag{11.24}$$

$$\therefore 0 = m^2 + n^2 + k^2. \tag{11.25}$$

So, we can easily arrive at a general solution here, but instead of the coefficients in the exponentials/sines of x being the same as those for y, they are offset by the value of k^2 given.

11.4 Diffusion or Heat Flow

The diffusion or heat flow equation is

$$\nabla^2 T = \frac{1}{\alpha^2}\frac{\partial T}{\partial t}\,. \tag{11.26}$$

One way to attack this problem is by separation of variables.

$$T = \sum_k \phi_k(x,y) T_k(t), \tag{11.27}$$

$$T_k(t)\nabla^2\phi_k(x,y) = \frac{\phi_k(x,y)}{\alpha^2}\frac{\partial T_k(t)}{\partial t}\,, \tag{11.28}$$

$$\frac{\nabla^2\phi_k(x,y)}{\phi_k(x,y)} = -k^2 = \frac{\dfrac{\partial T_k(t)}{\partial t}}{\alpha^2 T_k(t)}, \tag{11.29}$$

$$\nabla^2\phi_k(x,y) + k^2\phi_k(x,y) = 0 \qquad \text{and} \qquad \frac{\partial T_k(t)}{\partial t} + k^2\alpha^2 T_k(t) = 0. \tag{11.30}$$

So, we have two equations left. One is just the Helmholtz equation for ϕ, which we already know how to solve. The other is a first-order, homogeneous, constant-coefficient, linear differential equation for the time variation (which we also know how to solve).

Again, in both cases, we can assume a solution of the form

$$T = \sum_{m,n,\sigma} T_{mn\sigma} = \sum_{m,n,\sigma} A_{mn\sigma} e^{mx} e^{ny} e^{\sigma t}, \tag{11.31}$$

$$m^2 T_{mn\sigma} + n^2 T_{mn\sigma} = \frac{\sigma}{\alpha^2} T_{mn\sigma}, \tag{11.32}$$

$$\therefore \alpha^2(m^2 + n^2) = \sigma. \tag{11.33}$$

We use the boundary conditions to determine which m, n values are possible, and each pair will have a matching σ. The initial conditions will set the final choices of coefficients $A_{mn\sigma}$.

11.4.1 Initial Conditions and Boundary Conditions

Determining the relationship between coefficients and m, n in the general solution to the Helmholtz equation (11.23) that is the spatial part of the heat equation in (11.37) will require *boundary conditions*. Determining the remaining coefficients that come from (11.31) describing the time variations will come from initial conditions.

11.5 Wave Equations

11.5.1 Nondispersive Waves

There are 3 nondispersive wave equations that interest us. We will use the generic variable $\phi(x, y, z, t)$ to represent a wavelike part of the solution. In different kinds of waves, ϕ might be velocity, height above equilibrium, temperature anomaly, vorticity, or pressure. The main nondispersive wave equation is

$$c^2 \nabla^2 \phi = \frac{\partial^2 \phi}{\partial t^2} \, .$$

In simple applications of this equation, c is taken to be a constant and is called the wave *speed*. In more complicated problems, there may be more than one kind of wave speed and it may depend on location (as the medium through which the wave propagates changes, e.g., depth or density of a fluid, from location to location).

The auxiliary wave equations describe waves that only propagate in one direction, written here with only one spatial dimensional in terms of x, t, they are

$$c \frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial t} \, ,$$
$$-c \frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial t} \, .$$

The one-dimensional D'Alembert solutions for the wave equation, already studied in previous chapters, are $f(x - ct)$ and $g(x + ct)$, where f and g are any continuous, differentiable functions. A more general form of the solutions, valid for more than one dimension, is $f(\mathbf{k} \cdot \mathbf{r} - \sigma t)$, $f(\mathbf{k} \cdot \mathbf{r} + \sigma t)$, where now $\mathbf{k}$ is the wavenumber vector (a vector pointing in the direction of propagation of the wave), $\mathbf{r}$ is the position vector relative to the origin, and σ is the angular frequency. The one dimensional form is recovered if the wave is propagating purely in x, in which case $\mathbf{k} = |k|\hat{\mathbf{i}}$, and $c = \sigma/|k|$.

In a real problem, how do we determine the functions f and g? From matching the initial and boundary conditions. Generally, waves that result only from the initial conditions are called "free modes" of the system. Other wave modes are created by variability in the boundary conditions (or even in body forces such as gravity, e.g., tides), these are called "forced" modes. The D'Alembert solutions can easily be matched to initial conditions at all spatial locations, on ϕ and $\frac{\partial \phi}{\partial t}$.

We can also solve the wave equation by separation of variables. In wave terminology, this is often called "normal mode" analysis, because the separation of variable modes over which we sum to get the general solution are the "normal" ones in the system.

The standard, nondispersive, wave equation is very similar to the heat equation in that it can also be split into a Helmholtz in space and a linear, constant coefficient, homogeneous equation in time.

$$T = \sum_k \phi_k(x,y) T_k(t), \tag{11.34}$$

$$T_k(t)\nabla^2\phi_k(x,y) = \frac{\phi_k(x,y)}{c^2}\frac{\partial^2 T_k(t)}{\partial t^2}, \tag{11.35}$$

$$\frac{\nabla^2\phi_k(x,y)}{\phi_k(x,y)} = -k^2 = \frac{\dfrac{\partial^2 T_k(t)}{\partial t^2}}{c^2 T_k(t)}, \tag{11.36}$$

$$\nabla^2\phi_k(x,y) + k^2\phi_k(x,y) = 0 \qquad \text{and} \qquad \frac{\partial T_k(t)}{\partial t} + k^2\alpha^2 T_k(t) = 0. \tag{11.37}$$

11.6 Special Functions of Most Importance

In this section, we introduce some of the special functions that allow our guess and check and separation of variables methods to be more broadly applied.

11.6.1 Solutions of Helmholtz and Laplace

Reading: (Boas, 2006, 12.6)

Many of the coordinate systems we use are useful precisely because they are natural choices in terms of the Helmholtz and Laplace equations. That is, solving these equation is such a common step in applications that we tend to only like coordinate systems where these solutions are easy.

Rectangular Coordinates–Sines

The Helmholtz equation in rectangular coordinates is simply

$$0 = \nabla^2\phi + k^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + k^2\phi. \tag{11.38}$$

As we have seen many times, using separation of variables on this equation leads to exponentials, sines, and cosines as the basis for constructing a solution. These functions are accessible since x, y are separable in this equation. Our normal mode solution is of the form

$$\phi = \sum_{m,n} A_{mn} e^{mx} e^{ny}. \tag{11.39}$$

Which can also be expressed using sines and sinhs, etc., if desired.

The key reason why these functions are useful in series solutions is the same reason why they are useful as bases for series expansions: they have orthogonality relationships allowing us to use

Fourier's trick. Recall these relationships from our work on Fourier series:

$$\langle \sin mx \cos nx \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin mx \cos nx \, \mathrm{d}x = 0. \tag{11.40}$$

$$\langle \sin mx \sin nx \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin mx \sin nx \, \mathrm{d}x = \begin{cases} 0, & \text{if } m = n = 0, \\ 0, & \text{if } m \neq n, \\ \frac{1}{2}, & \text{if } m = n \neq 0. \end{cases} \tag{11.41}$$

$$\langle \cos mx \cos nx \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos mx \cos nx \, \mathrm{d}x = \begin{cases} 1, & \text{if } m = n = 0, \\ 0, & \text{if } m \neq n, \\ \frac{1}{2}, & \text{if } m = n \neq 0. \end{cases} \tag{11.42}$$

$$\sin(a+b) = \sin a \cos b + \sin b \cos a, \tag{11.43}$$

$$\cos(a+b) = \cos a \cos b - \sin b \sin a. \tag{11.44}$$

The complex exponential equivalent averages also come in handy.

$$\langle e^{imx} e^{inx} \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{imx+inx} \, \mathrm{d}x = \begin{cases} 0, & \text{if } m \neq -n, \\ 1, & \text{if } m = -n. \end{cases} \tag{11.45}$$

Cylindrical Coordinates–Bessel

Reading: (Boas, 2006, 12.12)

In cylindrical coordinates, the Helmholtz equation is,

$$0 = \nabla^2 \phi + k^2 \phi = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \phi}{\partial \theta^2} + \frac{\partial^2 \phi}{\partial z^2} + k^2 \phi \tag{11.46}$$

In θ, z, this equation is not much more difficult than the Cartesian one. However, in r, the first term is quite complicated. It is linear, but the coefficients are not constant.

The r equation is the tough one, and with a little work it can be made equivalent to

$$x^2 Z'' + xZ' + \left[x^2 - \left(n + \frac{1}{2} \right) \right] Z = 0. \tag{11.47}$$

The solutions to this equation are Bessel functions.

Bessel functions have a more complicated orthogonality relationship than sines and cosines and complex exponentials, but it still can be employed for a Fourier's trick-like method. It can be expressed on the interval from 0 to 1 as

$$\langle x J_m(ax) J_m(bx) \rangle = \int_0^1 x J_m(ax) J_m(bx) \, \mathrm{d}x = \begin{cases} 0, & \text{if } a = b = 0, \text{ and } m \neq 1. \\ 0, & \text{if } a \neq b, \\ \frac{1}{2} J_{m+1}(a), & \text{if } a = b \neq 0. \end{cases} \tag{11.48}$$

Spherical Coordinates–Legendre

Reading: (Boas, 2006, 12.2, 12.10)

As in the cylindrical coordinates case, the Laplacian takes on a different character in the spherical coordinate system.

$$0 = \nabla^2\Phi + k^2\Phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\Phi}{\partial\phi^2} + k^2\Phi \tag{11.49}$$

The r equation is the tough one, and with a little work separation of variables comes through. The solutions to this equation are the associated Legendre functions, $P_l^m(x)$ (Boas, 2006, 12.10).

The associated Legendre functions have an orthogonality relationship on both m and l over the interval -1 to 1, which is

$$\langle P_l^m(x)P_{l'}^m(x)\rangle = \int_{-1}^{1} P_l^m(x)P_{l'}^m(x)\,\mathrm{d}x = \begin{cases} 0, & \text{if } l \neq l', \\ \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}, & \text{if } l = l'. \end{cases} \tag{11.50}$$

$$\left\langle \frac{P_l^m(x)P_l^{m'}(x)}{1-x^2} \right\rangle = \int_{-1}^{1} \frac{P_l^m(x)P_l^{m'}(x)}{1-x^2}\,\mathrm{d}x = \begin{cases} 0, & \text{if } m \neq m', \\ \frac{(l+m)!}{m(l-m)!}, & \text{if } m = m'. \end{cases} \tag{11.51}$$

11.6.2 Leading Edge of a Wave: The Airy Function

In the Laplace problem, we found oscillatory and exponential solutions solving, respectively,

$$\phi'' + k^2\phi = 0, \tag{11.52}$$

$$\phi'' - k^2\phi = 0. \tag{11.53}$$

Sometimes, we arrive at problems where there is a transition from oscillatory to exponential behavior in a particular location (the leading edge of a Tsunami is one). In this case, the Airy function is interesting. It solves the simplest such problem which is

$$\phi'' - x\phi = 0 \tag{11.54}$$

Thus, for all negative x, it is oscillatory and for all positive x is is exponential. This function is a useful guide in what to expect under such circumstances.

11.6.3 Inhomogeneous Solutions–The Green Function Approach

Reading: (Boas, 2006, 13.8)

We have intensively studied the homogeneous equations featuring the Laplacian: Laplace, Wave, Heat, & Helmholtz. However, what about the innocent-looking Poisson equation?

$$\nabla^2\phi = f(x, y, z) \tag{11.55}$$

This problem seems as though its solutions would depend intimately on the choice of the function f. However, there is an approach, called the Green function approach, which builds on the linearity of the Laplacian.

Suppose that we could break down f into a combination of simpler functions (e.g., point sources). Then, we could add all of these together to find a solution to the full function f. This is the essential idea behind the formula (Boas, 2006, (8.9)),

$$\phi(x,y,z) = \frac{-1}{4\pi} \iiint \frac{f(x',y',z')}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}} \,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}z' \tag{11.56}$$

which is a general solution to (11.55) for a localized region of nonzero $f(x,y,z)$ in an unbounded domain. In bounded domains, or when $f(x,y,z)$ extends to infinity in one or more directions, this solution must be adapted.

The Green function is the function that results from a point source, or

$$\nabla^2 G(\mathbf{r};\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \tag{11.57}$$

Where I use the Dirac delta symbol, δ, to express the following idea,

$$\iiint f(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}z' = f(\mathbf{r}). \tag{11.58}$$

That is, the Dirac delta is a special function-like object that is nonzero (and infinite) at only one location. It is exactly the right magnitude to select the value of a function at only that location when an integral containing that location is performed. In the Poisson equation case, it allows us to consider "point sources" with δ times f that sum up to the full f over all space when integrated, or

$$\phi(\mathbf{r}) = \iiint G(\mathbf{r},\mathbf{r}')f(\mathbf{r})\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}z' \tag{11.59}$$

Note that this form results from integrating both sides of (11.55).

11.7 Example Problems

11.7.1 Application

Example 11.1 *Problem 13.1.4 of Boas (2006). Obtain the heat flow equation (1.3) as follows: The quantity of heat Q flowing across a surface is proportional to the normal component of the (negative) temperature gradient, $(?\nabla T)\cdot\hat{n}$. Compare Chapter 6, equation (10.4), and apply the discussion of flow of water given there to the flow of heat. Thus show that the rate of gain of heat per unit volume per unit time is proportional to $\nabla\cdot\nabla T$. But $\partial T/\partial t$ is proportional to this gain in heat; thus show that T satisfies (1.3).*

Consider any volume that is fixed in time. According to the statement of the problem, the heat flow through the sides of the volume will be proportional to the opposite of the gradient of temperature normal to the surface. Thus, the heat flowing in will be proportional to

$$\oiint \nabla T\cdot\hat{\mathbf{n}}\,\mathrm{d}A = \iiint \nabla\cdot\nabla T\,\mathrm{d}V = \iiint \nabla^2 T\,\mathrm{d}V.$$

where the equality is due to the divergence theorem. The rate of change of heat in the volume is proportional to

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint T\,\mathrm{d}V = \iiint \frac{\partial T}{\partial t}\,\mathrm{d}V.$$

The derivative can come inside of the integral by Leibniz's rule, since the bounds of the volume are fixed in time. Now, we equate the rate of change of heat inside to the flow of heat inward times a (positive) constant to find

$$\iiint \frac{\partial T}{\partial t}\,\mathrm{d}V = \alpha^2 \iiint \nabla^2 T\,\mathrm{d}V$$

Because the volume we chose was arbitrary, the only way this relationship can hold regardless of which volume we choose is if the integrands themselves are equal, thus

$$\frac{\partial T}{\partial t} = \alpha^2 \nabla^2 T$$

Q.E.D.

11.7.2 Scheming Schematics and Articulate Analysis

Example 11.2 *Problem 13.2.5 of Boas (2006). You do not need to make a computer plot of your results.*

Show that the solutions of (2.5) can also be written as and the other sides are at $0°$. $X = (e^{\pm ikx}, Y =$ $\sinh ky$ *or* $\cosh ky$ *Also show that these solutions are equivalent to (2.7) if* k *is real and equivalent to (2.18) if* k *is pure imaginary. (See Chapter 2, Section 12.) Also show that* $X = sink(x?a)$,$Y =$ $sinhk(y?b)$ *are solutions of (2.5).*

The solutions to (2.5) can be either of

$$X = A_+ e^{ikx} + A_- e^{-ikx},$$
$$X = B_s \sin kx + B_c \cos kx.$$

As differentiating any of these four functions satisfies $X'' = -k^2 X$ and each pair is linearly independent. We can convert from one to the other using the Euler relation, e.g., $A_+ e^{ikx} = A_+ \cos kx + A_+ i \sin kx$, and the conversions can be written in matrix form as:

$$\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} A_+ \\ A_- \end{bmatrix} = \begin{bmatrix} B_s \\ B_c \end{bmatrix},$$
$$\begin{bmatrix} A_+ \\ A_- \end{bmatrix} = \begin{bmatrix} -i/2 & 1/2 \\ i/2 & 1/2 \end{bmatrix} \begin{bmatrix} B_s \\ B_c \end{bmatrix}.$$

Therefore, any solution expressible with A_+, A_- can be converted to one in B_s, B_c and vice versa. The real exponential part of the solution

$$Y = C_+ e^{kx} + C_- e^{-kx},$$
$$Y = D_s \sinh kx + D_c \cosh kx.$$

either pair is a complete solution as differentiating any of these four functions satisfies $Y'' = k^2 Y$ and each pair is linearly independent. Again, we can express the conversions as a linear matrix problem:

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} C_+ \\ C_- \end{bmatrix} = \begin{bmatrix} D_s \\ D_c \end{bmatrix},$$
$$\begin{bmatrix} C_+ \\ C_- \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} D_s \\ D_c \end{bmatrix}.$$

If k is pure imaginary, the $e^{ikx} = e^{-}|k|x$ and $e^{-ikx} = e^{|}k|x$. Thus, the general solution with A_+, A_- and B_+, B_- is just the same, except the oscillations have been switched from x to y, and the names of the coefficients differ.

$$X = \sin k(x-a),$$
$$Y = \sinh k(y-b).$$

Are also solutions to $X'' = -k^2X$, $Y'' = k^2Y$ as differentiation shows. In fact, they are the general solutions, as

$$X = A\sin k(x-a) = \frac{Ae^{-ika}e^{ikx} - Ae^{ika}e^{-ikx}}{2i},$$
$$Y = \sinh k(y-b) = \frac{Ae^{-kb}e^{ky} - Ae^{kb}e^{-ky}}{2}.$$

By choosing a, b, A cleverly, any of the functions above can be recreated.

Example 11.3 *Define, in your own words: a) field equation, b) boundary condition, c) initial condition, d) general solution, e) particular solution, f) normal mode, g) forced mode.*

Definition 11.1 *Field Equation: In a spatio-temporal PDE, the field equation is the equation that applies throughout the interior of a region for all times.*

Definition 11.2 *Boundary Condition: In a spatio-temporal PDE, a boundary condition is an equation that applies only on the boundary of a region for all times. It generally reduces the set of solutions that satisfy the field equation alone. When combined with an initial condition, a unique solution can be sought.*

Definition 11.3 *Initial Condition: In a spatio-temporal PDE, an initial condition is an equation that applies throughout the region, but only at a given time. It generally reduces the set of solutions that satisfy the field equation alone. When combined with boundary conditions, a unique solution can be sought.*

Definition 11.4 *General Solution: A general solution is the name given to a solution of the field equation, but still possessing coefficients or parameters that will be selected in order to match boundary or initial conditions.*

Definition 11.5 *Particular Solution: A particular solution is the name given to a solution of an inhomogeneous differential equation that satisfies the field equation in one case. It may not satisfy the initial or boundary conditions, which require other degrees of freedom. Considering the deviation from the particular solution typically renders an inhomogeneous field equation into a homogeneous one.*

Definition 11.6 *Normal Mode: An ansatz for a class of solutions, typically for constant-coefficient, linear, homogeneous differential equations. Exponentials, sines and cosines are common in Cartesian normal mode solutions, but spherical harmonics or other orthogonal functions (e.g., Bessel, Airy) might be used as normal modes in other coordinate systems. Often the normal mode solutions are based on solutions of the Helmholtz equation.*

Definition 11.7 *Forced Mode: A particular solution resulting from boundary conditions or external forces applied within the field equations. It is typically not a normal mode solution.*

Example 11.4 *Verify that each of the one-dimensional D'Alembert wave solutions satisfies one of the nondispersive auxiliary wave equations, but not both, while both D'Alembert wave solutions satisfy the general nondispersive wave equation.*

$$
\begin{aligned}
c^2\nabla^2\phi(x-ct) = c^2\phi''(x-ct) &= \frac{\partial^2\phi(x-ct)}{\partial t^2} = -c\,\frac{\partial\phi'(x-ct)}{\partial t} = c^2\phi''(x-ct),\\
c^2\nabla^2\phi(x+ct) = c^2\phi''(x+ct) &= \frac{\partial^2\phi(x+ct)}{\partial t^2} = -c\,\frac{\partial\phi'(x+ct)}{\partial t} = c^2\phi''(x+ct),\\
c\,\frac{\partial\phi(x-ct)}{\partial x} = c\phi'(x-ct) &\neq \frac{\partial\phi(x-ct)}{\partial t} = -c\phi'(x-ct)\\
c\,\frac{\partial\phi(x+ct)}{\partial x} = c\phi'(x+ct) &= \frac{\partial\phi(x+ct)}{\partial t} = c\phi'(x+ct)\\
-c\,\frac{\partial\phi(x-ct)}{\partial x} = -c\phi'(x-ct) &= \frac{\partial\phi(x-ct)}{\partial t} = -c\phi'(x-ct)\\
-c\,\frac{\partial\phi(x+ct)}{\partial x} = -c\phi'(x+ct) &\neq \frac{\partial\phi(x+ct)}{\partial t} = c\phi'(x+ct)
\end{aligned}
$$

Example 11.5 *Find two different particular solutions to the nondispersive wave equation for waves $\phi(x,y,t)$ on an infinite two-dimensional plane (x,y) using D'Alembert's general solutions. The particular solutions a, b in question should satisfy the initial conditions: a) $\phi(x,y,0) = \cos(kx) + \sin(ky)$, $\left.\frac{\partial\phi}{\partial t}\right|t=0 = 0$ and b) $\phi(x,y,0) = 0$, $\left.\frac{\partial\phi}{\partial t}\right|t=0 = \cos(kx)+\sin(ky)$. HINT: Add together a wave propagating in the y direction to a wave propagating in the x direction.*

a) We know that the general solution will be of the form $\phi = f(x-ct) + g(x+ct)$ for waves propagating in x and $\phi = f(y-ct) + g(y+ct)$ for waves propagating in y. Thus, we consider a solution of the form $\phi = \phi_{x+}(x+ct) + \phi_{x-}(x-ct) + \phi_{y+}(y+ct) + \phi_{y-}(y-ct)$, for four arbitrary functions. We need to simultaneously satisfy the initial conditions on ϕ and $\frac{\partial\phi}{\partial t}$, so for a), we have

$$
\begin{aligned}
\phi(x,y,0) = \cos(kx) + \sin(ky) &= \phi_{x+}(x) + \phi_{x-}(x) + \phi_{y+}(y) + \phi_{y-}(y),\\
\left.\frac{\partial\phi}{\partial t}\right|_{t=0} &= 0 = c\phi'_{x+}(x) - c\phi'_{x-}(x) + c\phi'_{y+}(y) - c\phi'_{y-}(y).
\end{aligned}
$$

It is clear that these are solved if

$$
\begin{aligned}
\frac{1}{2}\cos(kx) &= \phi_{x+}(x) = \phi_{x-}(x),\\
\frac{1}{2}\sin(ky) &= \phi_{y+}(y) = \phi_{y-}(y).
\end{aligned}
$$

Which means that the full solution is

$$
\phi = \frac{1}{2}\cos[k(x+ct)] + \frac{1}{2}\cos[k(x-ct)] + \frac{1}{2}\sin[k(y+ct)] + \frac{1}{2}\sin[k(y-ct)].
$$

b) Again, we know that the general solution will be of the form $\phi = f(x-ct) + g(x+ct)$ for waves propagating in x and $\phi = f(y-ct) + g(y+ct)$ for waves propagating in y. Thus, we consider a solution of the form $\phi = \phi_{x+}(x+ct) + \phi_{x-}(x-ct) + \phi_{y+}(y+ct) + \phi_{y-}(y-ct)$, for four arbitrary

functions. We need to simultaneously satisfy the initial conditions on ϕ and $\frac{\partial \phi}{\partial t}$, so for a), we have

$$\phi(x,y,0) = 0 = \phi_{x+}(x) + \phi_{x-}(x) + \phi_{y+}(y) + \phi_{y-}(y),$$
$$\left.\frac{\partial \phi}{\partial t}\right|_{t=0} = \cos(kx) + \sin(ky) = c\phi'_{x+}(x) - c\phi'_{x-}(x) + c\phi_{y+}(y) - c\phi_{y-}(y).$$

It is clear that these are solved if

$$\frac{1}{2c}\cos(kx) = \phi'_{x+}(x) = -\phi'_{x-}(x),$$
$$\frac{1}{2c}\sin(ky) = \phi'_{y+}(y) = -\phi'_{y-}(y).$$

Which means that the full solution is

$$\phi = \frac{1}{2ck}\sin[k(x+ct)] - \frac{1}{2ck}\sin[k(x-ct)] - \frac{1}{2ck}\cos[k(y+ct)] + \frac{1}{2ck}\cos[k(y-ct)].$$

11.7.3 Application

Example 11.6 *The equations of motion for linear, irrotational ($\nabla \times \mathbf{u} = 0$) waves in a uniform depth fluid are, where η is the displacement of the surface from $z = 0$, ϕ is the velocity potential so that derivatives are the horizontal and vertical velocity ($\frac{\partial \phi}{\partial x} = u, \frac{\partial \phi}{\partial z} = w$), D is the depth, and g is the gravitational acceleration* ($9.81\,\mathrm{m\,s^{-2}}$).

$$\nabla^2 \phi = 0 \tag{11.60}$$
$$\frac{\partial \eta}{\partial t} = \frac{\partial \phi}{\partial z} \qquad at\ z = 0 \tag{11.61}$$
$$\frac{\partial \phi}{\partial t} = -g\eta \qquad at\ z = 0 \tag{11.62}$$
$$\frac{\partial \phi}{\partial z} = 0 \qquad at\ z = -D \tag{11.63}$$

a) Assume a normal mode/plane wave form for $\phi(x,y,z,t)$ and $\eta(x,y,t)$ (note that η is at the surface only, so does not depend on z). Show that the condition for nontrivial solutions (remember how to avoid applying Cramer's rule–set the determinant to zero!) is $\sigma^2 = g\kappa \tanh \kappa D$, where σ is the frequency in time of each mode and $\kappa = \sqrt{k^2 + l^2}$ is wavenumber of the wave, and k and l are the wavenumbers in the x and y directions.

We begin by noting that all of the equations are linear, constant-coefficient, homogeneous PDEs. Thus, we guess a solution of the form:

$$\phi = Pe^{kx}e^{ly}e^{mz}e^{i\sigma t},$$
$$\eta = Ee^{kx}e^{ly}e^{i\sigma t}.$$

Note that we have arbitrarily assumed that the coefficient with time will be imaginary. We could find otherwise as we proceed. Indeed, we will find otherwise for m, n. The field equation results in

$$\nabla^2\phi = 0 \to k^2 + l^2 + m^2 = 0 \to k^2 + l^2 = -m^2.$$

Thus, we see that satisfying the field equation means that some of the exponential coefficients have to be imaginary, while others must be real. The symmetry of the problem is such that we expect the x direction and the y direction to be the same (as there are no boundaries in x, y) while the situation in z might be different since there are upper and lower boundaries. A particular worry is the boundary condition at $z = -D$, which suggests that

$$\frac{\partial \phi}{\partial z} = 0 \qquad \text{at } z = -D \rightarrow mPe^{kx}e^{ly}e^{mz}e^{i\sigma t} = 0.$$

The latter relationship would be resolved most easily by $P = 0$ or $m = 0$, but either would be a trivial solution.
a) To avoid this, we note that $k^2 + l^2 = -m^2$ has two roots for m, $m = \pm\kappa$, and add them together.

$$\phi = P_+ e^{ikx}e^{ily}e^{\kappa z}e^{i\sigma t} + P_- e^{ikx}e^{ily}e^{-\kappa z}e^{i\sigma t},$$
$$\eta = Ee^{ikx}e^{ily}e^{i\sigma t}.$$

With these results, the boundary condition equations become

$$i\sigma E = \kappa P_+ - \kappa P_-$$
$$i\sigma P_+ + i\sigma P_- = -gE$$
$$\kappa P_+ e^{-\kappa D} - \kappa P_- e^{\kappa D} = 0$$

Or, in matrix form,

$$\begin{bmatrix} i\sigma & -\kappa & \kappa \\ g & i\sigma & i\sigma \\ 0 & \kappa e^{-\kappa D} & -\kappa e^{\kappa D} \end{bmatrix} \begin{bmatrix} E \\ P_+ \\ P_- \end{bmatrix} = 0$$

To find a nontrivial solution, we set the determinant of the matrix in the last equation to zero, which yields after some algebra,

$$\sigma^2 = g\kappa \tanh \kappa D$$

11.8 Homework Problems

11.8.1 Manipulation

Exercise 11.1 *Problem 13.1.2 of Boas (2006). (a) Show that the expression $u = \sin(x \pm vt)$ describing a sinusoidal wave (see Chapter 7, Figure 2.3), satisfies the wave equation (1.4). Show that, in general, $u = f(x - vt)$ and $u = f(x + vt)$ satisfy the wave equation, where f is any function with a second derivative. This is the d'Alembert solution of the wave equation. (See Chapter 4, Section 11, Example 1.) The function $f(x - vt)$ represents a wave moving in the positive x direction and $f(x + vt)$ represents a wave moving in the opposite direction.*

(b) Show that $u(r, t) = (1/r)f(r - vt)$ and $u(r, t) = (1/r)f(r + vt)$ satisfy the wave equation in spherical coordinates. [Use the first term of (7.1) for $\nabla^2 u$ since here u is independent of θ and ϕ.] These functions represent spherical waves spreading out from the origin or converging on the origin.

Exercise 11.2 *Problem 13.2.10 of Boas (2006). You do not need to make a computer plot of your results.*

Find the steady-state temperature distribution in a metal plate 10 cm square if one side is held at 100° and the other three sides at 0°. Find the temperature at the center of the plate. (The answer, but not the solution method, is written out in Boas).

Exercise 11.3 *Problem 13.3.2 of Boas (2006). You do not need to make a computer plot of your results.*

A bar 10 cm long with insulated sides is initially at $100°$*. Starting at* $t = 0$*, the ends are held at* $0°$*. Find the temperature distribution in the bar at time* t*. (The answer, but not the solution method, is written out in Boas).*

11.8.2 Manipulation

Exercise 11.4 *Write the Helmholtz equation in earth coordinates. Do you think the separable solutions will be the same or different from those found in spherical coordinates? Why or why not?*

11.8.3 Application

Exercise 11.5 *The equations of motion for linear, irrotational (*$\nabla \times \mathbf{u} = 0$*) waves in a uniform depth fluid are, where* η *is the displacement of the surface from* $z = 0$*,* ϕ *is the velocity potential so that derivatives are the horizontal and vertical velocity (*$\frac{\partial \phi}{\partial x} = u, \frac{\partial \phi}{\partial z} = w$*),* D *is the depth, and* g *is the gravitational acceleration* $(9.81\,\mathrm{m\,s^{-2}})$*.*

$$\nabla^2 \phi = 0 \tag{11.64}$$

$$\frac{\partial \eta}{\partial t} = \frac{\partial \phi}{\partial z} \qquad at\ z = 0 \tag{11.65}$$

$$\frac{\partial \phi}{\partial t} = -g\eta \qquad at\ z = 0 \tag{11.66}$$

$$\frac{\partial \phi}{\partial z} = 0 \qquad at\ z = -D \tag{11.67}$$

(Part a) is the example problem already solved above in example 11.6. You don't need to do it again, begin with part b. I repeat it here only to point out that part b builds on this result.) Assume a normal mode/plane wave form for $\phi(x, y, z, t)$ *and* $\eta(x, y, t)$ *(note that* η *is at the surface only, so does not depend on* z*). Show that the condition for nontrivial solutions (remember how to avoid applying Cramer's rule–set the determinant to zero!) is* $\sigma^2 = g\kappa \tanh \kappa D$*, where* σ *is the frequency in time of each mode and* $\kappa = \sqrt{k^2 + l^2}$ *is wavenumber of the wave, and* k *and* l *are the wavenumbers in the* x *and* y *directions.*

b) In very deep water, $\sigma^2 = g\kappa \tanh \kappa D \approx g\kappa$*. Redo the normal mode analysis in a), but in the semi-infinite plane (i.e., infinite depth) ocean.*

c) Compare the deep water case in b) to the nondispersive waves of previous problems. Show that the phase of deep water waves propagates at different speeds depending on the wavenumber, whereas nondispersive waves always have the same speed regardless of κ*.*

11.8.4 Scheming Schematics and Articulate Analysis

Exercise 11.6 *Look at* http://tinyurl.com/mljujml *and* http://tinyurl.com/ol3al47. *a) Contrast these against the separation of variables in the Cartesian coordinate cases. b) Why aren't the solutions sines and cosines? c) How can it matter which coordinate system we choose– that is, what is so special about separable solutions?*

Chapter 12

Continuum Mechanics

12.1 Basics of Continua

All matter is made up of molecules, which are made up of atoms, which are made up of electrons, protons, and neutrons, and so on. However, on large scales, where the separations between molecules can be considered to be tiny, matter behaves as though it is continuous. Such materials are called continua, and the formulation of their equations of motion is called continuum mechanics. In a continuum mechanics, the motion of the individual molecules and sub-molecular particles won't be considered directly, only the motion of the large bodies will. Some molecular effects do affect the larger scales (e.g., viscosity, diffusivity), but only collectively rather than one molecule at a time. Mathematically, assuming a material is a continuum means that we can use all of the power of calculus, differential geometry, and partial differential equations, assuming that taking limits to infinitesimal distances and time scales (i.e., taking derivatives in space and time) is acceptable even though we know that in the real world there is a limit to how far down we can go. We use the term *parcel* to imply an infinitesimal piece (i.e., a differential mass or volume element) of the continuum, which is still larger than the separations of the particles that make up the continuum.

Continuum mechanics includes the study of all states of matter. Thus, fluids (both liquids and gasses) and solids are covered, as well as more esoteric phases such as plasmas. Even for phases where quantum mechanics effects are important (e.g., Bose-Einstein condensates, superfluids) continuum mechanics can still be a useful guide. In geophysics, we also consider multiphase systems, such as liquids propagating through a solid matrix, ice crystals forming into clouds or sea ice, as well as deformable materials such as mineral phases, plastics, and viscoelastics. This chapter will briefly introduce the equations that govern the motions of the simplest such continua.

A pedagogical reason for including continuum mechanics in this course is to further motivate the study of partial differential equations in the last chapter and ordinary differential equations in preceding chapters. There, scenarios–warming a solid, disturbing a liquid, budgets for the earth's energy, or budgets for temperature and salinity in the ocean–were used to motivate study of the heat equation, wave equation, or coupled sets of ordinary differential equations. Here, the relationships between conservation of mass, momentum, and energy and these differential equations will be made plain.

12.2 Conservation

All of the equations of motion of continuua are framed as conservation principles (as indeed are almost all of the equations of physics). The fundamental ones are conservation of mass, momentum, and energy. A few other conserved quantities, such as conservation of angular momentum, vorticity, and potential vorticity can be useful as well.

12.2.1 Displacements

When a continuum moves, it is often useful to think of every point in the continuum as having an original location $\mathbf{X}$ and a present location $\mathbf{x}$. Since every point came from somewhere, we can think of original location $\mathbf{X}$ as a function of space $\mathbf{x}$ and time t, i.e., $\mathbf{X}(\mathbf{x}, t)$. We can also invert this relationship to find present position $\mathbf{x}(\mathbf{X}, t)$, that is the location at present of a point in the continuum that originated at each point $\mathbf{X}$ occupied by the continuum at time $t = 0$.

The displacement from initial position is then

$$\Delta\mathbf{x} = \mathbf{x} - \mathbf{X}. \tag{12.1}$$

We can consider the displacement as a function of present location and time $\Delta\mathbf{x}(\mathbf{x}, t)$, which is called the Eulerian description, or initial location and time $\Delta\mathbf{x}(\mathbf{X}, t)$ which is called the Lagrangian description. Note that writing it this way abuses notation a bit, since the function $\Delta\mathbf{x}$ in $\Delta\mathbf{x}(\mathbf{x}, t)$ can't be the same function as $\Delta\mathbf{x}$ in $\Delta\mathbf{x}(\mathbf{X}, t)$ even though we use the same symbol for it. A slightly more formally correct way to write the second function is $\Delta\mathbf{x}(\mathbf{x}(\mathbf{X}), t)$. The arguments of the function are used to distinguish between the two. The continuum velocity at a point is the rate at which $\Delta\mathbf{x}$ is changing, for a given parcel of continuum (i.e., hold $\mathbf{X}$ constant):

$$\mathbf{v} = \tfrac{\partial\Delta\mathbf{x}}{\partial t}|_{\mathbf{X}\text{ fixed}} = \tfrac{\partial\mathbf{x}}{\partial t} \tag{12.2}$$

There is actually quite a bit of thought that goes into Eulerian versus Lagrangian descriptions. For example, suppose we consider a flow property that is conserved following a parcel, ϕ. It should be determined from only the initial field of ϕ distributed over all space, that is, $\phi = \phi(\mathbf{X}, t)$. We can consider taking the derivative of $\phi(\mathbf{X}, t)$ with respect to time following a fluid parcel,

$$\frac{D\phi(\mathbf{X}, t)}{Dt} = \tfrac{\partial\phi(\mathbf{X},t)}{\partial t}. \tag{12.3}$$

This is the Lagrangian perspective, and we call $\frac{D}{Dt}$ the Lagrangian or *material* derivative, as it is the derivative following the flow of the material–or the derivative with respect to time holding $\mathbf{X}$ fixed. Or, instead, we could think of the value of ϕ as depending on $\mathbf{x}$ (again, implying a different function ϕ), thus

$$\frac{D\phi(\mathbf{x}(\mathbf{X}, t), t)}{Dt} = \tfrac{\partial\phi(\mathbf{x}(\mathbf{X},t),t)}{\partial t} + \tfrac{\partial\mathbf{x}(\mathbf{X},t)}{\partial t} \cdot \nabla\phi(\mathbf{x}(\mathbf{X}, t), t) = \tfrac{\partial\phi(\mathbf{x}(\mathbf{X},t),t)}{\partial t} + \mathbf{v} \cdot \nabla\phi(\mathbf{x}(\mathbf{X}, t), t). \tag{12.4}$$

Here ∇ is the gradient operator with respect to the components of $\mathbf{x}$, and it appears here just by ordinary application of the chain rule.

Notice that at any moment, (12.2) implies that the Eulerian velocity and the Lagrangian velocity are the same everywhere $\mathbf{v}(\mathbf{x}, t) = \mathbf{v}(\mathbf{x}(\mathbf{X}, t), t)$. This is true instantaneously, but a moment later

the parcel initially located at $\mathbf{X}$ will not be at this same location, i.e., $\mathbf{x}(\mathbf{X},t) \neq \mathbf{x}(\mathbf{X},t+\Delta t)$. Likewise, we can think of a Lagrangian-averaged velocity (the average velocity following a parcel trajectory or holding $\mathbf{X}$ fixed) and an Eulerian-averaged velocity (the velocity at a point averaged in time or holding $\mathbf{x}$ fixed).

12.2.2 Mass

In continuum mechanics, the motion of the continuum moves around the mass of the material. The continuity equation ensures that during these motions, no mass is created or destroyed. The continuity equation is often written in two forms. The first form is the "flux form,"

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \tag{12.5}$$

The density of the continuum (mass per unit volume) is ρ and the velocity of the continuum is $\mathbf{v}$. We can integrate this equation over a volume that is fixed in time and space (a *control* volume delineating a range of $\mathbf{x}$) to find

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho \,\mathrm{d}V + \int_V \nabla \cdot (\rho \mathbf{v}) \,\mathrm{d}V = \frac{\mathrm{d}}{\mathrm{d}t} M + \oint_S \rho \mathbf{v} \cdot \hat{\mathbf{n}} \,\mathrm{d}S = 0. \tag{12.6}$$

The integral of the density becomes the mass contained within the volume (M), and using the divergence theorem allows us to understand that $\rho\mathbf{v}$ represents the mass per unit area flowing into and out of the control volume. Notice that the divergence theorem (8.69) is the key mathematical tool allowing us to express the concept of "stuff flowing in and out" versus "accumulation of stuff inside."

Often, the conservation of mass is taken in different forms. One common form is to expand out the divergence and recombine the terms,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = \underbrace{\frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho}_{\frac{D\rho}{Dt}} + \rho \nabla \cdot \mathbf{v} = \frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0. \tag{12.7}$$

Here the *material derivative* $\frac{D\cdot}{Dt} \equiv \frac{\partial \cdot}{\partial t} + \mathbf{v} \cdot \nabla \cdot$ appears. This derivative has the meaning of the rate of change of a continuum property along a path following along with the motion of the continuum. Such a perspective is often called *Lagrangian*, while the idea of sitting still at a point and letting the continuum pass by is called the *Eulerian* perspective. Note that the change of perspective is mathematically just an application of the chain rule. The total derivative with time of a property $f(x,y,z,t)$ that moves with the continuum depends both on the partial derivative with respect to time and the partial derivatives with respect to space:

$$\frac{df(x,y,z,t)}{dt} = \frac{\partial f(x,y,z,t)}{\partial t} + \left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial z}{\partial t}\right) \cdot \nabla f(x,y,z,t), \tag{12.8}$$

$$\frac{Df(x,y,z,t)}{Dt} = \frac{\partial f(x,y,z,t)}{\partial t} + (u,v,w) \cdot \nabla f(x,y,z,t), \tag{12.9}$$

$$\frac{Df(x,y,z,t)}{Dt} = \frac{\partial f(x,y,z,t)}{\partial t} + \mathbf{v} \cdot \nabla f(x,y,z,t). \tag{12.10}$$

What distinguishes the material derivative from the total derivative, is that in a material or continuum, the way that the continuum varies in space with respect to time (e.g., $\frac{\partial x}{\partial t}$) is governed solely

by the velocity of the continuum's motion $(u, v, w) = \mathbf{v}$. Some authors like to use $\frac{\mathrm{d}}{\mathrm{d}t}$ to denote the material derivative, but this glosses over the fact that other propagation through space would be determined by a similar law, but would not depend on the motion of the continuum. For example, the propagation of light through a transparent gas obeys the following

$$\frac{df(x,y,z,t)}{dt} = \frac{\partial f(x,y,z,t)}{\partial t} + \left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial z}{\partial t}\right) \cdot \nabla f(x,y,z,t) = \frac{\partial f(x,y,z,t)}{\partial t} + \mathbf{c} \cdot \nabla f(x,y,z,t). \tag{12.11}$$

However, the speed of light $\mathbf{c}$ doesn't depend much on the motion of the gas, so $\mathbf{c} \neq \mathbf{v}$, so using $\frac{D}{Dt}$ would be inappropriate for light, while using $\frac{\mathrm{d}}{\mathrm{d}t}$ remains appropriate.

We might also be interested in the conservation of mass of only a constituent of the continuum, such as the salt in seawater or humidity in air. Well-mixed constituents such as these mostly travel along with the continuum, so do not require a different velocity. The equation of continuity for a well-mixed constituent is just

$$\frac{\partial \rho c}{\partial t} + \nabla \cdot (\rho c \mathbf{v}) = \nabla \cdot \mu \nabla c. \tag{12.12}$$

Here c is the mass fraction of the constituent's mass to the continuum mass in a control volume. Thus, the constituent mass per unit volume is just ρc. There is now a dynamic diffusivity (μ, dimensions M/LT) of the constituent added to the right hand side. Sometimes, a kinematic diffusivity ($\kappa = \mu/\rho$, dimensions L^2/T) is used if density variations are small. Here ρc is the mass of tracer per unit volume, and $\rho c \mathbf{v}$ is the *advective flux* of tracer. This *flux form* of the equation is useful for tracers measured in quantity per unit volume.

We can distribute out the derivatives and use the continuity equation to make another form of the well-mixed tracer equation, assuming negligible diffusivity

$$\frac{\partial \rho c}{\partial t} + \nabla \cdot (\rho c \mathbf{v}) = c\frac{\partial \rho}{\partial t} + \rho \frac{\partial c}{\partial t} + c \nabla \cdot (\rho \mathbf{v}) + \rho \mathbf{v} \cdot \nabla c \approx 0, \tag{12.13}$$

$$c\underbrace{\left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v})\right)}_{0} + \rho \left(\frac{\partial c}{\partial t} + \mathbf{v} \cdot \nabla c\right) \approx 0, \tag{12.14}$$

$$\frac{Dc}{Dt} \approx 0. \tag{12.15}$$

Thus, the material derivative is a useful way to think about the transport of mass fractions of well-mixed quantities being transported by a continuum.

Liquids, unlike gasses, tend to be difficult to compress. Thus, their density changes very little while flowing. We can express incompressibility as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = \underbrace{\frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho}_{\frac{D\rho}{Dt}} + \rho \underbrace{\nabla \cdot \mathbf{v}}_{0} = 0, \tag{12.16}$$

$$\frac{D\rho}{Dt} = 0 \text{ and } \nabla \cdot \mathbf{v} = 0. \tag{12.17}$$

A fully incompressible continuum has density that is constant for each fluid parcel, although density may vary among fluid parcels. In the more interesting geophysical cases, the density of a fluid parcel may vary, but only by a small fraction of a background value of the density. In this approximation, called the Boussinesq approximation, we can still assume $\nabla \cdot \mathbf{v} = 0$ as our continuity equation, but forces resulting from variations in density are not neglected.

Incompressibility also has a simple form when the displacements are taken to be the variable, rather than the velocities. Since each displacement is just the time integral of each velocity, incompressibility becomes just

$$\rho(\mathbf{X}, 0) = \rho(\mathbf{X}, t) \text{ and } \nabla \cdot \Delta\mathbf{x} = 0. \tag{12.18}$$

12.2.3 Momentum

The fundamental equation of motion of continuua is the Cauchy Equation of Motion, which is the continuum version of Newton's Second Law ($\mathbf{F} = m\mathbf{a} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}$) for the conservation of momentum $\mathbf{p}$. If we consider the force and momentum per unit volume over a control volume, we arrive at

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho\mathbf{v}\,\mathrm{d}V = \int_V \frac{\partial}{\partial t}\rho\mathbf{v}\,\mathrm{d}V + \oint_S \rho\mathbf{v}\wedge\mathbf{v}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \int_V \rho\mathbf{g}\,\mathrm{d}V + \oint_S \sigma\cdot\hat{\mathbf{n}}\,\mathrm{d}S \tag{12.19}$$

The momentum per unit volume of a continuum is just $\rho\mathbf{v}$. Using the divergence theorem and continuity equation results in

$$\frac{\partial\rho\mathbf{v}}{\partial t} + \nabla\cdot(\rho\mathbf{v}\wedge\mathbf{v}) = \rho\frac{D\mathbf{v}}{Dt} = \nabla\cdot\sigma + \rho\mathbf{g}. \tag{12.20}$$

In this form, it is easy to make the association between a control mass and the normal mass times acceleration: $\int_{V_M}\rho\frac{D\mathbf{v}}{Dt}\,\mathrm{d}V = m\mathbf{a}$.

Here body forces per unit volume, such as gravity, are included in $\rho\mathbf{g}$. Body forces are usually conservative, in which case we might write them as the gradient of a potential, $\rho\mathbf{g} = -\nabla\Phi$. Forces that result from stresses across interfaces within the continuum result in the *stress tensor* σ. The rate of change of momentum is on the left side, both as a flux form for momentum per unit volume $\rho\mathbf{v}$ and as an acceleration following a parcel per unit volume $\rho\frac{D\mathbf{v}}{Dt}$.

The stress tensor is probably unfamiliar, so it deserves a few words. Recall that we said that a tensor is an object like a scalar, a vector, or a matrix, but with the special property that under changes of coordinates it behaves the way we expect it to. Examine the momentum equation above, we take the divergence of the stress tensor, and the result is a vector. Since the divergence involves a dot product which combines three directions, just as $\nabla\cdot\mathbf{v}$ produces a scalar from a vector $\mathbf{v}$, that means that the stress tensor must have more directions in it than a vector. In fact, it's like a matrix, in that you can take its dot product from either the left or the right! Let's write out one example of the stress tensor for concreteness, the Reynolds stress tensor which results from averaging over a turbulent flow (we'll use overbar to denote averaging). The Reynolds stress tensor is (in the absence of background flow):

$$\sigma^R = -\rho\overline{\mathbf{v}\wedge\mathbf{v}}, \tag{12.21}$$

$$\sigma^R_{ij} = -\rho\overline{\mathbf{v}_i\mathbf{v}_j}. \tag{12.22}$$

The wedge product here implies that we consider nine different things, the x component of the velocity times all of the other components, the y component times all of the other components, etc. This forms a matrix, which we can see more clearly using the index notation as shown on the second line. The stress tensor is a second-rank tensor (i.e., it is like the wedge product of two vectors, or a square matrix) and it is always symmetric $\sigma_{ij} = \sigma_{ji}$. This is obviously true for the Reynolds stress tensor, where the two velocities are just the velocity of the fluid and hence interchangeable. In general, the symmetry of the stress tensor results in the conservation of angular momentum.

Three important examples of dynamic stress tensors will be given below: one for elastic solids, one for Newtonian incompressible liquids, and one for Newtonian gasses. The former resists deformation, and the latter resists the rate of deformation. Much of the business of continuum mechanics, especially solid mechanics, is involved in finding a good approximation of the stress tensor for the material at hand. In both fluids and solids, however, there is a simple form of the static stress tensor for undeformed solids or inviscid fluids, which is sometimes called the perfect form. It is

$$\sigma = -p\mathbf{I} \tag{12.23}$$

The equations of motion in this case become

$$\rho\frac{D\mathbf{v}}{Dt} = -\nabla p + \rho\mathbf{g}, \tag{12.24}$$

$$\tfrac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0. \tag{12.25}$$

These are called the Euler equations of motion for perfect (i.e., inviscid and diffusionless) fluids.

Often, the combination of the conservation of mass and conservation of momentum are not enough to fully describe the motion of a continuum. For example, let's count up variables and equations. We have the continuity equation and the momentum equation for three components of the acceleration, thus four equations. We have the unknowns ρ, p and three components of velocity: five! So, we need another equation to close the set.

One way to close the set is to assume incompressibility, because in this case the one continuity equation (12.16) becomes two separate equations (12.17), for five equations in five unknowns. However, even in this simplified case of perfect, incompressible fluids, the fluxes are all nonlinear in the unknowns, so we expect it to be difficult to solve these equations!

12.3 Angular Momentum

The conservation of linear momentum times a moment arm $\mathbf{r}$ is also conserved.

$$\tfrac{\mathrm{d}}{\mathrm{d}t}m(\mathbf{r}\times\mathbf{v}) = \mathbf{r}\times\mathbf{F}. \tag{12.26}$$

Taking the time derivative following a constant mass, this is

$$\cancelto{0}{m(\mathbf{v}\times\mathbf{v})} + m\mathbf{r}\times\mathbf{a} = \mathbf{r}\times\mathbf{F}. \tag{12.27}$$

Above, in the continuum case, we split the forces acting on a mass into a body force $\rho\mathbf{g}$ and a surface stress σ. In order to satisfy the conservation of angular momentum, it turns out that σ must be symmetric. Otherwise, the continuum version of the preceding equation gains an extra term, which is like the cross-product of σ with itself.

12.4 Solids

The simplest stress tensor for a solid relates the stress to the deformation of the solid. Such a relationship is called a constitutive relationship. Suppose, like a spring, a solid has a stress-free

initial location of all of the parcels $\mathbf{X}$. We might consider all $\Delta\mathbf{x}$ displacements from this initial state. If all deviations produce a stress proportional to the gradient of this deformation, i.e., proportional in any given direction to $\nabla_i \Delta x_j$. This is almost a complete form, except we need a symmetric tensor if we want to produce a stress tensor, so how about

$$\sigma_{ij} = -p\delta_{ij} + S\frac{1}{2}\left(\nabla_i \Delta x_j + \nabla_j \Delta x_i\right) = -p\delta_{ij} + S\epsilon_{ij}. \tag{12.28}$$

The strain tensor *epsilon* is symmetric by construction. However, it turns out that this is not a very good model for a solid, because not all kinds of strain are the same as the others, and so they do not produce the same amount of stress. Thus, we can consider a generalized relationship, where

$$\sigma_{ij} = -p\delta_{ij} + S_{ijkl}\frac{1}{2}\left(\nabla_k \Delta x_l + \nabla_l \Delta x_k\right) = -p\delta_{ij} + S_{ijkl}\epsilon_{kl}. \tag{12.29}$$

Now the proportionality can be different in each direction.

12.5 Liquids: Incompressible Fluids

Incompressible Newtonian fluids have the stress tensor which looks much the same, but there is a velocity in the place of the displacement. So, fluids don't resist displacement, they resist the rate of displacement.

$$\sigma_{ij} = -p\delta_{ij} + 2\mu\frac{1}{2}\left(\nabla_i v_j + \nabla_j v_i\right) = -p\delta_{ij} + 2\mu D_{ij}. \tag{12.30}$$

Here we have assumed that the liquid is incompressible. The difference between a liquid and a solid, then, is that the liquid resists the rate of strain D_{ij} with stresses while a solid resists the strain itself ϵ_{ij} with stresses. Both have a static pressure p.

12.6 Gasses: Compressible Fluids

A gas expands to fill its container, and is also compressible if the size of the container changes. Therefore, we need to preserve the full form of the continuity equation, which allows for changing mass within a given volume.

Additionally, there is the possibility that the gas will viscously oppose the compression. So, a general linear form for the stress tensor is

$$\sigma_{ij} = (-p + \lambda D_{kk})\delta_{ij} + 2\mu D_{ij}. \tag{12.31}$$

The reason for the new term appears since $D_{kk} = \nabla_k v_k$, which is zero for an incompressible liquid. In a gas, this term need not disappear. It is often assumed that $\lambda = -\frac{2}{3}\mu$, which avoids any appearance of the viscous terms in a motionless gas.

12.6.1 Energy and Entropy

Let's review the situation with unknowns and equations in the compressible equations. The momentum equation has three unknown velocity components and three equations (one for each direction):

$$\rho \frac{D\mathbf{v}}{Dt} = -\nabla p + \rho \mathbf{g}. \tag{12.32}$$

However, it also introduces the density and pressure, which are two other unknowns. The continuity equation introduces no new unknowns:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \tag{12.33}$$

It is only one constraint. Thus, we have four equations in five unknowns.

If incompressibility is not assumed, then additional constraints, such as the First Law of Thermodynamics (conservation of energy) or the Second Law of Thermodynamics are required. The internal energy and kinetic energy of a continuum generally flow along with the motion of the continuum, but potential energy depends on position, and the motion of the continuum itself may represent work being done. Dissipation of continuum energy, through viscosity or diffusion, really represents an exchange of internal energy (the energy of the molecular motions within the continuum) and the kinetic energy of the continuum's motion.

The Second Law of Thermodynamics (conservation or increase of entropy) can also be a useful thermodynamic constraint on continuum mechanics. Often the flows of entropy in a system are categorized into reversible (entropy-conserving) and irreversible (entropy-increasing) phenomena. Like the conservation of energy, care is required to account for the dissipative effects which act to make continuum mechanics irreversible. If we write the equation of motion for the entropy of a continuum parcel, and assume that all processes are perfectly isentropic (dissipationless, inviscid, reversible), then the entropy s just obeys

$$\frac{Ds}{Dt} = 0. \tag{12.34}$$

This adds another equation to our set, but also another variable (the entropy). However, for many materials, there are known relations among the thermodynamic variables, such as entropy, density, and pressure, called the equation of state. So, if we add in the equation of state for a perfect fluid, where entropy is just a function of density and pressure, then

$$s = s(\rho, p). \tag{12.35}$$

Now we have six equations in six unknowns.

Furthermore, solids, fluids, and gasses are significantly different in the way they respond to changes in energy and entropy, which is governed by the equation of (thermodynamic) state of the continuum. Technically, most continuum flows are not in thermodynamic equilibrium (e.g., because there are gradients in temperature, entropy, etc., throughout the continuum). However, in a similar spirit to the continuum assumption, continuua are usually assumed to be in *local* thermodynamic equilibrium, which means that averaging over the molecular motions in a hunk of continuous material yields a meaningful definition of temperature, energy, etc., of that hunk.

Chapter 13

Statistics & Probability

13.1 Introduction–Probably a Stat for That

Reading: Boas (2006, 15.1)

So far, almost everything we've worked on has been deterministic. That is, there is one, and only one answer being sought for each mathematical problem. Sometimes, as in the case of truncated series, our answer was only an approximation to the one answer. Sometimes, a series would not converge and there would be no meaningful answer. Sometimes, as in the case of some sets of linear equations, there was no consistent answer or there was an infinite number of possible choices of answers. Even when we studied chaos, it was still *deterministic* chaos where there was an exact (but complicated to determine or estimate) answer to each question.

Statistics and probability are topics that confront a different kind of mathematical problem, those where it is not possible to choose one answer over others from a set of possibilities, or where measurement accuracy or imperfect models prevent an exact answer from being possible. The mindset and approach to these topics is quite different from that of deterministic mathematics, and many people have a strong preference for one or the other. As geoscientists, we have an interest in familiarity with both, as deterministic methods are the tools for developing theories and models, while statistics and probability are the mathematics of experiment.

13.1.1 Probability

Assessing probability is the science (and sometimes art) of determining how likely a given outcome is from the total set of possible similar outcomes.

Definition 13.1 (Probability) *The ratio of the number of times a particular outcome is expected to the number of all possible total outcomes, where each of the number is equally likely and mutually exclusive.*

Many people who enjoy probabilistic thinking enjoy games of chance and gambling, and many clean example problems to illustrate probabilistic thinking are framed in these types of games.

13.1.2 Statistics

Statistics are manipulations of observations or other random or uncertain data to complement and inform our understanding of their probability. A statistic, a metric, a unit, a dimension, and a measurement are all related concepts involved in the gathering and generation of data. A statistic differs in that it need not involve new observations to be created, it may just be a manipulation of existing metrics or measurements.

13.1.3 A Bunch of Randos

Definition 13.2 (Rando) *An unpredictable, awkward and often creepy individual. a rando is always there in the background, even though no one invited him. Randos tend to be isolated within large groups; they sit in corners and lurk in the shadows of others. Even though everyone is aware of the presence of a rando, no one knows anything about him or her and therefore randos are often unwittingly ignored. Source: The Urban Dictionary.*

There are many closely related applications of probabilistic and statistical mathematics. Something is random if it can be repeatedly sampled with a variety of different outcomes. Probability is the quantification of the likelihood of those outcomes. Some topics of recent mathematical investigation are stochastic dynamics, which is similar to dynamical systems but involve random variables and are therefore nondeterministic. Fuzzy math is a similar idea, except for logic of random processes rather than dynamical systems. Uncertainty is the quite the same as randomness, in that it describes how measurable something is rather than intrinsic randomness. However, we often associate repeated measurements under uncertainty as being usefully modeled by random processes. The Copenhagen interpretation of quantum mechanics has made the unknowable a part of physical sciences, but the idea of unknowable truths is much older in nonscientific realms such as religion. The methods of statistics and probability in quantum mechanics allow quantification of the unknowable even without there being a knowable variable underlying the probability, this odd situation is partly a consequence of the peculiarities of quantum mechanics but mainly a demonstration of the power of probabilistic thinking. In geosciences, where many processes are poorly modeled because of limited understanding or limitations of computers or measurements, probabilistic thinking can be used to supplement models. The limited capability of models can have its "uncertainty quantified," and stochastic models or stochastic parameterizations are sometimes used to represent incomplete theories even in otherwise deterministic models. Deterministic chaos can also be tamed somewhat using probabilistic thinking, hence we have ensemble forecasts of the weather where many different versions of deterministic or stochastic weather models are run simultaneously to try and map the probabilities of different outcomes. Finally, much of geosciences, particularly planetary sciences and oceanography, is grossly undersampled because of the exceptional cost per measurement. Probabilistic methods can be used to better understand how representative limited observations are likely to be of the whole system were it possible to sample it better.

13.2 Histogram and Probability Density Function

Reading: Boas (2006, 15.3) A histogram is a method for mapping the estimated probability based on a set of repeated measurements which are thought to be drawn from the same set of potential outcomes. It is easy to create a histogram. One just makes a chart where the x axis gives a different

location for each outcome and as the data is reported the y axis accumulates the number of times that outcome occurs. Fig. 13.1 shows two examples, one where x is numerical and one where it is not. In both cases, the set of possibilities is organized into a discrete set of bins–even the numerical example is bunched by tens. This discretization is always true of histograms, since there is only a finite number of observations available.

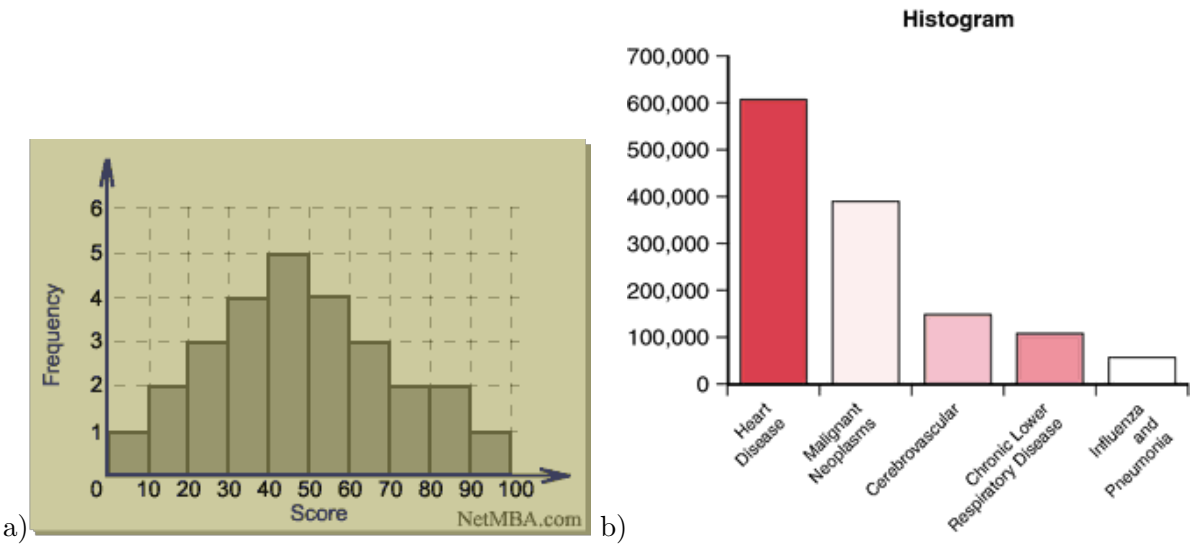


Figure 13.1: Two examples of histograms from the web.

Sometimes, the y axis of a histogram is *normalized*, that is, it is divided by the total number of occurrences. If this is true, then the sum of all of the histogram columns is one, and each column is less than or equal to one. The reason one might normalize the histogram is to better study the shape of the histogram, which is independent of the number of observations, instead of a particular history of observations.

A *probability distribution function* (PDF) is an idealization of a normalized histogram. If we could take an infinite number of observations, what shape would the normalized histogram take? Or, if we take only a limited number of samples from such an idealized distribution, the histogram would be the approximation to this distribution based on a limited sample. Also, when we consider an infinite number of observations, we can make the bins of the histogram infinitely thin. Thus, we can begin to think of the ordinate (x) in the histogram as a continuous, rather than discrete-valued, variable. In this case, the PDF is called a *probability density function* (PDF). The sum of a probability density function over all values is always 1. The integral of a probability density function over all values is likewise 1.

Definition 13.3 (Histogram) *A histogram $h(x)$ is a table or chart of the number of times each of a set of particular, exclusive outcomes occurs out of a collection of events.*

Definition 13.4 (Probability) *For exclusive, equally likely outcomes, the probability of a particular outcome is the number of times they are positive taken out of the total number of outcomes. See Fig. 13.2 for a graphical representation. (Positive is taken here to indicate a favorable result, i.e., do they count toward a particular outcome? For example, are they true or within the desired*

range of values?)

$$p = \frac{Number\ of\ positive\ outcomes}{All\ possible\ outcomes} \tag{13.1}$$

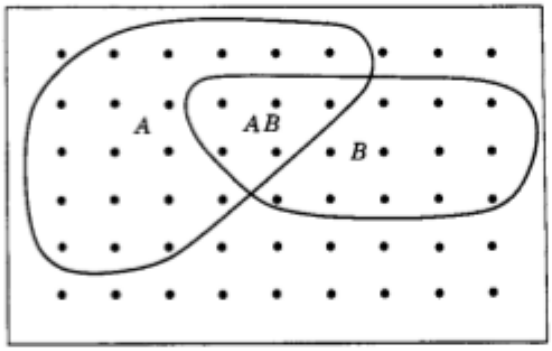


Figure 3.1

Figure 13.2: A figure from Boas (2006) graphically illustrating probability as a Venn diagram. Each point on the graph represents an equally likely, mutually exclusive outcome of an experiment.

Fig. 13.2 is a graphical representation of probability. Each point represents an equally likely, mutually exclusive result of an experiment. Some of these results contribute to outcome A indicated by the contour labeled A, some contribute to B, some contribute to both A and B which are labelled AB. The probability $p(A)$ is the ratio of the number of points inside the A contour ($N(A)$) to the total number of points N, $p(A) = N(A)/N$. Similarly, the probability $p(B)$ is the ratio of the points in B to the total points, $p(B) = N(B)/N$. The probability of both A and B, $p(AB)$ is the ratio of the number of points in the intersection to the total number of points, $p(AB) = N(AB)/N$. The *conditional probability* of B given that A is true, $p_A(B)$ (sometimes written as $p(A|B)$) is the ratio of the number of points within the AB intersection to the number of points within A–*which is considerably different from the overall probability of* $p(AB)$ *which is the ratio of points in* AB *to all points in the domain*, $p_A(B) = N(AB)/N(A) \neq N(AB)/N$. Similarly, $p_B(A)$ is the ratio of the number of points within the AB intersection to the number of points in B, $p_B(A) = N(AB)/N(B) \neq N(AB)/N$.

Definition 13.5 (Probability Distribution) *For exclusive, equally likely outcomes, the probability distribution is the function expressing the probability of a each particular outcome versus an index covering each outcome in turn. It is the limit of the normalized histogram over infinitely many repeated experiments. The sum of the probability density over all outcomes equals one.*

$$p(x_i) = \frac{Number\ of\ outcomes\ positive\ to\ value\ x_i}{All\ possible\ outcomes} = \lim_{N\to\infty} \frac{h(x_i)}{N}. \tag{13.2}$$

Definition 13.6 (Probability Density) *If outcomes are indexed by a continuous variable, the probability density function, $\rho(x)$, is the probability of the outcome lying within an infinitesimal distance of a particular value x. The definite integral of the probability density function gives the probability of the events lying between the endpoints of the integral. The integral of the probability*

density function over all values is one.

$$\rho(x) = \frac{\textit{Number of outcomes positive to values between } x \textit{ and } x + \mathrm{d}x}{\textit{All possible outcomes}}, \tag{13.3}$$

$$p(a \leq x \leq b) = \int_a^b \rho(x)\,\mathrm{d}x. \tag{13.4}$$

13.2.1 Sample Estimates

Considering the histogram to be an estimate of the probability distribution function for a limited sample of data is our first indication of a much larger idea–that of sampling and representativeness. It is easy to imagine that one sample or another might be an incomplete representation of the PDF, so that histogram would be an odd one. Thus, the histogram is not the PDF, but just an estimate of it based on a finite size sample. We can only consider the PDF as the limit to which a histogram converges with an infinite number of observations.

13.2.2 Do PDFs exist? Frequentists and Bayesians

This concept of defining an object that cannot be observed as the limit of an infinite number of observations irritates many people. Some statisticians prefer the "frequentist" approach, where all of probability and statistics can only be inferred from actual observations rather than through experiments. The other camp, the Bayesians, like to perform thought experiments that suggest PDFs of various processes or experiments, even if these experiments cannot be carried out.

Plato vs. Aristotle

This argument between frequentists and Bayesians is quite old. In some sense, it stems from the same intellectual thread as the Platonists versus the Aristotelians. Plato liked to frame arguments in terms of comparisons to "ideals" that exist outside of reality–indeed, if they did exist they could not be an ideal. Bayesians tend to think of ideal probability distribution functions in a similar vein, as derivable from thought experiment and then sampled data is an imperfect representation. Aristotle on the other hand argued for measuring the essences (the what) and seeking causes (the why), which is the nature of frequentist statistics. Frequentists define and measure statistics (the what), and from these definitions can relate one statistic to another (the why).

13.3 What do you Mean Mean?

Reading: Boas (2006, 15.5, 15.6)

13.3.1 Mean

We all learned how to take a mean or an average in second or third grade. You sum the values of all of the elements of a sample and then divide by the number of elements N.

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i = \frac{x_1 + x_2 + \cdots + x_N}{N}. \tag{13.5}$$

One big difficulty of the mean is that the operation of taking the mean does not commute with nonlinear functions. That is, for a generic function F,

$$F(\langle x \rangle) \neq \langle F(x) \rangle. \tag{13.6}$$

However, if F is a linear function, this is true (Why? Can you prove this?).

Geometric & Harmonic Mean

The fact that the mean only commutes with functions that are linear leads us to sometimes consider other kinds of means. The geometric mean and the harmonic mean are

$$\langle x \rangle_g = \left(\prod_{i=1}^{N} x_i \right)^{1/N} = \sqrt[N]{x_1 x_2 \ldots x_N}, \tag{13.7}$$

$$\langle x \rangle_h = \frac{N}{\sum_{i=1}^{N} \frac{1}{x_i}} = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_N}} \tag{13.8}$$

These means are less common than the arithmetic mean, but they are useful sometimes. I include them here primarily to accentuate that the mean is not the only way to consider a distribution of values.

13.3.2 Median

The median is the n^{th} 50 percentile value. That is, for a given set of measurements, it is the value that sits in the middle, with half of the set being smaller than it and half being larger (if there are an even number of measurements, some people prefer to take the mean of the middle two as the median, but this is a detail). One can also consider other percentiles or quartiles or deciles, etc., where the idea is that a given percentage of the measurements are below and the remainder are above.

13.3.3 Mode, Maximum Likelihood, Range, Etc.

The mode, or statistical mode, or maximum likelihood value is the value where the probability distribution function or the probability density function have their maximum value. It is therefore the most likely value. The maximum value and the minimum value are likewise the ends of the nonzero values of the PDF (for bounded pdfs). The difference between the maximum and the minimum values is the range.

These statistics are particularly prone to error, since a single spurious value can skew the result significantly. Thus, percentiles are usually preferred when the dataset is reasonably large.

13.3.4 The Cumulative Distribution and Density Functions

A related idea is the cumulative distribution function or the cumulative density function (CDF), which are the sum and integral of the probability distribution function and the probability density function, respectively. Thus, the CDF has a value of 0 at the minimum of the range and 1 at the maximum of the range. Any percentile or the median is easily read from the CDF, as the inverse of the CDF is the value at which that that fraction of the number of events is smaller than the value. The PDF can be found from the CDF by differentiation or differencing. Constructing a sample estimate of the CDF can be more robust than a sample histogram, for the same reason that integrals are less noisy than derivatives.

13.3.5 The Mean as a Moment

There is another way to calculate the mean of a random variable which is closely related to the mean over a volume or area from chapter 6. Suppose we have a (not normalized) histogram $h(x)$ over a set of n different values of x over N different experimental results. We could sum up all of the columns of the histogram like this

$$N = \sum_{j=1}^{n} h(x_j). \tag{13.9}$$

to arrive at the total number of experiments. Or, we could form a normalized histogram like this

$$1 = \sum_{j=1}^{n} \frac{h(x_j)}{N}. \tag{13.10}$$

Note that here j runs over all of the possible outcomes of x, hitting each only once. Since each value of the histogram is associated with one particular value of the variable x_i, we can compare this to the ordinary average,

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i. \tag{13.11}$$

Suppose we rearranged the order of the sum in the average, so that we first summed all of the occurrences that equal the first possible value x_1, then the occurrences of the second value x_2, etc. The number of terms of each type in the sum would just be the *histogram*! Thus,

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i = \frac{1}{N} \sum_{j=1}^{n} x_j h(x_j). \tag{13.12}$$

We can think of a few ways of expressing the same idea,

$$\langle x \rangle = \frac{1}{N} \sum_{j=1}^{n} x_j h(x_j) = \sum_{j=1}^{n} \frac{x_j h(x_j)}{N} = \frac{\sum_{j=1}^{n} x_j h(x_j)}{\sum_{j=1}^{n} h(x_j)}. \tag{13.13}$$

Each of these is exactly equal to the sample average over the experiments included in the histogram.

Taking this idea to the limit of infinite experiments, we see that

$$\langle x \rangle = \sum_{j=1}^{n} x_j p(x_j). \tag{13.14}$$

For a probability density function of a continuous variable x, the equivalent form is

$$\langle x \rangle = \int_{-\infty}^{\infty} x\rho(x)\,\mathrm{d}x. \tag{13.15}$$

Notice that the normalization of the probabilities means that we do not need the denominator if we sum or integrate over all possible values.

Sometimes, the term *moment* is used to describe an operation like the average above over a distribution or density function. Thus, the first moment of the probability density function is the average:

$$\mu = \langle x \rangle = \int_{-\infty}^{\infty} x\rho(x)\,\mathrm{d}x. \tag{13.16}$$

The second moment is

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} x^2\rho(x)\,\mathrm{d}x. \tag{13.17}$$

The m^{th} moment is

$$\langle x^m \rangle = \int_{-\infty}^{\infty} x^m\rho(x)\,\mathrm{d}x. \tag{13.18}$$

This name comes from the close relationship between this form and moments in physics, such as the moment of inertia.

The moments are sometimes centralized and normalized. The central moments are the moments about the mean. The m^{th} centralized moment is

$$\mu_m = \langle \big(x - \langle x \rangle\big)^m \rangle = \int_{-\infty}^{\infty} \big(x - \langle x \rangle\big)^m \rho(x)\,\mathrm{d}x. \tag{13.19}$$

The moments can be normalized using the standard deviation $\sigma_x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$ (see below). The m^{th} normalized moment is

$$\frac{\langle x^m \rangle}{\sigma_x^m} = \frac{\langle x^m \rangle}{\left(\sqrt{\langle x^2 \rangle - \langle x \rangle^2}\right)^m}. \tag{13.20}$$

The standardized moment, which is just the normalized, centralized moment is

$$\frac{\mu_m}{\sigma_x^m} = \frac{\langle \big(x - \langle x \rangle\big)^m \rangle}{\left(\sqrt{\langle x^2 \rangle - \langle x \rangle^2}\right)^m}. \tag{13.21}$$

The *expectation value* of any function $f(x)$ is just the evaluation of that function weighted by the probability of x or probability density of x if x is continuous.

$$\langle f(x) \rangle = \sum_{j=1}^{n} f(x_j) p(x_j), \tag{13.22}$$

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x)\rho(x)\,\mathrm{d}x. \tag{13.23}$$

Thus, the moments are the expectation values of the powers of x.

13.4 Error and Uncertainty

The average, mean, median, and percentile values of a distribution are all useful concepts, but we will need a few more to be able to compare measurements. In particular, we might like to know if a measured quantity A from a set of repeated experiments has a different value than a measured quantity B. If both A and B are uncertain, or involve random noise in their measurement, then what we need to do is compare the probability distribution of A and that of B. It is not just the mean of these distributions we compare, since the distributions might be so wide that they are indistinguishable. That is, even if their mean values differ, they must differ by an amount that exceeds a quantification of the noise or spread of the distribution.

Thus, we associate the ideas of error, uncertainty, and precision with the width of the distribution.

13.4.1 Variance and Standard Deviation

The variance is a statistic to quantify the width of the distribution. It can be defined as (using the standard notation) as

$$\sigma_x^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \langle x\rangle\right)^2. \tag{13.24}$$

The standard deviation is the square root, or just σ_x. Thus, the variance is the second centralized moment.

We can also associate the standard deviation with the moments of the probability distribution or probability density function.

$$\begin{aligned}
\sigma_x^2 &= \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \langle x\rangle\right)^2 && (13.25)\\
&= \frac{1}{N}\sum_{i=1}^{N}\left(x_i^2 - 2x_i\langle x\rangle + \langle x\rangle^2\right) && (13.26)\\
&= \frac{1}{N}\sum_{i=1}^{N}\left(x_i^2\right) - \frac{2\langle x\rangle}{N}\sum_{i=1}^{N}(x_i) + \frac{\langle x\rangle^2}{N}\sum_{i=1}^{N}(1) && (13.27)\\
&= \frac{1}{N}\sum_{i=1}^{N}\left(x_i^2\right) - 2\langle x\rangle^2 + \langle x\rangle^2 && (13.28)\\
&= \frac{1}{N}\sum_{i=1}^{N}\left(x_i^2\right) - \langle x\rangle^2 && (13.29)\\
&= \langle x^2\rangle - \langle x\rangle^2. && (13.30)
\end{aligned}$$

This is a nice reminder of the fact that the average of a function is not the function of the average, otherwise the variance would always be zero.

13.4.2 Skewness and Kurtosis

Sometimes it is useful to consider quantities similar to variance, but based on higher moments that the first and second. The primary ones are the skewness (the third centralized moment) and the kurtosis. The kurtosis is the fourth standardized moment minus 3. The reason behind the minus three is that the fourth centralized moment of the normal distribution is $3\sigma^4$, so the fourth standardized moment is 3. Thus, the kurtosis is zero for a normal distribution. The skewness indicates whether the largest deviations from the mean tend to occur above (positive skewness) or below (negative skewness) the mean value. The kurtosis indicates how likely extreme values (i.e., those far away from the mean) are. A positive kurtosis indicates more extreme values than is typical for a normal distribution while a negative value indicates fewer extreme values.

13.4.3 Accuracy and Precision

Two concepts (related to variance, skewness, and kurtosis) are importantly differentiated in statistics: accuracy and precision. The *accuracy* of a measurement provides information about the reliability of the mean value. That is, a measurement is accurate (or sometimes called unbiased) if the average over many repeated experiments converges to the true value. Thus, accuracy is about $\langle x \rangle$. A measurement is *precise* if the distribution of experimental results is narrow (small variance). A useful definition of precision might be $\sigma_x \ll \langle x \rangle$, for example.

However, the fact that a measurement is precise does not mean that it is accurate. Suppose for example that I throw three darts at a dartboard and uniformly surround the bullseye, but do not score any darts in the bullseye. That is accurate, but imprecise, throwing. If instead I miss the dartboard all three times but all three of darts are clustered together in a bunch with the darts touching each other, then that is precise, but inaccurate, dart throwing. If I throw all three darts into the bullseye, that is both precise and accurate throwing.

13.5 Probability Theorems

There are a number of probability theorems that allow one to compare probabilities, add probabilities (what is the probability of A and B occurring together?), determine whether occurrences are independent or not (does A depend on B?), determine if they are mutually exclusive (does A preclude the occurrence of B?), and related ideas.

13.5.1 Venn Diagrams

You probably all learned how to make Venn diagrams in school. They are a map of probability distributions! Here are some of my favorite examples, which help illustrate the idea that some of the outcomes that occur overlap or do not overlap with other outcomes. Probability is defined as the fraction of positive outcomes You get the idea.

In the discussion of Fig. 13.2, we considered the number of experimental results (points) that fell into a particular category or outcome (contour). We can readily interpret Venn diagrams in such a manner. These interpretations lead us to quickly establish the fundamental theorems of combining

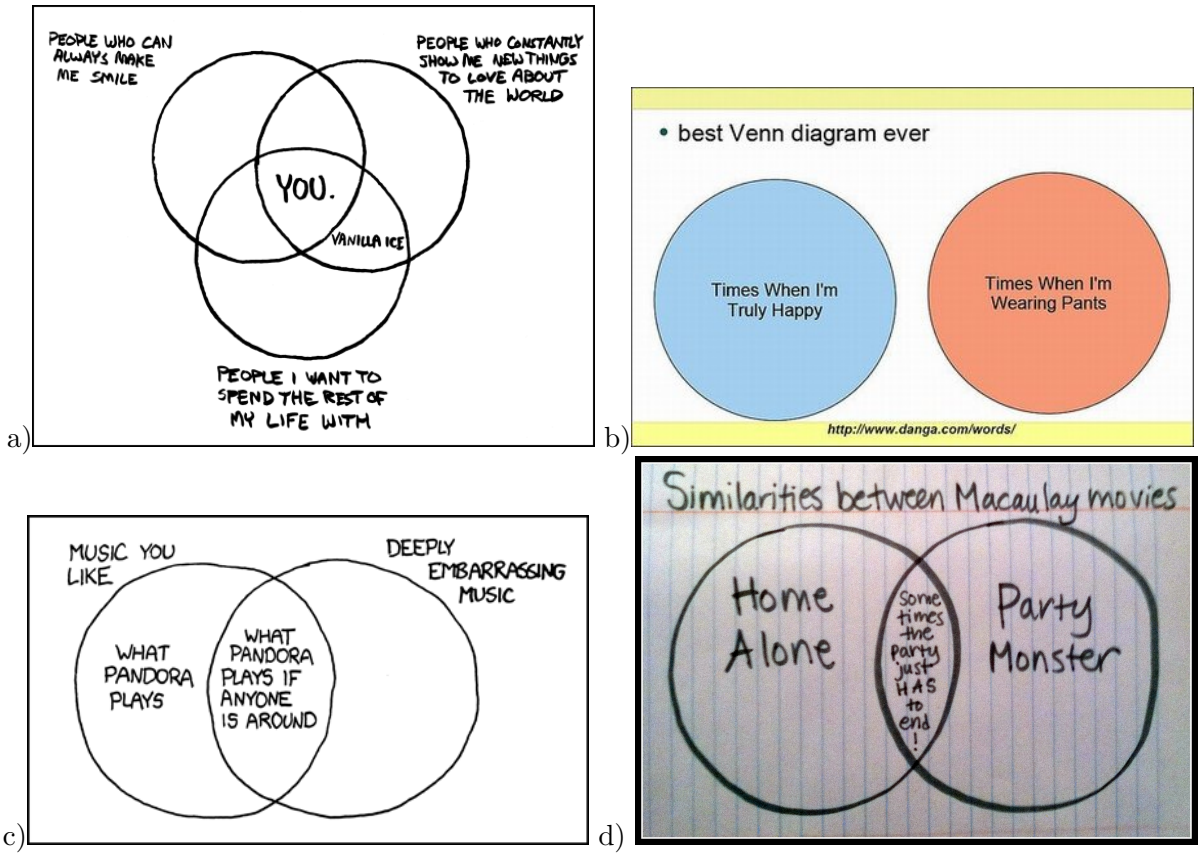
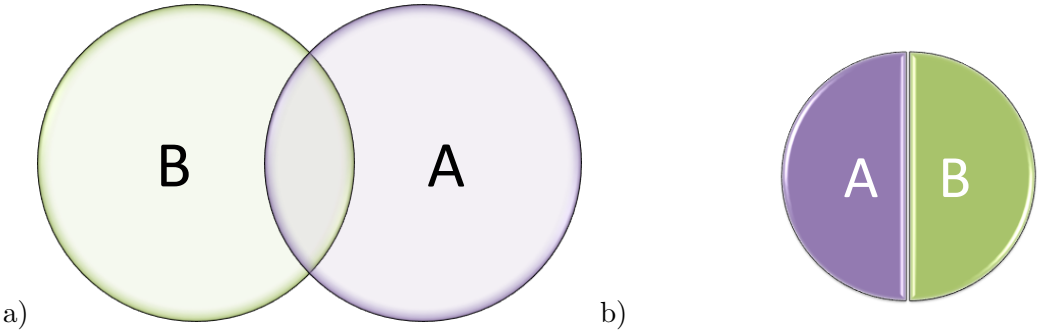


Figure 13.3: Examples of funny Venn diagrams, a) and c) from Randall Munroe (http://xkcd.com), b) from http://www.danga.com/words, and d) one by Adriana DiGennaro from http://www.funnyordie.com.

and comparing probabilities. Some less exciting Venn diagrams can help us to understand combining and joining probabilities (Fig. 13.4).



a) b)

Figure 13.4: Examples of less funny Venn diagrams, a) indicates probability of $A, p(A)$, probability of $B, p(B)$, and probability of A and B, $p(AB)$ where the two intersect (compare to Fig. 13.3c-d). b) Indicates the a case where A and B are mutually exclusive, in this case $p(A+B) = p(A)+p(B)$ (compare to Fig. 13.3b).

We have already defined the probability of outcomes A, B, and both A and B or AB, and the conditional probabilities $p_A(B)$ and $p_B(A)$. We can add to this list the probability of *either* A or B or both, $p(A+B)$. Fig. 13.4a shows all of these examples.

$$p(A) = \frac{N(A)}{N} \tag{13.31}$$

$$p(B) = \frac{N(B)}{N} \tag{13.32}$$

$$p(AB) = \frac{N(AB)}{N} \tag{13.33}$$

$$p(A+B) = \frac{N(A)+N(B)-N(AB)}{N} \tag{13.34}$$

$$p_B(A) = \frac{N(AB)}{N(B)} \tag{13.35}$$

$$p_A(B) = \frac{N(AB)}{N(A)}. \tag{13.36}$$

We can also make relationships between the probabilities by combining the counting formulas in (13.31-13.36), such as

$$p(A+B) = p(A)+p(B)-p(AB). \tag{13.37}$$

$$p(AB) = p_B(A)\cdot p(B) = p(A)\cdot p_A(B). \tag{13.38}$$

$$\text{Bayes' Theorem: } p_A(B) = \frac{p(AB)}{p(A)}. \tag{13.39}$$

Bayes' Theorem is the most famous of these relationships, for reasons we shall see later.

We can examine the probability of relationships between mutually exclusive outcomes and independent outcomes,

$$\text{if mutually exclusive: } p(A+B) = p(A)+p(B), \therefore p(AB) = 0, \tag{13.40}$$

$$\text{if independent: } p(AB) = p(A)\cdot p(B), \therefore p_A(B) = p(B), p_B(A) = p(A). \tag{13.41}$$

Note that two mutually exclusive outcomes of nonzero probability cannot be independent, because Bayes' Theorem says that $p_A(B) \neq 0$ only if $p(AB) \neq 0$. To put it in words, if two events are mutually exclusive, then knowing that one did occur means knowing that the other did not occur, so they are not independent. Likewise, two independent outcomes of nonzero probability cannot be mutually exclusive!

As examples of how logic and probability combine, we can examine the probability of combining outcomes using the Boolean logical operators.

$$\begin{aligned}
\text{And (both): } p(AB) &= p_B(A) \cdot p(B) = p(A) \cdot p_A(B). \\
\text{Not: } \sim p(A) &= 1 - p(A). \\
\text{Inclusive Or (either or both), OR: } p(A+B) &= p(A) + p(B) - p(AB). \\
\text{Not Or (none), NOR: } \sim p(A+B) &= 1 - p(A) - p(B) + p(AB). \\
\text{Exclusive Or (either but not both), XOR: } p(A+B) - p(AB) &= p(A) + p(B) - 2p(AB). \\
\text{Not And, NAND (either or none): } 1 - p(AB) &= 1 - p_B(A) \cdot p(B) = 1 - p(A) \cdot p_A(B). \\
\text{Not XOR, XNOR (both or none): } \sim (p(A+B) - p(AB)) &= 1 - p(A) - p(B) + 2p(AB).
\end{aligned}$$

13.6 Parametric Distributions

One way to estimate a distribution is by building histograms from large numbers of experiments. However, if we take a more Bayesian attitude, there are many distributions that can be generated from thought experiments. It is useful to be familiar with some of these standard distributions, as they serve a good approximations to many types of experiments.

13.6.1 Uniform

The uniform distribution is the simplest distribution. It has either a discrete set of possible outcomes, each of which is equally likely, or a continuous range out outcomes for which the probability density is equal.

The uniform probability distribution for n values is

$$p(x; n) = \frac{1}{n}. \tag{13.42}$$

The probability density function for the range of values between a and b, where $b > a$ is

$$p(x; a, b) = \frac{1}{b-a}. \tag{13.43}$$

The meaning of expectation value for this distribution is a bit odd, as all values are equally likely. So, the expected value is just the average of the possible values. Tables 13.1 and 13.2 table summarizes important statistics about the binomial distribution.

name	statistic	result
pdf	$p(x; n, \wp)$	$\frac{1}{n}$
mean	$\langle x \rangle$	$\frac{\sum_{i=1}^{n} x_i}{n}$
variance	$\langle x^2 \rangle - \langle x \rangle^2$	$\frac{\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n}$

Table 13.1: Statistics of the discrete uniform distribution.

name	statistic	result
pdf	$p(x; b, a)$	$\frac{1}{b-a}$
mean	$\langle x \rangle$	$\frac{a+b}{2}$
variance	$\langle x^2 \rangle - \langle x \rangle^2$	$\frac{(a-b)^2}{12}$

Table 13.2: Statistics of the continuous uniform distribution.

13.6.2 Binomial & Bernoulli

Reading: Boas (2006, 15.7)

The binomial distribution is the probability of finding x successes in n independent experiments with a probability of success of $\wp$. The special case of $n = 1$ is called the Bernoulli distribution, and a single experiment with a yes or no outcome is called a Bernoulli trial. A Bernoulli distribution with equal probability is one example of a uniform distribution. The distribution is given by

$$p(x; n, \wp) = \binom{n}{x} \wp^x (1 - \wp)^{n-x} = \frac{n!}{x!(n-x)!} \wp^x (1 - \wp)^{n-x} . \tag{13.44}$$

An example is the odds of getting exactly x heads out of 8 coin tosses. In this case $\wp = \frac{1}{2}$, $n = 8$. So,

$$p(x; 8, \frac{1}{2}) = \binom{8}{x} \frac{1}{2^x} \frac{1}{2^{8-x}} = \frac{8!}{x!(8-x)!} \frac{1}{2^8} . \tag{13.45}$$

Another example is drawing a card from a well-shuffled deck, replacing the card randomly back into the deck and reshuffling between draws. Fig. 13.5 from (Boas, 2006) gives examples of this distribution for varying $n, \wp$. Table 13.3 summarizes important statistics about the binomial distribution.

13.6.3 Counting: Combinations and Permutations

Bernoulli trials and related experiments where outcomes are repeatedly drawn from a uniform distribution lead one to consider forming probabilities by counting the number of possibilities. In such problems, the permutation and the combination formulas are useful.

name	statistic	result
pdf	$p(x; n, \wp)$	$\frac{n!}{x!(n-x)!} \wp^x (1 - \wp)^{n-x}$
mean	$\langle x \rangle$	$n\wp$
variance	$\langle x^2 \rangle - \langle x \rangle^2$	$n\wp(1 - \wp)$

Table 13.3: Statistics of the binomial distribution.

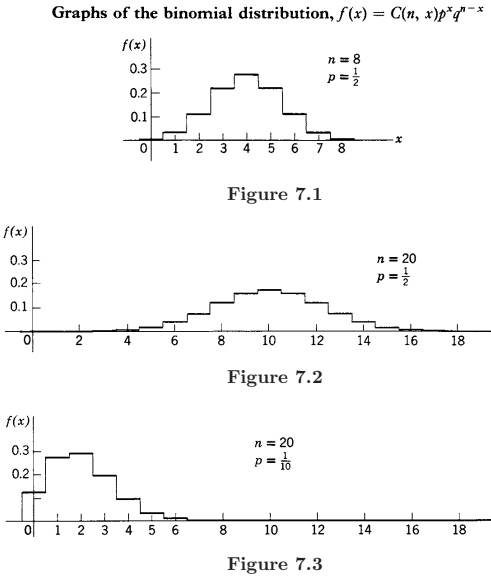


Figure 13.5: Examples of binomial distributions for varying $n, \wp$.

If we repeatedly draw from a uniform distribution, and the order in which the results are received is important, and particular outcomes can be repeated, then the odds of a particular outcome is

$$p = \left(\frac{1}{N}\right)^n \tag{13.46}$$

where N is the number of total possibilities, and n is number of samples chosen. Imagine drawing a card from a shuffled deck, writing down the answer, and then replacing the card and reshuffling before drawing again. If N is the number of cards in the deck and n is the number of times this process is repeated, then an ordered list made in this way will occur with the probability p. This count is sometimes called the permutation with replacement.

A reordering is an accounting of how many ways one can resort an ordered list of things. Consider an ordered deck of n cards. Shuffle it. How many possible shuffles are there? There are n choices for the first card in the deck, $n-1$ for the second, $n-2$ for the third, etc. Thus, the number of reorderings is $n!$ and the probability of a given reordering is

$$p = \frac{1}{n!} \tag{13.47}$$

There are surprisingly many ways to reshuffle or permute a list!

A permutation is the similar to the case above, except instead of considering the order of the whole reshuffled deck, suppose we are interested in only the first r cards of the shuffled deck, but we still care about the order in which they occur. Clearly, $r \leq n$, and when $r = n$ we recover the whole deck shuffle. The number of ways of reordering n cards choosing r at a time is $P(n, r)$, can is called the number of permutations of n items choosing r at a time. The number of such permutations and the probability of any one of them is

$$P(n, r) = \frac{n!}{(n-r)!}, \tag{13.48}$$

$$p(n, r) = \frac{(n-r)!}{n!}. \tag{13.49}$$

Finally, suppose we don't worry about the ordering of the choice we find. Then, reshuffling is irrelevant unless we choose a number of cards less than n. This kind of counting is called the number of combinations of n things r at a time, and the count and probability of one are

$$C(n,r) \equiv \binom{n}{r} = \frac{n!}{(n-r)!r!}, \tag{13.50}$$

$$p(n,r) = \frac{(n-r)!r!}{n!} \tag{13.51}$$

Note that this count is precisely the number of permutations divided by the number of ways r things can be reordered! Thus, if order doesn't matter, then these are all equivalent and we divide by the number of these to arrive at the count. The n choose r notation is used here, which we have seen before but in a different context (Section 5.2.1 on multivariate Taylor series)! The result is the same as the formula for the coefficients of a polynomial expansion!

There is one more count that is useful to consider, which is very important for the bootstrapping technique for estimating uncertainty we will learn later. This count is the number of combinations *with replacement*. This is the number of ways that one can choose from n things r at a time, where order does not matter, but *duplicates are allowed*. Think of drawing from a deck of cards, writing down the card, replacing and shuffling, drawing the next card, etc., r times in a deck of n cards. However, unlike the permutation with replacement, we don't consider the order in which we draw each card, just the fact that it was drawn. In the bootstrapping case, for example, we will be interested only in an unordered statistic (e.g., the average value) of the cards drawn in this manner. This count of combination with replacement results in

$$\mathcal{C}(n,r) = C(n+r-1,r) \equiv \left(\!\binom{n}{r}\!\right) = \binom{n+r-1}{r} = \frac{(n+r-1)!}{(n-1)!r!}. \tag{13.52}$$

The number of things to choose from is greater, because of the increased choice (freedom to choose a duplicate). This makes the factorial in the numerator a bit bigger, but also the factorial in the denominator a bit bigger. Overall, the larger numerator wins out.

13.6.4 Normal

The normal distribution (a.k.a. the Gaussian distribution, a.k.a. the bell curve) is one of the most important ideas in statistics. The probability density function is

$$\rho(x;\sigma,\mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \tag{13.53}$$

This bell-shaped curve is extremely important, as it is the limiting distribution for the average of any uniform, independent data drawn from the same distribution. Since we are often in the business of averaging data, this distribution describes the randomness in our averages.

There are two adjustable parameters, the mean (μ) and the variance (σ^2) of the distribution. If one samples from this probability distribution often enough, the sample mean and the sample variance will converge to these values. That is, $\lim_{n\to\infty}\langle x\rangle = \mu$ and $\lim_{n\to\infty}\langle (x-\langle x\rangle)^2\rangle = \sigma^2$ where n is the number of samples taken from the distribution.

For finite n, the sampled mean and sampled variance will not be equal to these limits. Tests, such as the *t-test*, allow one to estimate the uncertainty in these sample estimates versus the true

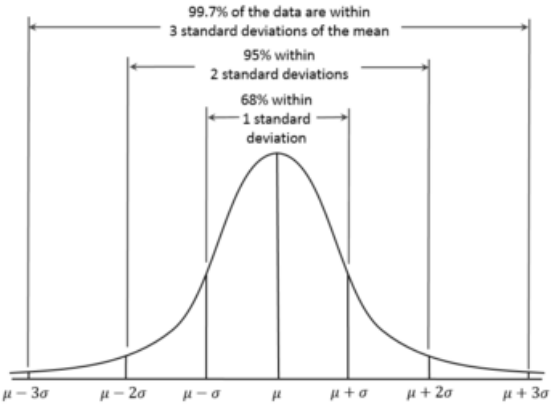


Figure 13.6: The Gaussian, normal, or bell-curve distribution, normalized and scaled by its mean and variance (Wikimedia Commons).

values. The *central limit theorem* says that any average of independent samples drawn repeatedly from the same distribution will eventually converge to being normally distributed. This effect gives the variance of these averages, which can be used to determine a confidence interval on sample estimates of the average.

13.6.5 Student's t distribution

Just as the Gaussian distribution is the limit of samples of the mean, the Student t distribution is the distribution of sample variances that result from sampling from a Gaussian distribution. Thus, if one wants to test whether the variance is different in one set of measurements from another (is one noisier?), the samples can be matched to a Student t distribution. The name t-test means exactly this kind of comparison. The "student" who discovered this distribution and test was working for the Guinness company at the time, and these statistics are useful in testing and guaranteeing the quality of ingredients from limited subsamples. The author used "student" as a pen name on the paper.

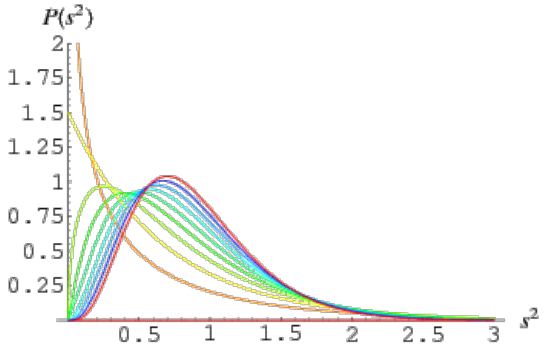


Figure 13.7: The Student's t distribution, for sample sizes of 1 to 10 (Wolfram Alpha).

13.6.6 Dirac Delta

The Dirac delta distribution is sometimes called a function, but it really is not since it is discontinuous and infinite. We touched on its use in solving differential equations, but it is also a useful concept in statistics. It is the distribution that results from the examination of a set of possible results where only one one of them ever occurs.

Discrete

For a discrete set of values, the probability distribution function for the Dirac delta is

$$p(x;x^*) = \begin{cases} 1 & \text{if } x = x^* \\ 0 & \text{if } x \neq x^* \end{cases} \tag{13.54}$$

Continuous

For a continuous set of values, the probability density function for the Dirac delta is

$$\rho(x;x^*) = \delta(x^* - x) \tag{13.55}$$

Which means that the probability (integral of probability density function) over any interval is 1 if the interval includes x^* and 0 if not.

13.6.7 Exponential

Intermittent phenomena, such as velocity jumps in turbulence, may be nearly Gaussian for moderate values but closer to exponential e^{-x} in the tails of the distribution. Radioactive decay also has an exponential probability of occurrence with time, the Boltzmann-Maxwell distribution is exponential in terms of the energy of moving parcels, and the Beer-Lambert (Beer's) law for the attenuation of light traveling through a medium of molecules also predicts an exponential distribution of distances traveled by molecules in a gas. A few of the homework problems address exponential distributions, largely because they are easy to integrate and work with based on our known rules.

13.6.8 Errorbars and Confidence Limits

One important aspect of estimating the probability distribution of a statistic is to provide estimates of that statistic with errorbars or confidence limits. Generally, errorbars are in the units of the measurement (so they can be plotted), whereas confidence intervals are in the units of percentile. There is no consistent definition for how wide an errorbar should be, or what level of confidence should be plotted. Here are some common choices,

1. To mirror the standard notation of the mean of a statistical distribution plus or minus 1 standard deviation ($\mu \pm \sigma$), errorbars may be chosen at $\pm 1\sigma$. Note that only 68% of the data is expected to occur within these errorbars.

2. Sometimes the mean plus or minus 2 standard deviations ($\mu \pm 2\sigma$) is chosen for errorbars at $\pm 2\sigma$. Note that 95.4% of the data is expected to occur within these errorbars (for normally distributed data).
3. In other cases, errorbars may indicate the unit distance equivalent confidence intervals. Commonly, the 90% confidence interval spans from the 5th percentile to the 95th. The interquartile range of 25th to 75th percentile is also common.
4. 50% (interquartile), 60% (n^{th} 20 to n^{th} 80 percentiles), 90% (n^{th} 5 to n^{th} 95 percentiles), or 95% ($2.n^{\text{th}}$ 5 to $97.n^{\text{th}}$ 5 percentiles)
5. A box and whisker plot is commonly used to denote median (middle line), interquartile range (box width), and range of the (non-outlier) data (whiskers). It is somewhat subjective as to what is an outlier, so sometimes people use the n^{th} 1 and n^{th} 99 percentile values instead of the range.

It is important to distinguish whether the errorbars denote the uncertainty in the mean, also known as the *standard error*, $\mu \pm \sigma/\sqrt{n}$ or the range expected when samples are drawn from the distribution $\mu \pm \sigma$. This is a context-dependent choice.

13.7 Combining Distributions

Distributions can be combined using the rules for probabilities above, or sometimes using analytic "error propagation" formulas. Quite often in the modern era, however, distributions are combined by using "Monte Carlo" methods. This simple idea is to just get a computer to simulate drawing randomly from all of the distributions needed repeatedly to form the distribution of the final answer. If the final answer is usefully approximated, e.g., by a normal distribution, then the details of this distribution (mean, variance, moments) can be quickly calculated by Monte Carlo methods. Monte Carlo methods are particularly powerful for exploring the addition of possible extra constraints on the final outcome.

13.7.1 Uncertainty of a Mean–Central Limit Theorem

In class, we showed that the distribution of an average is unlike the distribution of the original variable. In fact, there is a theorem that states that the distribution of the average of repeated, multiple, independent selections from a the same distribution tends toward the normal distribution.

Theorem 13.1 (Central Limit Theorem) *Let $x_1, x_2, \ldots$ be independent, identically distributed random variables with mean μ and finite nonzero variance σ^2. Then the probability distribution of the sum Σ_n of n of these variables approaches a Gaussian, or uniform or normal, distribution with average $n\mu$ and variance $n\sigma^2$ in the limit $n \to \infty$. Similarly, the probability distribution of the average Σ_n/n of n of these variables approaches a Gaussian distribution with average μ and variance σ^2/n as $n \to \infty$.*

Definition 13.7 (Standard Error) *For an observation that can be considered to be derived from the average over n independent observations of drawn from the same distribution, the standard error is taken to be $\sigma/\sqrt{n}$, where σ is the standard deviation of each of the independent observations and n is the number of independent observations included in the average.*

Thus, the statistics of an average tend to be less noisy and converge more quickly on the average of the distribution. The primary difficulties lie in determining when the observations are independent.

13.7.2 Error Propagation of Normal Distributions

We'd like to figure out how to evaluate the error of a function combining together normally distributed random variables. To do so, we need to consider how simple arithmetic combinations of variables affect the statistics of the result.

Sum of Two Gaussian Variables

If x and y are two independent random variables that are normally distributed (means: μ_x, μ_y and variances: σ_x^2 and σ_y^2), then the distribution of their sum $z = x + y$ is also a normally distributed random variable with mean $\mu_z = \mu_x + \mu_y$ and variance $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$.

Multiplication of a Gaussian Variable by a Constant

As you might suspect from dimensional analysis, multiplication of a Gaussian or normally-distributed variable does not fundamentally affect the probability distribution of the variable (so long as the conversion factor is not itself a random variable!). If x is a random variable that is distributed with a normal (or other well-behaved) distribution with mean μ and variance σ^2, then the distribution of the variable times a constant ax will be a variable that is distributed with the same distribution, but with mean $a\mu$ and variance $a^2\sigma^2$. Indeed, all of the moments of the variable will be rescaled by a, as can be quickly shown by change of variables on the integrals defining the moments.

Product and Convolution of Two Gaussian Distributions

If x and y are two independent random variables that are normally distributed (means: μ_x, μ_y and variances: σ_x^2 and σ_y^2), then the distribution of their product $z = xy$ is not in general a normally distributed random variable.

Don't Confuse Operations on the Distributions as Operations on the Values of the Variables

Misleadingly, simply multiplying together the functions that generate their pdfs *does* result in a Gaussian, but that is not the same as sampling from their distributions. Instead it is the joint probability of independent processes (see (13.41)) both occurring. It has little or nothing to do with the product of the results. Likewise, the sum of two Gaussian functions is not likely to be a Gaussian function, but the distribution of two normally-distributed variables is drawn from a normal distribution. Again, the sum of the functions themselves means something different (13.40) than the sum of the results of variables drawn from the distribution. Finally, multiplying a probability distribution by a constant doesn't even make sense, since the probability distribution must always remain normalized. However, multiplying the values drawn from a distribution by a constant makes lots of sense, and is the equivalent to changing the units of the variable.

Truncated Taylor Series

We already know how to approximate complicated functions with locally linear functions: Taylor series expansion and truncation! Let's examine how we can use this together with the results above to analyze error propagation in functions of normally-distributed variables. Let μ_f and σ_f^2 denote the mean and variance of a function $f(A, B, C)$ that depends on variables A, B, C. The Taylor series of F, truncated at the linear order, about the mean values of these variables μ_A, μ_B, μ_C is just

$$f(A,B,C) \approx f(\mu_A,\mu_B,\mu_C) + (A-\mu_A)\frac{\partial f(\mu_A,\mu_B,\mu_C)}{\partial A} \tag{13.56}$$

$$+ (B-\mu_B)\frac{\partial f(\mu_A,\mu_B,\mu_C)}{\partial B} + (C-\mu_C)\frac{\partial f(\mu_A,\mu_B,\mu_C)}{\partial C} + \ldots \tag{13.57}$$

If A, B, C are normally distributed, then since the derivatives of the function are all known constants, then the truncated Taylor series is just a constant plus a sum of centralized, normally distributed variables scaled by constants, so

$$\mu_f \approx f(\mu_A,\mu_B,\mu_C), \tag{13.58}$$

$$\sigma_f^2 \approx \left(\sigma_A \frac{\partial f(\mu_A,\mu_B,\mu_C)}{\partial A}\right)^2 + \left(\sigma_B \frac{\partial f(\mu_A,\mu_B,\mu_C)}{\partial B}\right)^2 + \left(\sigma_C \frac{\partial f(\mu_A,\mu_B,\mu_C)}{\partial C}\right)^2. \tag{13.59}$$

Errors in Products Propagated by Logarithms

As we have already seen (Section 3.10 and Section 5.3.2), taking the logarithm of a product converts the product to a sum. We can then apply our error propagation formulas quickly.

$$f(A,B,C) = A^a B^b C^c, \tag{13.60}$$

$$\ln f(A,B,C) = \ln\left(A^a B^b C^c\right) = a\ln(A) + b\ln(B) + c\ln(C), \tag{13.61}$$

$$\frac{\partial f}{f} = a\frac{\partial A}{A} + b\frac{\partial B}{B} + c\frac{\partial C}{C}. \tag{13.62}$$

If the variations of each variable are taken as Gaussian, normal distributions, and the values of the variables are taken at their mean and the coefficients are not considered to be random, then

$$\frac{\sigma_f^2}{\mu_f^2} = a^2\frac{\sigma_A^2}{\mu_A^2} + b^2\frac{\sigma_B^2}{\mu_B^2} + c^2\frac{\sigma_C^2}{\mu_C^2}. \tag{13.63}$$

Note that the results will be identical to those found by Taylor series, only a bit faster to derive using logarithms. To see this, just note that Taylor expanding the product about the means is

$$\begin{aligned} A^a B^b C^c &\approx \mu_A^a\mu_B^b\mu_C^c + (A-\mu_A)a\mu_A^{a-1}\mu_B^b\mu_C^c + (B-\mu_B)b\mu_A^a\mu_B^{b-1}\mu_C^c + (C-\mu_C)c\mu_A^a\mu_B^b\mu_C^{c-1} + \ldots \\ &\approx \mu_A^a\mu_B^b\mu_C^c\left[1 + (A-\mu_A)\frac{a}{\mu_A} + (B-\mu_B)\frac{b}{\mu_B} + (C-\mu_C)\frac{c}{\mu_C} + \ldots\right], \\ &\approx \mu_A^a\mu_B^b\mu_C^c\left[1 + a\frac{(A-\mu_A)}{\mu_A} + b\frac{(B-\mu_B)}{\mu_B} + c\frac{(C-\mu_C)}{\mu_C} + \ldots\right]. \end{aligned}$$

This formula will give the same result as (13.63).

13.7.3 Error Propagation by Monte Carlo

When we do not know the distributions of all of the variables, or when the magnitudes of the errors are so large in comparison to their mean values that the Taylor series is a poor approximation of the function, or when strange effects (such as exponents are taken to be random variables as well) then we can use the Monte Carlo method to generate an approximate distribution of the final function. Parametric distributions can then be fit to the Monte Carlo estimate, or hypothesis testing can occur directly using the estimated distribution.

13.7.4 Bootstrap & Jackknife Methods

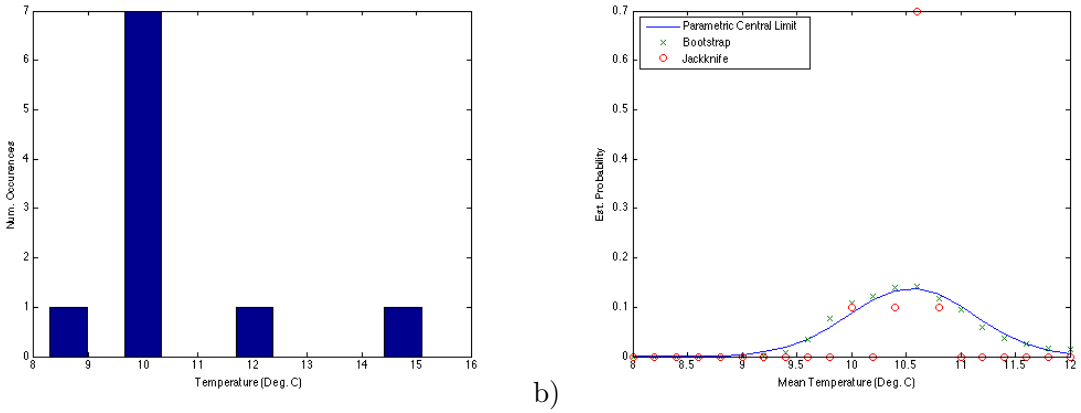
Recent statistical literature has begun to develop methods that are computationally expensive, but statistically simple, to develop nonparametric ways to estimate distributions from a set of measurements. Two popular methods are the bootstrapping method and the jackknife method (Emery and Thomson, 2001). These methods are extremely useful for making estimates of the uncertainty of *virtually any statistic of random variables*, with or without a pre-conceived parametric notion of what the distribution will be. They have some difficulties–moderate computational cost, theoretically shaky in some circumstances, etc.–but they are so easy to calculate and understand that they are becoming very widely used. Historically, the jackknife came first, and then the bootstrap improved upon the idea.

In the jackknife method, one considers what might have occurred if not all of the data had been collected. In particular, suppose one datum less was collected. Assuming there is still a reasonably large number of data points left, we could still calculate all of the same statistics on this slightly smaller set. Now replace the withheld datum, and select another datum to withhold, recalculate all of the statistics, etc. In this way, we create a set of estimates of each of the desired statistics, each of which features a different datum withheld. We can then make histograms of any of these statistics from this set, and proceed to estimate their uncertainty. Unfortunately, there are only as many new estimates as we have data to withhold. Thus, these histograms tend to be spotty, and so in comparison to parametric methods (such as using the standard error for the uncertainty in an average) the jackknife idea usually doesn't perform as well. This is the reason for the name "jackknife," it is an all-purpose tool, but not as good at a specific task (e.g., driving a screw in) as a dedicated tool (e.g., a screwdriver).

The bootstrap method allows one to "pull the data up by its own bootstraps." Just as the name is meant to imply, this seems like a silly idea, but it often works! In the bootstrap method, one assigns uniform likelihood to each of the data collected. Then synthetic datasets (often of the same length as the original, but not necessarily so) are made by resampling from this distribution repeatedly (with replacement after selection, so one datum might get selected more than once within a synthetic dataset). Unlike the jackknife approach, the upper limit for number of possible datasets is huge! For a dataset of length n, there are $(2n-1)!/(n!(n-1)!)$ distinct bootstrap datasets. For example, a dataset of length 10 has 92,378 possibilities. Usually, we don't need this many to estimate a good histogram, so 1000 or 10,000 suffice.

Here is an example of bootstrapping and jackknifing where we also know the parametric result from the central limit theorem. Suppose we make the following 10 measurements of temperature in degrees Celsius: 8.3, 9.8, 9.9, 10.0, 10.0, 10.1, 10.1, 10.2, 12.0, 15.1. The histogram is quite spotty (Fig. 13.8a), but we can still make a sample mean ($\sigma^2 = 10.55$ degrees) and variance (3.3317

degrees^2). As each measurement is independent, the standard error formula predicts that the variance of the distribution of the mean will be $\sigma^2/10$. From this theory, we can draw the blue line in Fig. 13.8b. The jackknife estimate is limited, as there are only 10 different estimates of the mean, but they do cluster roughly around the central limit estimate (red circles). However, the bootstrap estimate, based on 10,000 synthetic datasets of the same length as the original, made by resampling the original with replacement, fits the theoretical optimal set almost perfectly (blue crosses).



a) b)

Figure 13.8: a) The histogram of our dataset, and b) the result of central limit theorem, jackknife, and bootstrap methods for estimating the probability distribution of the mean of these data.

So, while the jackknife method is limited by small datasets, the bootstrap method has enormous potential. Not surprisingly, there are a lot of statisticians studying this method. For our purposes, it is an easy go-to method for many applications. One beautiful advantage of the bootstrap method over other methods is that you can estimate an uncertainty for *any statistic*, even one for which there is no theory, just as easily as for one where there is (such as an average or variance). It is not guaranteed that the bootstrap method is optimal, but seems very unlikely to yield an unreasonable estimate.

13.8 Hypothesis Testing

Hypothesis testing is a method for deciding the statistical likelihood of comparisons using statistical distributions. The statistical distributions can be functions with parameters chosen to fit a dataset (parametric hypothesis testing) or can be constructed from the data itself or histograms of the data (nonparametric hypothesis testing).

The simplest case is to consider a random variable and its likelihood of being above or below a particular value. This idea is implicit in our discussion of percentiles and medians. The percentile is exactly the likelihood of finding that value or higher from the statistical distribution. The cumulative distribution function relates to the percentile as described above, and can be estimated from the data histogram by cumulatively summing the columns. One can also integrate or sum a parametric probability density function or probability density function to arrive at this result. Here are some useful rules of thumb:

1. In a uniform probability distribution the likelihood of finding the value above or below a value

is as simple as adding up the number of possible values above or below.

2. In a not uniform probability distribution the likelihood of finding the value above or below a value is found by adding up the number of possible values above or below times their individual probability.

3. For a normal distribution, 50% of the probability lies within $\pm 0.675\sigma$ of the mean.

4. For any distribution, 50% of the probability lies within the interquartile range (25th percentile to 75th percentile) surrounding the median (50th percentile).

5. For a normal distribution, 90% of the probability lies within $\pm 1.645\sigma$ of the mean.

6. For a normal distribution, 95% of the probability lies within $\pm 1.960\sigma$ of the mean.

7. For a normal distribution, 99% of the probability lies within $\pm 2.576\sigma$ of the mean.

8. For a normal distribution, being within $\pm\sigma$ of the mean has a likelihood of 68%.

9. For a normal distribution, being within $\pm 2\sigma$ of the mean has a likelihood of 95.4%.

10. For a normal distribution, being within $\pm 3\sigma$ of the mean has a likelihood of 99.7%.

11. *Chebyshev's inequality*: The probability p of a random number x with variance σ^2 deviating from its mean value μ by more than any number t obeys $p < \frac{\sigma^2}{t^2}$.

In more elaborate hypothesis testing, one may wonder what the likelihood of two distributions being the same are, or the likelihood of one being greater than the other, etc. Many of these tests are easily boiled down to a function combining the variables, and then using the error propagation formulas above can be reframed into evaluating percentiles of the combined distribution. For more complicated problems, such as the likelihood that the kurtosis of a sample of a normally-distributed variable exceeds a particular value, or whether a set of samples deviates from what one expects when drawing from a given probability distribution, more complicated tests may be required.

An important set of tests, called t-tests, give statistical significances for the sample estimates of a mean of measurements drawn from a normal distribution, but where the true whole population standard deviation is unknown. While the percentile estimates above give likelihoods for known distributions, the t-tests give the equivalents based on samples of finite size.

Of course, the bootstrapping method is ideal for building up guesses of the probability distributions of different datasets. These can be used to estimate how much overlap there is between them, and therefore give quick (except for computation) nonparametric hypothesis testing.

13.8.1 Bayesian Theory

One of the most powerful applications of Bayes' Theorem is in reading the probability of hypothesis given that a particular event occurred or did not occur. So, we can say, "the patient in the experiment recovered after being given the drug," which then lowers the probability of accidental cures. This kind of experimental design is extremely common, and complicated to analyze statistically. The following discussion intends to shed light on why.

Why Most Published Research Findings Are False

There are many problems in the way this approach is commonly used, particularly in medical trials (Ioannidis, 2005). The biggest problems arise when many different drugs or diet plans or lifestyles are tested versus a desired outcome. I think most of us would guess that there would be false positives in this case, e.g., if there is a 95% confidence that a treatment was effective, then there was a 5% chance that this was a fluke. If none of these "cures" was likely to work (i.e., there was no identified mechanism between the action of the "cure" and the desired outcome), the likelihood of false positives becomes overwhelming if many such tests are repeated. If each test is biased (e.g., ineffective drugs not reported, "big results" published over "little ones") then the likelihood of false positives totally overwhelms the likelihood of true findings. In fact, in this limit, published studies are a better measure of *bias* than *true correlation*.

I reproduce the tables from Ioannidis (2005) to illustrate. Table 13.4a gives the standard, unbiased single-study results one expects including false positives (type II error) and false negatives (type I error). Most studies can be designed with these effects in mind. However, the sociology of science comes into play powerfully in skewing the odds. If a researcher is biased in terms of which results are reported, tending to report positive findings over negative ones, then Table 13.4b is more appropriate. For surprisingly modest bias, normal measures of statistical significance are overwhelmed. Similarly, when multiple groups are working on the same problem (e.g., what diets prevent cancer?) then the odds of false positives and false negatives go up, *even if the researchers don't share data until after publication!*

Of course, Ioannidis (2005) was stimulated by medical research, not geophysical research. Maybe our field is less likely to have these problems? Possibly. We are a small field, so n tends to be small and this effect is typically not very significant. However, we are a field that strongly rewards positive results (with PhDs and tenure), while largely ignoring negative results, so reporting bias u is likely to be quite high. Bias is particularly difficult to detect when everyone invents his or her own way to measure an effect. A flexibility of design tends to inflate bias u, and our field is rife with such inventive design. We are often working in noisy fields with small signals, and that works against us. The fewer the number of likely relationships versus the number of tested relationships (R), the worse these effects are. Our fields are particularly prone to this effect. We often regress entire datasets or maps of data versus another data stream (e.g., what climate variables are correlated with El Nino?) This is a recipe for making R very small, because the number of tested but unlikely relationships is very high. We have a way of detecting these screw-ups, which we usually refer to as "correlation is not causation." That is, it is not enough in geophysics to prove correlation between events, one must also provide a mechanism and demonstrate its effectiveness theoretically (usually done in a numerical model). However, in the days of "big data" this failsafe is being neglected more and more often. The hotter a field is (both in terms of number of investigators and tendency to be biased), the less likely the results are to be true. Our field is relatively slow and plodding, but certain corners (e.g., climate change) probably suffer from this effect.

13.9 Data Processing & Modeling

Here I will not describe these methods in detail, but merely describe what they do. You can look them up on Wikipedia or Wolfram if you need this type of functionality.

Table 1. Research Findings and True Relationships

Research Finding	True Relationship		
	Yes	**No**	**Total**
Yes	$c(1-\beta)R/(R+1)$	$c\alpha/(R+1)$	$c(R+\alpha-\beta R)/(R+1)$
No	$c\beta R/(R+1)$	$c(1-\alpha)/(R+1)$	$c(1-\alpha+\beta R)/(R+1)$
Total	$cR/(R+1)$	$c/(R+1)$	c

a) DOI: 10.1371/journal.pmed.0020124.t001

Table 2. Research Findings and True Relationships in the Presence of Bias

Research Finding	True Relationship		
	Yes	**No**	**Total**
Yes	$(c[1-\beta]R+uc\beta R)/(R+1)$	$c\alpha+uc(1-\alpha)/(R+1)$	$c(R+\alpha-\beta R+u-u\alpha+u\beta R)/(R+1)$
No	$(1-u)c\beta R/(R+1)$	$(1-u)c(1-\alpha)/(R+1)$	$c(1-u)(1-\alpha+\beta R)/(R+1)$
Total	$cR/(R+1)$	$c/(R+1)$	c

b) DOI: 10.1371/journal.pmed.0020124.t002

Table 3. Research Findings and True Relationships in the Presence of Multiple Studies

Research Finding	True Relationship		
	Yes	**No**	**Total**
Yes	$cR(1-\beta^n)/(R+1)$	$c(1-[1-\alpha]^n)/(R+1)$	$c(R+1-[1-\alpha]^n-R\beta^n)/(R+1)$
No	$cR\beta^n/(R+1)$	$c(1-\alpha)^n/(R+1)$	$c([1-\alpha]^n+R\beta^n)/(R+1)$
Total	$cR/(R+1)$	$c/(R+1)$	c

c) DOI: 10.1371/journal.pmed.0020124.t003

Table 13.4: Table from Ioannidis (2005) showing the likelihood of research findings versus true findings of correlations between events. R is the ratio of "true relationships" to "unrelated potential relationships," that is, plausible but not true relationships, in the field. The type I error rate, or the rate of obtaining a "false positive" of a research result of true when the actual relationship is false is α. The type II error rate, or rate of failing to reject a false negative (a research result that reads untrue when the relationship is actually true), is β. The number of different tests being considered by all investigators is c. a) shows the traditional rates of detection, including these effects. b) shows the effects when investigation bias u toward a positive result is included. This u is the proportion of over-reporting of positive results and under-reporting of negative ones. c) Shows the results when n independent studies of equal power are working on the same problem (neglecting bias).

13.9.1 Regression and Fitting

Regression methods generally refer to methods where a line (or other function of a few parameters) is fit to a set of data.

13.9.2 Least Squares

Least squares is one form of regression, and it is called this because the method fits a line that is "best" in the sense that it minimizes the squares of the data scatter about the predictions of the line. It is closely related to a method in linear algebra called the Moore-Penrose pseudo-inverse or the Singular Value Decomposition.

13.9.3 Power Laws, Products, and Log-Normal Errors

Oftentimes, rather than a linear function, one anticipates a product of factors (e.g., from the Buckinham Pi Theorem) or a power law dependence. A crucial tool in these types of problems is good old log paper (or the loglog plotting function in MATLAB, etc.) As already shown, the logarithm of a product of powers becomes a linear function. Thus, by first taking the logarithm, and then doing a linear fit, and then returning to the original space, a power law or product of power laws can be fit easily. The only problem with this approach is that minimizing the squares in logarithmic space may not be the best way to select an optimal fit. Nonetheless, this pragmatic approach to nonlinear function fitting was heavily emphasized, especially in engineering, in the 20th century and is still a go-to for most scientists, but more complicated computational approaches are beginning to be more common now.

13.9.4 SVD and EOFs

We have already described the basics of the Empirical Orthogonal Function method. It is a linear algebra method to compress data into a few distinct modes that vary over time. It is closely related to least squares, singular value decomposition, and eigenvalues and eigenvectors!

13.9.5 Spectral Data Processing

The fast Fourier transform, or generally the discrete Fourier transform, is a method to estimate a Fourier series of a finite length dataset with finite sampling frequency. Thus, instead of a time series, one gets back

13.9.6 AR and ARMA Models

Autoregression modeling and autoregression-moving-average modeling is a method for fitting a stochastic model, rather than a deterministic one, to a dataset. Depending on the method, it may be generated to reproduce different statistics of the original data–variances, covariances, lag covariances–all other aspects of the model prediction are chosen at random. These methods are

commonly used to provide a control model against which a new theory (with more stringent predictive power) can be tested. Thus, if you think your model works, then prove it is better than a random (but educated) guess!

13.9.7 Estimates of Nonlinear Functions–A Challenge

As we saw in the chapter on chaos, even deterministic systems can be essentially unpredictable if they fall into a chaotic regime and measurements of initial conditions are slightly uncertain. Such systems greatly tax our ability to model them simply, and the great breakthroughs from chaos science to build predictive capability in real-world chaotic systems are often over-hyped. However, it is still early days, and the combination of cleverness and computation has only just begun.

13.10 Example Problems

13.10.1 Jargon to Argot

Example 13.1 *Define or look up definitions for: statistic, metric, unit, dimension, and measurement. Be sure to make the definitions sufficiently contrasting the differences between these concepts.*

I provide only the relevant noun definitions, among the others.

Definition 13.8 (Statistic) *1. A numerical datum. 2. A numerical value, such as standard deviation or mean, that characterizes the sample or population from which it was derived. (Source: American Heritage Dictionary by way of the Free Dictionary)*

Definition 13.9 (Metric) *1. A standard of measurement. 2. Mathematics A geometric function that describes the distances between pairs of points in a space. (Source: American Heritage Dictionary by way of the Free Dictionary)*

Definition 13.10 (Unit) *4. A precisely specified quantity in terms of which the magnitudes of other quantities of the same kind can be stated. (Source: American Heritage Dictionary by way of the Free Dictionary)*

Definition 13.11 (Dimension) *1. A measure of spatial extent, especially width, height, or length. 4. Mathematics: a. The least number of independent coordinates required to specify uniquely the points in a space. b. The range of such a coordinate. 5. Physics: A physical property, such as mass, length, time, or a combination thereof, regarded as a fundamental measure or as one of a set of fundamental measures of a physical quantity: Velocity has the dimensions of length divided by time. (Source: American Heritage Dictionary by way of the Free Dictionary)*

Definition 13.12 (Measurement) *1. the act or process of measuring 2. an amount, extent, or size determined by measuring 3. a system of measures based on a particular standard (Source: Collins English Dictionary by way of the Free Dictionary).*

Example 13.2 *Problem 15.5.8 Boas (2006). Would you pay $10 per throw of two dice if you were to receive a number of dollars equal to the product of the numbers on the dice? Hint: What is your expectation? If it is more than $10, then the game would be favorable for you.*

Here is the table of equally likely values (the multiplication table).

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

where first die roll value is columns and second die roll is rows. We can form the mean (expected value) by averaging the entries in the table, which yields ($\langle x\rangle = \sum x/N = 441/36 = 12.25$). Thus, we want to take the bet, as we expect to make \$2.25 profit, and the odds of winning are good 17/36 possible rolls are wins and 2/36 break even.

Example 13.3 *Problem 15.6.2 Boas (2006). It is shown in the kinetic theory of gases that the probability for the distance a molecule travels between collisions to be between x and $x + dx$, is proportional to $e^{-x/\lambda}dx$, where λ is a constant. Show that the average distance between collisions (called the mean free path) is λ. Find the probability of a free path of length $\geq 2\lambda$. Hint: the normalized probability density function for the Boltzman distribution is $\rho(x;\lambda) = \frac{1}{\lambda}e^{-x/\lambda}$, where the coefficient is needed so that the integral over all x from 0 to ∞ is one.*

The probability density function for collisions to be between x and $x + \mathrm{d}x$ is $\rho(x) \propto e^{-x/\lambda}$. We know that if we integrate over all (positive) distances, we will find a collision somewhere, so if we choose a coefficient C, such that $\rho(x) = Ce^{-x/\lambda}$, then

$$\begin{aligned} 1 = \int_0^\infty \rho(x)\,\mathrm{d}x = \int_0^\infty Ce^{-x/\lambda}\,\mathrm{d}x = \int_0^\infty C\lambda e^{-x/\lambda}\frac{\mathrm{d}x}{\lambda} = C\lambda\int_0^\infty e^{-y}\,\mathrm{d}y = C\lambda(1-0), \\ C = \frac{1}{\lambda}, \\ \rho(x) = \frac{e^{-x/\lambda}}{\lambda}. \end{aligned}$$

The expectation value and probability of $x > 2\lambda$ are therefore

$$\begin{aligned} \langle x\rangle = \int_0^\infty x\rho(x)\,\mathrm{d}x = \int_0^\infty \frac{xe^{-x/\lambda}}{\lambda}\,\mathrm{d}x = \lambda, \\ p(x \geq 2\lambda) = \int_{2\lambda}^\infty \rho(x)\,\mathrm{d}x = \int_{2\lambda}^\infty \frac{e^{-x/\lambda}}{\lambda}\,\mathrm{d}x = \int_2^\infty e^{-y}\,\mathrm{d}y = e^{-2} - 0 = e^{-2}. \end{aligned}$$

Example 13.4 *Problem 15.5.10 Boas (2006). Let μ be the average of the random variable x. Then the quantities $(x_i?\mu)$ are the deviations of x from its average. Show that the average of these deviations is zero. Hint: Remember that the sum of all the p_i must equal 1.*

If μ is the average of a given set of N measurements with M possible values, then the average of

the deviation from the average is

$$\mu = \sum_{i=1}^{N} \frac{x_i}{N},$$
$$\langle x - \mu \rangle = \sum_{i=1}^{N} \frac{x_i - \mu}{N} = \left(\sum_{i=1}^{N} \frac{x_i}{N} \right) - \left(\mu \sum_{i=1}^{N} \frac{1}{N} \right) = \mu - \mu \frac{N}{N} = 0,$$

Alternatively, if we form the mean from the first moment of the probability distribution,

$$\mu = \sum_{i=1}^{M} x_i p(x_i),$$
$$\langle x - \mu \rangle = \sum_{i=1}^{M} (x_i - \mu) p(x_i) = \left(\sum_{i=1}^{M} x_i p(x_i) \right) - \left(\mu \sum_{i=1}^{M} p(x_i) \right) = \mu - \mu \cdot 1 = 0.$$

Example 13.5 *Problem 15.8.1 Boas (2006). Verify that for a random variable x with normal density function $f(x)$ as in (8.1), the mean value of x is μ, the standard deviation is σ, and the integral of $f(x)$ from (6.3), and (6.4) from ??? $-\infty$ to ∞ is 1 as it must be for a probability function. Hint: Write and evaluate the integrals $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x, \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x, \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,\mathrm{d}x$, see equations (6.2-6.4).*

The variable has a normal (Gaussian) probability density function, so

$$\rho(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)},$$
$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2}\,\mathrm{d}t,$$
$$\int_{-\infty}^{\infty} \rho(x;\mu,\sigma) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}\,\mathrm{d}x = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-x^2/(2\sigma^2)}\,\mathrm{d}x/(\sqrt{2}\sigma)$$
$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2}\,\mathrm{d}t = \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-t^2}\,\mathrm{d}t + \frac{1}{\sqrt{\pi}} \int_{-\infty}^{0} e^{-t^2}\,\mathrm{d}t = \frac{1}{2}(\mathrm{erf}(\infty) - \mathrm{erf}(-\infty)) = 1.$$
$$\langle x \rangle = \int_{-\infty}^{\infty} x\rho(x;\mu,\sigma)\,\mathrm{d}x = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}\,\mathrm{d}x = \int_{-\infty}^{\infty} (y+\mu) \frac{1}{\sigma\sqrt{2\pi}} e^{-y^2/(2\sigma^2)}\,\mathrm{d}y = 0 + \mu$$

The last result for $\langle x \rangle$ relies on a change of variables to $y = x - \mu$. Then there are two integrals, the one y times $\rho(y)$, which is zero since it's an odd function integrated over even bounds. The other one is just μ times $\rho(x)$, which integrates to $\mu \cdot 1$.

$$\langle (x-\mu)^2 \rangle = \int_{-\infty}^{\infty} (x-\mu)^2 \rho(x;\mu,\sigma)\,\mathrm{d}x = \int_{-\infty}^{\infty} y^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-y^2/(2\sigma^2)}\,\mathrm{d}y = \sigma^2$$

I just looked this one up.

13.11 Homework Problems

13.11.1 Manipulation

Exercise 13.1 *Problem 15.1.10 Boas (2006). A shopping mall has four entrances, one on the North, one on the South, and two on the East. If you enter at random, shop and then exit at random, what is the probability that you enter and exit on the same side of the mall?*

Exercise 13.2 *Problem 15.3.15 Boas (2006). Don't use the reduced sample space, just use Bayes' theorem. Use Bayes formula (3.8) to repeat these simple problems previously done by using a reduced sample space. (a) In a family of two children, what is the probability that both are girls if at least one is a girl? (b) What is the probability of all heads in three tosses of a coin if you know that at least one is a head?*

13.11.2 Application

Exercise 13.3 *Problem 15.6.5 Boas (2006).*

Exercise 13.4 *Make up a small dataset of 10 or so data. a) Calculate the mean, variance, and standard error of the mean. b) Describe what statistics of the data are expected to have a distribution where the variance describes the spread and what statistics are expected to have a distribution where the standard error describes the spread of the statistic. c) Describe how jackknife estimation and bootstrap estimation can be used to produce a histogram categorizing the uncertainty in the mean. d) (optional) You may carry out the bootstrap and jackknife estimates computationally for extra credit.*

Exercise 13.5 *Use Table 13.4 to consider the following scenario. a) Suppose an unbiased single-investigator performs a study where ten independent possible linkages between evolution of angiosperms and plate tectonics are tested. The different possibilities are mutually exclusive and estimated to be equally likely, but only one is true. This investigator works hard on the method, and she estimates only a 5% risk of a false positive and a 5% risk of a false negative, but there is only the right kind of data to test 4 of the possible linkages (each with equal power of detection). i) What are the odds that her experiment will result in a yes relationship with the correct true linkage? ii) What are the odds that her experiment will result in a no relationship with the true correct linkage? iii) What are the odds that her experiment will result in a yes relationship with a false linkage? iv) What are the odds that her experiment will result in a no relationship with a false linkage?*
b) A different investigator works on finding relationships performs a study where ten independent possible linkages between water in the mantle and melting are tested. The different possibilities are mutually exclusive and estimated to be equally likely, but only one is true. This investigator works hard on the method, and he estimates only a 5% risk of a false positive and a 5% risk of a false negative, but there is only the right kind of data to test 6 of the possible linkages (each with equal power of detection). Unlike the investigator in a), this investigator is going up for tenure, and so really wants to publish a significant result. Thus, he does not report all 6 possible linkages tested, instead only reports 4, *and so when writing up the paper is drawn toward reporting primarily the linkages that were detected as "true" in his study. This "file-drawering" of negative results and "over-hyping" of positive results can be modeled with a* $u = 0.5$ *bias, meaning that he is 50% more likely to report a positive result than a negative one. i) What are the odds that his experiment will result in a yes relationship with the correct true linkage? ii) What are the odds that his experiment*

will result in a no relationship with the true correct linkage? iii) What are the odds that his experiment will result in a yes relationship with a false linkage? iv) What are the odds that his experiment will result in a no relationship with a false linkage?
c) Worldwide, the urgency of climate change has driven ten groups to consider possible linkages between temperature and carbon dioxide. Each of the 10 groups is able to reach the same experimental accuracies as the investigator looking into evolution and plate tectonics, and they are all unbiased (because they have built elaborate double-blind studies from fear of climate-gate like investigations by skeptics!) and independent from one another (i.e., they don't share data or methods until after the experiments and analysis are done). i) What are the odds that one group will report a yes relationship when testing the correct true mechanism? ii) What are the odds that one group will report a no relationship when testing the true correct mechanism? iii) What are the odds that one group experiment will report a yes relationship when testing a false mechanism? iv) What are the odds that one group will report in a no relationship when testing a false mechanism?
d) Do your results make you concerned about the state of accuracy in the scientific literature?

13.11.3 Scheming Schematics and Articulate Analysis

Exercise 13.6 *Find the zeroth, first, and second moments (not normalized or centralized) of the continuous uniform distribution. Use these to derive the mean and variance of the continuous uniform distribution given in Table 13.2.*

13.11.4 Evaluate & Create

Exercise 13.7 *Make a Venn diagram that describes an aspect of your life or work. Does it reflect independence or mutual exclusivity?*

Chapter 14

A Culture of Neglect: Asymptotics and Approximations

Chapter 15

Are You Crunchy? Computational Concepts

Bibliography

Aki, K., Richards, P. G., 2002. Quantitative seismology, 2nd Edition. University Science Books, Sausalito, Calif.

Arfken, G. B., Weber, H.-J., Harris, F. E., 2013. Mathematical methods for physicists: a comprehensive guide, 7th Edition. Elsevier, Amsterdam.

Batterson, J., 2008. A mathematical perspective of global warming.

Boas, M. L., 2006. Mathematical methods in the physical sciences, 3rd Edition. Wiley, Hoboken, NJ.

Bourguignon, J.-P., Jeltsch, R., Pinto, A. A., Viana, M., 2015. Mathematics of Energy and Climate Change. Springer.

Campillo, M., Ionescu, I. R., Paumier, J.-C., Renard, Y., 1996. On the dynamic sliding with friction of a rigid block and of an infinite elastic slab. Physics of the Earth and Planetary interiors 96 (1), 15–23.

Cushman-Roisin, B., Beckers, J. M., 2010. Introduction to geophysical fluid dynamics: Physical and Numerical Aspects. Academic Press.

Drazin, P. G., Reid, W. H., 2004. Hydrodynamic stability, 2nd Edition. Cambridge University Press, Cambridge, UK.

Einstein, A., 1931. Maxwell's influence on the development of the conception of physical reality. James Clerk Maxwell: A Commemoration 1831.

Einstein, A., 1956. The meaning of relativity, 5th Edition. Vol. 1921. Princeton University Press, Princeton.

Emery, W. J., Thomson, R. E., 2001. Data Analysis Methods in Physical Oceanography, 2nd Edition. Elsevier, Amsterdam.

Fowler, A. C., 2011. Mathematical geoscience. Vol. 36 of Interdisciplinary applied mathematics. Springer, London.

Fox-Kemper, B., Ferrari, R., Pedlosky, J., 2003. On the indeterminacy of rotational and divergent eddy fluxes. Journal of Physical Oceanography 33 (2), 478–483.

Ghil, M., Childress, S., 1987. Topics in geophysical fluid dynamics: atmospheric dynamics, dynamo theory, and climate dynamics. Vol. v. 60 of Applied mathematical sciences. Springer-Verlag, New York.

Gregory, J. M., Ingram, W., Palmer, M., Jones, G., Stott, P., Thorpe, R., Lowe, J., Johns, T., Williams, K., 2004. A new method for diagnosing radiative forcing and climate sensitivity. Geophysical Research Letters 31 (3).

Gutenberg, B., Richter, C. F., 1942. Earthquake magnitude, intensity, energy and acceleration. Bulletin of the Seismological Society of America 32, 163–191.

Haney, R. L., 1971. Surface thermal boundary conditions for ocean circulation models. Journal of Physical Oceanography 1, 241–248.

Hays, J., Imbrie, J., Shackleton, N., 1976. Variations in the earth's orbit: Pacemaker of the ice ages. Science 194 (4270), 1121–1132.

Holton, J. R., 2004. An introduction to dynamic meteorology, 4th Edition. Vol. v. 88. Elsevier Academic Press, Burlington, MA.

Imbrie, J., Berger, A., Shackleton, N., 1993. Role of orbital forcing: a two-million-year perspective. Global Changes in the Perspective of the Past, 263–277.

Imbrie, J., Imbrie, J. Z., 1980. Modeling the climatic response to orbital variations. Science 207 (4434), 943–953.

Ioannidis, J. P., 2005. Why most published research findings are false. PLoS medicine 2 (8), e124.

Kaper, H., Engler, H., 2013. Mathematics and climate. Vol. 131. Siam.

Knudsen, M., 1900. Ein hydrographischer lehrsatz. Ann. Hydrogr. Marit. Meteorol. 28, 316–320.

Kolmogorov, A. N., 1941. The local structure of turbulence in incrompressible viscous fluid for very large reynolds number. Dokl. Akad. Nauk. SSSR 30, 9–13.

Lahr, N. B., Reiter, C. A., 2014. Visualizing Newton's method on fractional exponents.

LeBlond, P. H., Mysak, L. A., 1978. Waves in the Ocean. No. 20 in Elsevier Oceanography. Elsevier Scientific Publishing Company, New York.

Longair, M., 2015. ... a paper... i hold to be great guns': a commentary on maxwell (1865)'a dynamical theory of the electromagnetic field. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 373 (2039), 20140473.

Lorenz, E. N., 1963. Deterministic nonperiodic flow. Journal of Atmospheric Sciences 20.

Mandelbrot, B. B., 1967. How long is the coast of britain. Science 156 (3775), 636–638.

Marshall, J., Plumb, R. A., 2008. Atmosphere, ocean, and climate dynamics: an introductory text. Vol. v. 93. Elsevier Academic Press, Amsterdam.

Maxwell, J. C., 1865. A dynamical theory of the electromagnetic field. Philosophical Transactions of the Royal Society of London, 459–512.

McWilliams, J. C., 2006. Fundamentals of geophysical fluid dynamics. Cambridge University Press, Cambridge.

Milankovitch, M., 1930. Mathematische klimalehre und astronomische theorie der klimashwankungengebruder borntraeger.

Ott, E., 2006. Basin of attraction. Scholarpedia.

Pedlosky, J., 1987. Geophysical fluid dynamics, 2nd Edition. Springer-Verlag, New York.

Pickard, G. L., Emery, W. J., 1990. Descriptive Physical Oceanography: An Introduction, 5th Edition. Butterworth Heinemann, Oxford.

Richter, C. F., 1935. An instrumental earthquake magnitude scale. Bull. Seism. Soc. Am 25 (1), 1–32.

Rundle, J. B., 1989. Derivation of the complete gutenberg-richter magnitude-frequency relation using the principle of scale invariance. Journal of Geophysical Research: Solid Earth (1978–2012) 94 (B9), 12337–12342.

Schubert, G., Turcotte, D. L., Olson, P., 2001. Mantle convection in the earth and planets. Cambridge University Press, Cambridge.

Sellers, W. D., 1969. A global climatic model based on the energy balance of the earth-atmosphere system. Journal of Applied Meteorology 8 (3), 392–400.

Snieder, R., 2004. A guided tour of mathematical methods for the physical sciences, 2nd Edition. Cambridge University Press, Cambridge, UK.

Stommel, H. M., 1961. Thermohaline convection with 2 stable regimes of flow. Tellus 13, 224–230.

Strogatz, S. H., 1994. Nonlinear dynamics and Chaos: with applications to physics, biology, chemistry, and engineering. Addison-Wesley Pub., Reading, Mass.

Taylor, G. I., 1938. The spectrum of turbulence. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences A 164 (919), 476–490.

Turcotte, D. L., 1997. Fractals and chaos in geology and geophysics, 2nd Edition. Cambridge University Press, Cambridge, U.K.

Turcotte, D. L., Schubert, G., 2002. Geodynamics, 2nd Edition. Cambridge University Press, Cambridge.

Vallis, G. K., 2006. Atmospheric and Oceanic Fluid Dynamics : Fundamentals and Large-Scale Circulation. Cambridge University Press, Cambridge.

Wilcox, D. C., 1997. Basic fluid mechanics. DCW Industries, La Cañada, Calif.

Wilks, D. S., 2011. Statistical methods in the atmospheric sciences, 3rd Edition. Vol. v. 100 of International geophysics series. Elsevier/Academic Press, Amsterdam.

Zaliapin, I., Ghil, M., 2010. Another look at climate sensitivity. Nonlinear Processes In Geophysics 17, 113–122.