Calculus Deconstructed

A Second Course in First-Year Calculus

Zbigniew H. Nitecki
Tufts University

May 28, 2008

Often I have considered the fact that most of the difficulties which block the progress of students trying to learn analysis stem from this: that although they understand little of ordinary algebra, still they attempt this more subtle art. From this it follows not only that they remain on the fringes, but in addition they entertain strange ideas about the concept of the infinite, which they must try to use.

Leonhard Euler
Introductio in Analysin Infinitorum (1755) [22, Preface]

Mathematicians are a kind of Frenchmen: whatever you say to them they translate into their own language and it is then something quite different.

Johann Wolfgang von Goethe
Maximen und Reflexionen

In memoriam

Martin M. Guterman

1941-2004

Colleague

Coauthor

Friend

Preface

Calculus is a collection of incredibly efficient and effective tools for handling a wide variety of mathematical problems. Underpinning these methods is an intricate structure of ideas and arguments; in its fullest form this structure goes by the name of *real analysis*. The present volume is an introduction to the elementary parts of this structure, for the reader with some exposure to the techniques of calculus who would like to revisit the subject from a more conceptual point of view.

This book was initially developed for the first semester of the Honors Calculus sequence at Tufts, which gives entering students with a strong background in high school calculus the opportunity to cover the topics of the three-semester mainstream calculus sequence in one year. It could, however, prove equally useful for a course using analysis as the topic for a transition to higher mathematics.

The challenge of the Honors Calculus course was twofold. On one hand, the great variability among high school calculus courses required us to cover all the standard topics again, to make sure students end up "on the same page" as their counterparts in the mainstream course. On the other hand, these bright students would be bored out of their minds by a straight repetition of familiar topics.

Keeping in mind these competing demands, I opted for a course which approaches the tools of calculus through the eyes of a mathematician. In contrast to "Calculus Lite" , the present book is "Calculus Tight": a review of often familiar techniques is presented in the spirit of mathematical *rigor* (hopefully without the *mortis*), in a context of ideas, and with some sense of their history. For myself as a mathematician, the intellectual exercise of working through all the theorems of single-variable calculus and constructing an elementary but coherent and self-contained presentation of this theory has been a fascinating experience. In the process, I have also become fascinated with the history of the subject, and I'm afraid my enthusiasm has infected the exposition.

Idiosyncracies

After three centuries of calculus texts, it is probably impossible to offer something truly new. However, here are some of the choices I have made:

Level: "Calculus" and "analysis"—which for Newton *et. al.* were aspects of the same subject—have in our curriculum become separate countries. I have stationed this book firmly on their common border. On one side, I have discussed at some length topics that would be taken for granted in many analysis courses, in order to fill any "gaps" in the reader's previous mathematical experience, which I assume to be limited to various kinds of calculation. I have also limited myself to mathematical topics associated with the basics of calculus: advanced notions such as connectedness or compactness, metric spaces and function spaces demand a level of sophistication which are not expected of the reader, although I would like to think that this book will lead to attaining this level. On the other side, I have challenged—and hopefully intrigued—the reader by delving into topics and arguments that might be regarded as too sophisticated for a calculus course. Many of these appear in optional sections at the end of most chapters, and in the *Challenge Problems* and *Historical Notes* sections of the problem sets. Even if some of them go over the heads of the audience, they provide glimpses of where some of the basic ideas and concerns of calculus are headed.

Logic: An underlying theme of this book is the way the subject hangs together on the basis of mathematical arguments. The approach is relatively informal. The reader is assumed to have little if any previous experience with reading definitions and theorems or reading and writing proofs, and is encouraged to learn these skills through practice. Starting from the basic, familiar properties of real numbers (along with a formulation of the Completeness Property)—rather than a formal set of axioms—we develop a feeling for the language of mathematics and methods of proof in the concrete context of examples, rather than through any formal study of logic and set theory. The level of mathematical sophistication, and the demands on the proof-writing skills of the reader, increase as the exposition proceeds. By the end of the book, the reader should have attained a certain level of competence, fluency and self-confidence in using the rhetorical devices of mathematics. A brief appendix at the end of the book (Appendix A) reviews some basic methods of proof: this is intended as a reference to supplement the discussions in the text with an overview.

Limits: The limit idea at the heart of this treatment is the limit of a *sequence*, rather than the standard $\varepsilon - \delta$ limit of a function (which is presented in the optional § 3.7). At the level of ideas, this concept is much easier to absorb, and almost everything we need can be formulated naturally using sequences.

Logarithms and exponentials: Most rigorous treatments of calculus define the logarithm as an integral and the exponential as its inverse. I have chosen the ahistorical but more natural route of starting with natural powers to define exponentials and then defining logarithms as their inverses. The most difficult step in this "early transcendentals" approach, from a rigor point of view, is the differentiability of exponentials; see § 4.3.

History: I have tried to inject some of the history of various ideas, in some narrative at the start of each chapter and in exercises denoted *History Notes*, which work through specific arguments from the initiators of the subject. I make no claims for this as a historical work; much of the discussion is based on secondary sources, although where I have had easy access to the originals I have tried to consult them as well.

Technology: There are no specific references to technology in the text; in my class, the primary tools are paper and pencil (or blackboard and chalk). This does not, of course, preclude the use of technology in conjunction with this book. I would, in fact, love to hear from users who have managed to incorporate computer or graphing calculator uses into a course based on the present exposition.

How to use this book

Clearly, it is impossible to cover every proof of every result (or anything near it) in class. Nevertheless, I have tried to give the reader the resources to see how everything is supported by careful arguments, and to see how these arguments work. In this sense, the book is part textbook, part background reference. The beginning of each chapter sketches the highlights of the topics and tools to come from a historical perspective. The text itself is designed to be read, often with paper and pencil in hand, balancing techniques and intuition with theorems and their proofs.

The exercises come in four flavors:

Practice problems are meant as drill in techniques and intuitive exploration of ideas.

Theory problems generally involve some proofs, either to elaborate on a comment in the text or to employ arguments introduced in the text in a "hands on" manner.

Challenge problems require more ingenuity or persistence; in my class they are optional, extra-credit problems.

Historical notes are hybrids of exposition and exercise, designed to aid an active exploration of specific arguments from some of the originators of the field.

Of course, my actual homework assignments constituted a selection from the exercises given in the book.

A note at the start of each set of exercises indicates problems for which answers have been provided in Appendix B; these are almost exclusively Practice problems. Solutions to *all* problems, including proofs, are given in the **Solution Manual** accompanying this text.

On average, one section of text corresponds to one hour in the classroom. The last section of each chapter after the first is optional (and usually skipped in my class), although the notion of uniform continuity from § 3.7 appears in the rigorous proof of integrability of continuous functions.

Acknowledgements

First of all I want to thank two wonderful classes of students who helped me develop this book, not necessarily by choice: a small cadre acted as guinea pigs for the "alpha" version of this book—Brian Blaney, Matt Higger, Katie LeVan, Sean McDonald and Lucas Walker—and a larger group—Lena Chaihorsky, Chris Charron, Yingting Cheng, Dan Eisenberg, Adam Fried, Forrest Gittleson, Natalie Koo, Tyler London, Natalie Noun, A.J. Raczkowski, Jason Richards, Daniel Ruben, and Maya Shoham—were subjected to the "beta" version, very close to the one here: they enthusiastically rooted out misprints, errors and obfuscations. I would like to particularly single out Jason Richards, who repeatedly questioned and suggested changes to a number of proofs, first as a student in the "beta" class and then as a homework grader in the course for several years.

I would also like to thank a number of colleagues for helpful conversations, including Fulton Gonzalez, Diego Benardete, Boris Hasselblatt, Jim Propp, George McNinch, and Lenore Feigenbaum.

Second, I want to acknowledge extensive email help with TeXnical issues in the preparation of this manuscript from Richard Koch, who patiently explained features of TeXShop to me, and to numerous contributors to `support@tug.org`, especially E. Krishnan, Dennis Kletzing, Lars Madsen, Tom DeMedts, Herbert Voss and Barbara Beaton, who helped me with a cornucopia of special packages.

Third, I would like to acknowledge the inspiration for the historical approach which I gained from Otto Toeplitz's *Calculus, a Genetic Approach* [52]. I have drawn on many sources for historical information, above all Charles Edwards' wonderful exposition, *The Historical Development of the Calculus* [20]. I have also drawn on Michael Spivak's *Calculus* [50] and Donald Small & John Hosack's *Calculus, an Integrated Approach* [48]; the latter was for many years the text for our course at Tufts. A more complete list of sources and references is given in the Bibliography at the end of this book. One mild disclaimer: when I could, I listed in the Bibliography an original source for some of my historical notes: such a listing does not necessarily constitute a claim that I have been able to consult the original directly. Citations in the text of the notes indicate most of my sources for my statements about history.

Finally, a word about Marty Guterman, to whose memory this work is dedicated. During the thirty or so years that we were colleagues at Tufts, and especially in the course of our joint textbook projects, I learned a great deal from this extraordinary teacher. His untimely death occurred before the present project got underway, but we had frequently discussed some of the ideas that have been incorporated in this text. His reactions and comments would have been highly valued, and he is missed.

Z. Nitecki

Contents

ἀγεωμέτρητος μηδει΄ς εἰσι΄τω
Let no one ignorant of geometry enter here.

Inscription over the entrance to Plato's Academy

Whoever thinks algebra is a trick in obtaining unknowns has thought it in vain. No attention should be paid to the fact that algebra and geometry are different in appearance. Algebras are geometric facts which are proved.

Omar Khayyam (*ca.* 1050-1130) [10, p. 265]

1 Precalculus

Mention "math" to most people, and they think "numbers". Columns of figures in the financial pages of the newspaper. Computer printouts. Sequences of digits in the calculator window.

Originally, "digital information" meant counting on your fingers, but years of arithmetic have extended our idea of numbers. The sophisticated notion of **real numbers** includes negatives, fractions, square roots, logarithms, and exotic items like π and e. Really, though, we know numbers by what we can do to them: add, subtract, multiply, divide, and compare. We feel a number has been given to us concretely when it has a decimal form, like

$$3.141592654\ldots$$

which we can use for calculation or "number-crunching".

It might be useful to reflect briefly on the history of these ideas, for our notion of a "number" is a relatively new concept.

1.1 Numbers and Notation in History

To the Greeks in the time of Pythagoras of Samos (*ca.* 580-500 BC) and his followers, "numbers" were counting numbers, and there was a presumption that all quantities could be described starting from "numbers"[1]–that is, they

[1] Numbers were represented by pebbles, or *calculi*—the origin of our word "calculus".

were *commensurable*. The discovery (probably around 410 BC) that the diagonal of a square is not commensurable with its side (in other words, that $\sqrt{2}$ is irrational–see p. 20) shook the Hellenic mathematical community. In reaction, the followers of Plato (429-348 BC) developed a purely geometric theory of magnitudes, in which (counting) numbers were distinct from geometric magnitudes like lengths, areas, and volumes; these quantities were compared by means of *ratios*.[2] For example, where we might give a formula for the area of a circle, the Platonists would say that the *ratio* between (the areas of) two *circles* is the same as the *ratio* between (the areas of) the *squares* built on their respective diameters (see Exercises 7-9 in § 5.1).

The *Elements* of Euclid[3] (ca. 300 BC), a summary of the mathematics of the Platonists, was (and continues to be) read and commented on by numerous scholars, probably more frequently than any book except the Old and New Testaments. It presented the mathematics of its time in a strictly deductive style which set the standard for mathematical rigor until the nineteenth century. Its geometric approach to numbers continued to be the primary language of mathematical argument up to and including Newton's *Principia* (1687)[41].

It should be noted that the geometric quantities considered by the Greeks were by their nature what we call *positive* numbers; the introduction of zero as a number, as well as negative numbers and our digital notation, appears to have occured in Indian mathematics [10, pp. 234-7, 261], [4, pp. 61-5]: the place-value notation we use, in which the positions of digits indicate multiplication by different powers of ten, was invented by 595AD, while the introduction of zero (looking like our 0) as a place-holder first occurs in an inscription of 876 AD. Brahmagupta (ca. 628 AD) treated zero as a number, giving rules for arithmetic using it, and treated quadratic equations with negative as well as positive roots.

The Romans had little interest in abstract mathematics, and in Western Europe from the fall of Rome in 476 AD until the twelfth century AD the Greek mathematical tradition was more or less lost; it was preserved, along with the Indian contributions noted above, by Byzantine and Islamic scholars and gradually came to Western Europe during the tenth through twelfth centuries AD. The contributions of Arabic mathematics included two books of Al-Khowârizmî (*ca.* 780-*ca.* 850 AD) (who worked in Baghdad): the first was an exposition of Indian arithmetic, while the second

[2]This interpretation of the effect of the discovery of irrationality is based on commentaries written in later classical times. It has recently been seriously questioned [25].

[3]Euclid taught mathematics at the Museum, an institute founded in Alexandria by Ptolemy I, the successor of Alexander the Great.

concerns the solution of linear and quadratic equations: its title, "*Hisâb al-jabr Wa'lmuquâbalah*", is the origin of our word "algebra" (via Chester's translation–see below); our word "algorithm" also apparently comes from a corruption of this author's name [10, pp. 251-8]. Some of these ideas were extended in a book on algebra by Omar Khayyam (*ca.* 1050-1130)[4] who looked at equations of higher degree [10, pp. 264-6].

During the twelfth century, a number of translations from Arabic to Latin were produced; for example of the *Elements* by Adelard of Bath (1075-1160) in 1142, and of the *Al-jabr* by Robert of Chester in 1145, under the Latin title *Liber algebrae et almucabola*. An influential exposition of Indian-Arabic arithmetic was the *Liber Abaci* by Leonardo of Pisa (ca. 1180-1250), better known as "Fibonacci". In the sixteenth century further progress was made in algebra; the *Ars Magna* of Girolamo Cardan (1501-1576) publicized earlier methods for solving equations of degree three and four; however, like Al-Khowarizmi, Cardan and his contemporaries did not have our literal symbolism; it is instructive to read the purely verbal exposition of the equation $x^2 + 10x = 39$ [51, pp. 56-60]; for a modern reader, it is very difficult to follow. The decisive step of introducing letters for unknowns (as well as for general coefficients) was made by François Vièta (1540-1603) in his *In artem analyticem isagoge* (Introduction to the Analytic Art) of 1591 [51, pp. 75-81]. The relation between coefficients of algebraic equations and their roots was studied by the English Thomas Harriot (1560-1621) and the Flemish Albert Girard (1590-1663)—the latter allowed negative and even imaginary roots in his 1629 book *Invention nouvelle en l'algèbre* [10, pp. 333-7]. John Napier (1550-1617), a Scottish laird, invented logarithms around 1594, as a computational tool in the form of rods that worked much like a slide rule [10, pp. 342-7]; our decimal notation for fractions first appeared in an English translation of one of his works in 1616; the idea of decimal notation for fractions (as opposed to just whole numbers) had been suggested earlier by the Flemish Simon Stevin (1548-1620) in *De thiende* ("The Tenth") (1585). Stevin also introduced a prototype of our superscript notation for powers, although the version we use first appeared in Descartes' *La Geometrie* (1637) and was further popularized by Newton in his *Principia* (1687). The use of negative and fractional exponents was pioneered by John Wallis (1616-1703).

[4] Yes, *that* Omar Khayyam (*A loaf of bread, a jug of wine, and thou*).

1.2 Numbers, Points, and the Algebra of Inequalities

Geometric imagination can help us think about numbers. As we saw in the preceding section, the Greeks built their mathematics almost exclusively on geometry, but the kind of picture we have of numbers today is very much tied up with the idea of coordinate systems and graphs. This idea can be traced back to the scholars of Merton College, Oxford who first tried to formulate quantitative concepts of motion in the second quarter of the fourteenth century, and in particular to the Parisian scholar Nicole Oresme (1323-1382) who took up these ideas and suggested the idea of a graphical representation of a varying quantity by vertical line segments over the points of a reference interval [20, pp. 88-90], [10, pp. 290-2]. The full flowering of this idea occured in the work of Pierre de Fermat (1601-1665) and René Descartes (1596-1650), who developed the idea of the graph of an equation and used it to develop some general theories of equations. The term **cartesian coordinates** comes from the latter's name.

We picture real numbers as points on the **number line**, with 0 in the middle, the other natural numbers $1, 2, 3, \ldots$ regularly spaced to the right, negatives to the left, fractions at in-between division points. Every point

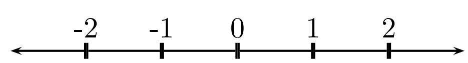


Figure 1.1: The number line

on this line is associated to a different real number; in fact we think of the two as the same thing, imagining numbers as decimal expressions when we calculate and as points on the number line when we compare them.

We compare numbers in two ways: relative position and size. We say x is **less** than y (*resp.* y is **greater** than x), and write $x < y$ (*resp.* $y > x$) if x lies to the left of y on the number line. These are *strict* inequalities: to express the weak inequality that allows also $x = y$, we write $x \leq y$ (*resp.* $y \geq x$). While the proper pronounciation of $x \leq y$ is "*x is less than or equal to y*", we are sometimes sloppy in English, saying "*x is less than y*" when we mean $x \leq y$; in such a context, we can specify $x < y$ by the phrase "*x is **strictly** less than y*". The *symbols*, however, are always interpreted precisely. A number is **positive** (*resp.* **negative**) if it is greater than (*resp.* less than) 0. Again, this is sometimes used in a sloppy way, but the correct terminology for $x \geq 0$ (*resp.* $x \leq 0$) is to call x **non-negative** (*resp.* **non-**

positive).

The relation between inequalities and arithmetic is slightly more complicated than for equations. We can add inequalities which go in the same direction:

if $a \leq b$ and $c \leq d$, then $a + c \leq b + d$,

and we can multiply both sides of an inequality by a *strictly positive* number

if $a < b$ and $c > 0$, then $ac < bc$.

However, switching signs can reverse an inequality

if $a \leq b$ then $-a \geq -b$

as can taking reciprocals (for *positive* numbers)

if $0 < a < b$ then $\frac{1}{a} > \frac{1}{b} > 0$

(see exercise 13).

Other relations between arithmetic and inequalities are given in Exercises 8-13.

We measure the size of a real number x by its **absolute value** $|x|$. When x is expressed as a decimal, possibly preceded by a (plus or) minus sign, then $|x|$ is the non-negative number expressed by dropping the sign. In general, $|x|$ is either x or $-x$, whichever is non-negative. This is specified in the formula

$$|x| = \begin{cases} x & \text{if} \quad x \geq 0 \\ -x & \text{if} \quad x < 0 \end{cases}$$

For example,

$$|-3| = -(-3) = 3.$$

The absolute value of x is the greater of the two numbers x and $-x$ (they are equal precisely if $x = 0$)

$$-x, x \leq |x| \text{ for any } x.$$

This means that for a non-negative number $c \geq 0$ the statement

$$|x| \leq c$$

is the same as the pair of inequalities

$$-x \leq c \quad \text{and} \quad x \leq c$$

or equivalently

$$-c \leq x \leq c.$$

Two important relations between absolute value and arithmetic are given in the following.

Proposition 1.2.1. *For any two real numbers x and y,*

1. $|x + y| \leq |x| + |y|$
2. $|xy| = |x| \, |y|$.

Proof. To see the first statement, we will use the observation above. Add together the inequalities $x \leq |x|$ and $y \leq |y|$ to get

$$x + y \leq |x| + |y| \, ,$$

but also note that adding $-x \leq |x|$ to $-y \leq |y|$ gives

$$-(x + y) = (-x) + (-y) \leq |x| + |y| \, .$$

These two inequalities are the version of the first statement given by substituting $c = |x| + |y|$ in our earlier observation.

We leave the proof of the second statement to you (Exercise 14). □

The absolute value is used to calculate the **distance** between real numbers. We take the space between successive natural numbers as our unit of distance. Then the distance between any two real numbers is the size of their difference:

$$\operatorname{dist}(x, y) = |x - y| \, .$$

Note that in particular the absolute value of any real number is the same as its distance from 0:

$$|x| = \operatorname{dist}(x, 0).$$

The following summarizes some important properties of the distance function.

Proposition 1.2.2 (Properties of distance:)**.** *Let x, y, z be any real numbers. Then*

1. ***positive-definite:*** $\operatorname{dist}(x, y) > 0$ *unless $x = y$, in which case* $\operatorname{dist}(x, y) = 0$.
2. ***symmetric:*** $\operatorname{dist}(x, y) = \operatorname{dist}(y, x)$.

3. ***triangle inequality:*** $\text{dist}(x, y) \leq \text{dist}(x, z) + \text{dist}(z, y)$.

The positive-definite property says that we can tell whether or not two numbers are the same by deciding whether the distance between them is zero or positive; we shall find this useful later in dealing with numbers defined by limits. The symmetric property says that the distance "from here to there" is the same as the distance "from there to here". These properties are both easy to verify. The triangle inequality says that the distance from one point to another is never more than the distance travelled when we go via a third point. If we take the inequality in Proposition 1.2.1 but replace x with $x - z$ and y with $z - y$, then $x + y = (x - z) + (z - y)$ gives us the triangle inequality.

Exercises for § 1.2

Answers to Exercises 1, 3, 5, 6ace, 7(1,3) are given in Appendix B.

Practice problems:

1. Which number is bigger, $\sqrt{5} - \sqrt{3}$ or $\sqrt{15} - \sqrt{13}$? How about $\sqrt{15} - \sqrt{13}$ and $\sqrt{115} - \sqrt{113}$? Of course, you can check the answer using a calculator or computer, but you are asked to verify it using only what you know about whole numbers. (*Hint:* Use the identity $a^2 - b^2 = (a - b)(a + b)$.)

2. Which number is closer to 1: $\frac{3}{4}$ or $\frac{4}{3}$? Can you discern any general pattern here? That is, which positive numbers are closer to 1 than their reciprocals?

3. Given a fraction with positive numerator and denominator, suppose you increase the numerator and denominator by the same amount. Does the value of the fraction go up or down? Can you prove your assertion?

4. Suppose m and n are integers, with $n \neq 0$.

   (a) Show that
   $$\frac{m^2}{n^2} = \frac{m}{n}$$
   precisely if either $m = 0$ or $m = n$.

(b) When does the inequality

$$\frac{m^2}{n^2} < \frac{m}{n}$$

hold?

5. Suppose you know $|a| > c > 0$. What can you conclude about the possible values of a?

Theory problems:

6. Each of the following statements is true whenever *all* the letters represent *strictly positive* numbers. For each one, either give an example (involving some negative numbers) for which it fails, or else explain why it is true for all (nonzero) numbers.

   (a) If $a < b$ then $|a| < |b|$.
   (b) If $a < b$ then $\frac{1}{a} > \frac{1}{b}$.
   (c) If $a < b$ and $c < d$ then $ac < bd$.
   (d) If $a < b$ then $-a > -b$.
   (e) $|a - b| \leq |a| + |b|$
   (f) $|a - b| \geq |a| - |b|$.

7. Translate each of the three assertions in Proposition 1.2.2 into a statement about absolute values, and explain why it is true.

8. *Show:* If $a - b > 0$, then $a > b$. (*Hint:* Add an appropriate quantity to both sides.)

9. *Show:* If $a \leq b$ then $-a \geq -b$. (*Hint:* Add $-a - b$ to both sides.)

10. *Show:* If $0 < a < b$ and $0 < c \leq d$, then $ac < bd$. (*Hint:* Compare both quantities to ad.)

11. Use exercise 10 to show that for any two *non-negative* numbers a and b, if $a^2 < b^2$, then $a < b$.

12. *Show:* If $0 < a < b$, then $\frac{b}{a} > 1$. (*Hint:* Multiply both sides by an appropriate quantity.)

13. *Show:* If $0 < a < b$, then $\frac{1}{a} > \frac{1}{b}$. (*Hint:* Consider the previous problem.)

14. *Show:* For any two real numbers a and b, $|ab| = |a|\,|b|$. (*Hint:* Consider the possible sign configurations.)

Challenge problems:

15. *Show:* For any two real numbers a and b, $|2ab| \leq a^2 + b^2$. (*Hint:* Use the fact that $(a+b)^2$ and $(a-b)^2$ are both non-negative.)

16. Suppose we have a collection of fractions

$$\frac{a_1}{b_1} \leq \frac{a_2}{b_2} \leq \cdots \leq \frac{a_n}{b_n}$$

with all denominators b_i positive.

(a) *Show* that
$$\frac{a_1}{b_1} \leq \frac{a_1 + a_2}{b_1 + b_2} \leq \frac{a_2}{b_2}.$$

(b) *Show* that
$$\frac{a_1}{b_1} \leq \frac{a_1 + a_2 + \cdots + a_n}{b_1 + b_2 + \cdots + b_n} \leq \frac{a_n}{b_n}.$$

(*Hint:* Use mathematical induction (Appendix A).)

This result was given by Cauchy in his *Cours d'analyse* (1821) and used in his proof of what is now called the Cauchy Mean Value Inequality (Exercise 14 in § 4.9).

1.3 Intervals

When we deal with functions, we often need to consider sets of real numbers. The most common kind of set is the interval between two numbers. Suppose a and b are real numbers with $a \leq b$. The **closed interval** with **left endpoint** a and **right endpoint** b, denoted

$$[a, b]$$

consists of all real numbers x that satisfy the pair of weak inequalities

$$a \leq x \leq b.$$

(See Figure 1.2.) Note that the endpoints a and b both belong to the closed interval $[a, b]$. We write

Figure 1.2: The closed interval $[a, b]$

$$x \in [a, b]$$

to indicate that x is an **element** of $[a, b]$—that is, it satisfies both of the inequalities above. If $a < b$, the **open interval** with endpoints a and b, denoted

$$(a, b)$$

consists of all numbers x satisfying the *strict* inequalities

$$a < x < b.$$

(See Figure 1.3.) Note that the endpoints do *not* belong to the open

Figure 1.3: The open interval (a, b)

interval (a, b). The notation for an open interval looks just like the coordinates of a point in the plane, but the context will usually make it clear which we mean.
There are also **half-open intervals** $[a, b)$ and $(a, b]$ which include one endpoint but not the other (figure 1.4):

$$\begin{array}{lll} x \in [a, b) & \text{if} & a \leq x < b \\ x \in (a, b] & \text{if} & a < x \leq b \end{array}$$

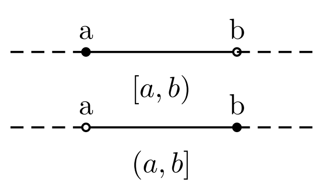


Figure 1.4: Half-open intervals

In addition to the interval between two numbers, we will consider intervals consisting of *all* numbers to the right (or left) of a given number; our notation can handle such intervals by means of the symbols ∞ and $-\infty$. Think of these as the right and left "ends" of the number line respectively. *Every* real number x satisfies $x < \infty$ and $-\infty < x$. Thus formally, for any number a, the notation

$$(a, \infty)$$

specifies the set of all numbers x satisfying $a < x$ and $x < \infty$. But since the second inequality holds for *every* number x, the specified set consists simply of all numbers x satisfying $a < x$, that is, all numbers to the right of a. By the same token,

$$(-\infty, a)$$

denotes the set of all x with $x < a$, and

$$(-\infty, \infty)$$

which formally denotes the set of numbers x for which $-\infty < x$ and $x < \infty$, is simply another notation for the whole number line. Another

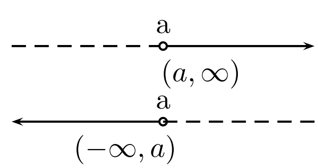


Figure 1.5: Unbounded intervals

notation for the set of all real numbers is $\mathbb{R}$. We shall often write $\boldsymbol{x} \in \mathbb{R}$ to mean either "x is a real number" or "x is *any* real number". Similar interpretations apply to the half-open intervals $[a, \infty)$ and $(-\infty, a]$. Note, however, that ∞ and $-\infty$ are *not* real numbers, so we *never* write notations like $[1, \infty]$ or $[-\infty, 3)$.
We distinguish intervals with *real* endpoints $a, b \in \mathbb{R}$ from those with "infinite endpoints" via the property of *boundedness*:

Definition 1.3.1. *Let I be any interval (open, closed, or half-open). An **upper bound** for I is any number β such that*

$$x \leq \beta$$

for every point x in I ($x \in I$)—this means β lies to the right *of I. A **lower bound** for I is any number α such that*

$$\alpha \leq x$$

for every $x \in I$.
An interval I is ***bounded*** *if there exist BOTH an upper bound for I and a lower bound for I.*

If a and b are the endpoints of I, then any number $\beta \geq b$ (*resp.* $\alpha \leq a$) is an upper (*resp.* lower) bound for I.
An interval with an infinite "endpoint" is not bounded: for example, if $I = [1, \infty)$ there is no upper bound for I, since whatever real number β we consider, it is always possible to find a point $x \in I$ for which the inequality $x \leq \beta$ *fails*.

Exercises for § 1.3

Answers to Exercises 1aceg, 2acegi, 5ad are given in Appendix B.

Practice problems:

1. Sketch each interval below on the number line, using the pictures in the text as a model, or explain why the notation doesn't make sense:

   (a) $[2, 3]$ (b) $(2, 4)$ (c) $(2, \infty)$ (d) $[1, 4)$
   (e) $[3, 2)$ (f) $(1, 3]$ (g) $(-\infty, 1]$ (h) $[-1, \infty]$

Theory problems:

2. For each of the following, either give an example or explain why no such example exists. Use the interval notation of this section to specify your examples.

   (a) Two intervals with no point in common. (Such intervals are called **disjoint**.)
   (b) Two disjoint intervals whose union is $[0, 1]$. (The **union** of two intervals I and J, denoted $I \cup J$, is the collection of points which belong to *at least one* of the two intervals I and J.)
   (c) Two disjoint *closed* intervals whose union is $[0, 1]$.
   (d) Two disjoint *open* intervals whose union is $(0, 1)$.
   (e) A *bounded* interval which contains an *unbounded* interval.
   (f) An open interval I and a closed interval J such that[5] $I \subset J$.

[5] The notation $I \subset J$ means I is a *subset* of J, in other words, I is *contained in* J.

(g) An open interval I and a closed interval J such that $J \subset I$.

(h) An unbounded interval which does not contain 0.

(i) Two disjoint unbounded intervals.

3. The **intersection** of two intervals I and J, denoted $I \cap J$, is the collection of all points that belong to *both* I *and* J. Suppose $I = [a, b]$ and $J = [c, d]$.

   (a) If $b = d$, show that $I \cap J = [\alpha, b]$, where $\alpha = \max(a, c)$.

   (b) If $a = c$ (without assuming $b = d$), show that $I \cap J$ is a closed interval. What are its endpoints?

   (c) *Show:* If $b < c$ or $d < a$, then I and J are disjoint (*cf* Exercise 2)–there are no points in $I \cap J$. (We say $I \cap J$ is the **empty set**, and write $I \cap J = \emptyset$.)

   (d) *Show:* If $b \geq c$ and $a \leq d$, then $I \cap J$ is a closed interval. What are its endpoints?

4. (a) *Show:* If I is an interval and $x, y, z \in \mathbb{R}$ with $x < z < y$ and $x, y \in I$, then $z \in I$.

   (b) Suppose S is a collection of real numbers with the following properties:

      i. There is a number $a \in S$ such that $a \leq x$ for every number $x \in S$. (a is called the **minimum** of S.)

      ii. There is a number $b \in S$ such that $b \geq x$ for every number $x \in S$. (b is called the **maximum** of S.)

      iii. If $x, y \in S$ then every number in between ($x \leq z \leq y$) also satisfies $z \in S$.

   Then show that $S = [a, b]$.

5. (a) Give an example of two intervals whose union is *not* a interval.

   (b) *Show:* If an endpoint of the interval J belongs to the interval I, then $I \cup J$ is an interval. (*Hint:* use Exercise 4b.)

   (c) *Show:* If *some* point of J (not necessarily an endpoint) belongs to I, then $I \cup J$ is an interval.

   (d) Give an example of two intervals I and J with $I \cap J = \emptyset$ for which $I \cup J$ is an interval. (*Hint:* Intervals can be open or closed.)

Errores quam minimi in rebus mathematicis non sunt contemnendi.
The very smallest errors in mathematical matters are not to be neglected.

Isaac Newton
Tractatus de quadratura curvarum (1693)
transl. J. Stewart (1745)

You can't always get what you want
but if you try sometime
you might just find
that you get what you need.

Mick Jagger and Keith Richards

2 Sequences and their Limits

The three fundamental ideas of calculus are *limits*, *derivatives*, and *integrals*. Of these, the notion of limit is most basic, since the other two are formulated as different kinds of limits. You have probably worked with limits of functions. However, on the level of ideas, limits of *sequences* are easier to think about. We will devote some time to understanding limits in this context as a warmup for looking at functions; in fact, limits of sequences will be a central notion throughout this book.

The Greeks did not have the idea of a limit *per se*, and in fact avoided trusting in infinite processes of any kind. However, they did have a technique for getting at certain ratios by what we would now think of as a sequence of successive approximations to a desired quantity. This was used primarily to study problems that we now think of as integration, namely the *rectification* and *quadrature* of curves; the first means finding the length (or in their terms, finding a straight line segment of the same length) of a given curve and the second means finding the area (*i.e.*, a rectangular figure with the same area) bounded by a curve. Some idea of how they proceeded is given in Exercises 7-9 and 11-13 in § 5.1. The method was called the **method of exhaustion**: its first version, due to Eudoxus of Cnidos (408-355 BC) and expounded in Euclid's *Elements*, consists of inscribing polygons in a given region and showing that the area inside the region which is not filled by a polygon can be made smaller than any specified number by giving the polygon an appropriately high number of sides. A more sophisticated version is the **method of compression** of Archimedes of Syracuse (*ca.* 212-287 BC) in which also *circumscribed* polygons are shown to have an excess area that can be made as small as possible. This is really very close to our definition of convergence of a sequence (Definition 2.2.2).

2.1 Real Sequences

A **sequence in** $\mathbb{R}$ (which we can also refer to as a **real sequence**) is just an ordered list of numbers. A finite sequence of n real numbers is essentially an n-tuple, but to consider limits we need *infinite* sequences, so whenever we say "sequence", we mean that it is infinite unless it is explicitly called finite.

Some sequences are understood once we establish a pattern with the first few terms. For example,

$$1, 2, 3, 4, \ldots$$

is normally recognized as the sequence of positive integers, while

$$\frac{1}{1}, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$$

is understood to be the reciprocals of powers of 2. This is, however, an unreliable method; we prefer a specification that *explicitly* tells us how to find the value of *each* number in the sequence. This is conventionally given in one of two notations.

The first of these uses a subscripted letter for each element. For example, the sequence of positive integers could be given as

$$x_i = i, \quad i = 1, 2, 3, ...$$

Here, the subscript (called the **index**) indicates the position of a given number in the sequence. The index need not start at 1; any integer can be used as the initial index value. For example, the second sequence above is most conveniently expressed as

$$y_n = \frac{1}{2^n}, \quad n = 0, 1, 2, ...$$

However, it is *always* assumed that, starting from its initial value, the index goes through *successive integer* values (skipping none). It is *not* permissible to express the odd integers as

$$a_j = j, \quad j = 1, 3, 5, ... \quad ☹$$

rather we must find a trick for getting successive *odd* values out of *successive integer* values, for example

$$a_j = 2j + 1, \quad j = 0, 1, 2, ...$$

or

$$a_j = 2j - 1, \quad j = 1, 2, ...$$

Notice that different *formulas* can give the same *sequence*: what matters is that the *same numbers* appear in the *same order*.
A second notation places the formula for an element in terms of the index inside braces, usually with a subscript and superscript indicating the range of the index. The sequence of odd integers could be written

$$\{2j + 1_j\}_{j=0}^{\infty} \quad \text{or} \quad \{2j - 1_j\}_{j=1}^{\infty}.$$

(The superscript "∞" indicates that the index goes on forever—*i.e.*, the sequence is infinite.)
These sequences were specified in **closed form**, telling us how to determine the value of each element directly from its index. Some sequences are more naturally specified by a **recursive formula**, in which an element is calculated from its immediate predecessor(s) in the sequence. For example, in a savings account earning 5% interest (compounded annually), the balance b_m at the end of year m can be calculated from that at the end of the previous year via the formula

$$b_m = (1.05)b_{m-1}.$$

Note that this doesn't help us calculate the *starting* balance (since it has no predecessor); we need to determine the value of b_m when m is at its initial value "by hand".
Some recursive formulas are easily translated into closed form. For example, the sequence above could be specified by

$$b_m = (1.05)^m b_0$$

(we still have to specify b_0). However, other recursive formulas don't translate easily into closed form. For example, the sequence defined recursively by

$$p_0 = \frac{1}{2}$$
$$p_{i+1} = p_i^2 - p_i$$

does *not* reduce to a closed form definition: the only way to calculate, say, p_{100} is to calculate each of the 100 preceding elements $p_0, p_1, \ldots p_{99}$ one-by-one, using the recursive formula. Such a procedure is difficult to carry out for more than a few steps by hand. "Do loops" and other programmed procedures on computers allow us to calculate vastly more—but never *all*—elements of the sequence. To say something about *all* elements of the sequence (or about its "eventual" behavior), we need more abstract tools.

A particularly important kind of recursive definition is illustrated by the formula

$$\begin{aligned} S_0 &= 1 \\ S_{n+1} &= S_n + \frac{1}{2^n}, \quad n = 0, 1, ... \end{aligned}$$

The first few terms of this sequence are

$$\begin{aligned} S_0 &= 1 \\ S_1 &= 1 + \frac{1}{2} &&= \frac{3}{2} \\ S_2 &= 1 + \frac{1}{2} + \frac{1}{4} &&= \frac{7}{4} \\ S_3 &= 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} &&= \frac{15}{8} \\ &\vdots && \vdots \end{aligned}$$

A more efficient way to write this kind of expression is **summation notation**. For example, the formula

$$S_n = \sum_{i=0}^{n} \frac{1}{2^i}$$

is pronounced "S_n equals the sum[1] from $i = 0$ to $(i =)n$ of $\frac{1}{2^i}$". This kind of sequence is called a **series**. In effect, we are dealing with *two* sequences here. First, the infinite list of numbers being added, called the **terms** or **summands** of the series, is the sequence $\left\{\frac{1}{2^i} i\right\}_{i=0}^{\infty}$. This is really the data we are given, but what we are interested in is the sequence of **partial**

[1] Σ is the Greek letter "sigma", corresponding to the first (Latin) letter of the word "sum".

sums S_n. Sometimes a series is written "directly" as an infinite sum using summation notation: the series above would in this way be denoted

$$\sum_{i=0}^{\infty} \frac{1}{2^i}.$$

See Figure 2.1

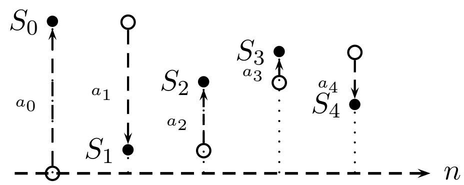


Figure 2.1: Partial sums of the series $\sum_{k=0}^{\infty} a_k$

Series play an important role in many areas of mathematics. The problem of deducing the behavior of the sequence of partial sums from the sequence of summands can be difficult, often involving subtle tricks. We shall touch upon some of these as we proceed.

Exercises for § 2.1

Answers to Exercises 1-27 (odd only) are given in Appendix B.

Practice problems:

The list of sequences $1 - 27$ below will be referred to again in the exercises to later sections. For each sequence on this list, write down all the terms with index up to and including 5, in an exact form and as a decimal expression correct to two decimal places. (Some of these may require a calculater or computer.)[2]

1. $a_j := 2j + 1$, $j = 0, ...$
2. $b_i := \frac{i}{i+1}$, $i = 0, ...$
3. $c_k := \frac{1}{k}$, $k = 1, ...$
4. $d_i := \frac{2^{i+1}}{3^i}$, $i = 0, ...$
5. $e_n := \frac{1}{n!}$, $n = 0, ...$
6. $f_j := \frac{j}{j+1} - \frac{j+1}{j}$, $j = 1, ...$

[2]In Exercises 5 and 17, $n!$ means n *factorial*: the product of all integers from 1 to n, if n is a positive integer, and $0! = 1$. In Exercises 9-12, remember that $\sin\theta$ means that the angle has size θ *radians*–see p. 94.

7. $g_k := (-2)^{-k}$, $k = 0, ...$
8. $x_k := \frac{k^2}{k+1}$, $k = 0, ...$
9. $s_n := \sin \frac{n\pi}{2}$, $n = 0, ...$
10. $\alpha_n := \sin n$, $n = 0, ...$
11. $\beta_n := \sin \frac{\pi}{n}$, $n = 1, ...$
12. $\gamma_k := \frac{1}{k} \sin k$, $k = 1, ...$
13. $A_n := \sum_{j=0}^{n}(2j+1)$, $n = 0, ...$
14. $B_n := \sum_{i=0}^{n} \frac{i}{i+1}$, $n = 0, ...$
15. $C_n := \sum_{k=1}^{n} \frac{1}{k}$, $n = 1, ...$
16. $D_N := \sum_{i=0}^{N} \frac{2^{i+1}}{3^i}$, $N = 0, ...$
17. $E_N := \sum_{n=0}^{N} \frac{1}{n!}$, $N = 0, ...$
18. $F_k := \sum_{j=1}^{k} \left(\frac{j}{j+1} - \frac{j+1}{j}\right)$, $k = 1, ...$
19. $G_n := \sum_{k=0}^{n}(-2)^{-k}$, $n = 0, ...$
20. $X_K := \sum_{k=0}^{K} \frac{k^2}{k+1}$, $k = 0, ...$
21. $x_0 = 100$, $x_m := (1.05)x_{m-1}$, $m = 1, ...$
22. $x_0 = 1$, $x_n := \frac{x_{n-1}}{x_{n-1}+1}$, $n = 1, ...$
23. $x_0 = 2$, $x_n := x_{n-1} - \frac{1}{x_{n-1}}$, $n = 1, ...$
24. $x_0 = 1$, $x_n := x_{n-1} + \frac{1}{x_{n-1}}$, $n = 1, ...$
25. $x_0 = x_1 = 1$, $x_n = x_{n-2} + x_{n-1}$ for $n = 2, ...$ (This is called the *Fibonacci* sequence.)
26. $x_0 = 0$, $x_1 = 1$, $x_n = \frac{1}{2}(x_{n-2} + x_{n-1})$, $n = 2, ...$
27. $x_0 = 0$, $x_1 = 1$, $x_n = \frac{1}{2}(x_{n-2} - x_{n-1})$, $n = 2, ...$

2.2 Limits of Real Sequences

How do we “know” a number? Counting is good, and using this we can get any whole number. Ratios are almost as good: by dividing the space between two integers into equal parts and *then* counting, we can locate any fraction *precisely* on the number line. As we saw in § 1.1, to the earliest Greek thinkers on the subject, notably Pythagoras of Samos (*ca.* 580-500 BC), whole numbers and their ratios were the basis of reasoning about quantities—a point of view reflected in our use of the word **rational** to mean both a quality of thought and the property, for a number, of being expressible as a *ratio* of integers.

Other numbers are more slippery. For example, $\sqrt{2}$ can be defined as the positive solution to the equation

$$x^2 = 2.$$

Assuming that the equation *has* a positive solution, and only *one* (we shall consider arguments for both of these assertions later: *e.g.*, see Exercise 29 and Proposition 3.2.4) this definition specifies $\sqrt{2}$ unambiguously. But it doesn't really give us $\sqrt{2}$ as a point on the number line, at least not in a way that lets us locate it by ratios.
In fact, any attempt to put our finger on $\sqrt{2}$ by the method outlined earlier is doomed to failure, because $\sqrt{2}$ is **irrational**: it is *not* expressible as a ratio of integers.
This is depressing. Legend has it that Hippasus of Metapontum (*ca.*400 BC) a member of Pythagoras' school, was either shipwrecked or exiled, with a tombstone put up for him, as if dead, for revealing this discovery outside the group, so serious a threat was it to the current world order. We, too, might be tempted to refuse to believe it, heroically looking for two integers whose ratio *does* give $\sqrt{2}$. To avoid wasting time on this, let's settle the question once and for all. The (standard) strategy we will use to show that no such integers exist is called **proof by contradiction** (in Latin, *reductio ad absurdum*—literally, "reduction to an absurdity"). The idea is to argue that some particular phenomenon can't possibly happen, because if it *did*, then some other patently false (*i.e.*, absurd) conclusion would necessarily follow. In our case, we will argue (following Aristotle (384-322 BC)) that if we *could* write $\sqrt{2}$ as a ratio of integers, then we could use this to produce an integer which is simultaneously *even* and *odd*.

Proposition 2.2.1. *$\sqrt{2}$ is irrational: that is, there does* not *exist a pair of integers $p, q > 0$ with*

$$\left(\frac{p}{q}\right)^2 = 2.$$

Proof. (*By contradiction.*) If $\sqrt{2}$ *were* rational, we could pick an expression for it as a fraction in lowest terms. That is, if there existed *some* pair $p, q > 0$ of integers satisfying the equation above, then (after cancelling out any common factors) we could produce a pair of integers $p, q > 0$ that not only satisfy the equation above—which we rewrite as

$$p^2 = 2q^2$$

—but also with p and q *relatively prime*. In particular, we could choose p and q so that *at least one* of them is odd (why?). Let's concentrate on such a pair.
Since the right side of this last equation, $2q^2$, is even, so is the left side, p^2, and so by Exercise 30, p itself is even. This means q **must be odd**.
But p, being even, can be expressed as twice another integer (which we will call m):

$$p = 2m.$$

Squaring both sides and substituting into the original equation, we have

$$p^2 = (2m)^2 = 4(m^2) = 2q^2.$$

Dividing both sides of the last equality by 2, we have

$$2m^2 = q^2.$$

Now, an argument like the one showing p is even also shows that q **must be even**.
Whoa! We have just established that if $\sqrt{2}$ is rational, then there is an integer q which is odd and even at the same time. *That's absurd!* So $\sqrt{2}$ must be irrational.[3] □

OK, that's settled. $\sqrt{2}$ is definitely *not* rational. How, then, do we locate it on the number line? The best we can do is to devise a scheme that lets us position $\sqrt{2}$ relative to the rationals (whose location we can get *exactly*). We shall consider two schemes to do this. The first—and most familiar one—is to find a **decimal expansion** for $\sqrt{2}$. We start by noticing where $\sqrt{2}$ sits relative to the integers: since

$$1^2 < (\sqrt{2})^2 < 2^2$$

(why?) we conclude from Exercise 11 that $\sqrt{2}$ lies between 1 and 2. This means that a decimal expression for $\sqrt{2}$ would have to start with 1 before the decimal point[4]:

$$\sqrt{2} = 1.****\ldots$$

Now, if we try to append a digit in the first place after the decimal, we have ten possibilities. By trial and error, we find that

$$(1.4)^2 < (\sqrt{2})^2 < (1.5)^2$$

[3]A *conundrum* (look it up): we use contradiction to reason about the irrational!
[4](the * 's indicate as yet unknown digits)

so $\sqrt{2}$ lies between 1.4 and 1.5. We can now fill in the first digit:

$$\sqrt{2} = 1.4 * * * \dots.$$

Having decided this, we now test all the possibilities for the *next* digit and discover

$$(1.41)^2 < (\sqrt{2})^2 < (1.42)^2$$

so

$$\sqrt{2} = 1.41 * * \dots.$$

Continuing this process, we can fill in *all* the digits.

Well, yes... in principle. But it would take forever, and even if we were finished, what would we have? A string of digits. Unending: we couldn't fit it on this page, or on any other. Nonetheless, the scheme *does* (at least in principle) allow us to decide, for any particular decimal fraction φ, whether $\varphi < \sqrt{2}$ or $\varphi > \sqrt{2}$: just run the scheme until the first time that the digit produced for $\sqrt{2}$ differs from the correspoinding digit for φ. In fact, the same kind of comparison can be made with any other number for which we have a decimal expression.[5]

There is another way to look at this, as well. The various decimal fractions which we produce in this process

$$\begin{aligned} x_0 &= 1 \\ x_1 &= 1.4 \\ x_2 &= 1.41 \end{aligned}$$

and so on, form a sequence $\{x_k\}$ in $\mathbb{R}$ (where x_k has precisely k digits after the decimal point). We have chosen each element x_k of the sequence to be the highest k-digit decimal fraction which is less than $\sqrt{2}$, and so we can say that for $k = 0, 1, 2, ...$

$$x_k < \sqrt{2} < x_k + 10^{-k}$$

(why?). In particular, x_k is closer to $\sqrt{2}$ than it is to the *next* decimal fraction, so

$$\left| x_k - \sqrt{2} \right| < 10^{-k}.$$

This means, if we tried to use x_k in place of $\sqrt{2}$ in some calculation, our input would be off by an amount (called the **error**) which we don't know

[5]We conveniently gloss over the possibility of an infinite string of successive 9's, which can confuse matters, but only for pedants.

exactly, but which we can be sure is less than 10^{-k}. This is called an **error estimate** or **error bound**. Since 10^{-k} goes *down* as k increases, we actually know that the same error estimate holds for any *later* element of the sequence: for *all* $j \geq k$,

$$\left| x_j - \sqrt{2} \right| < 10^{-k}.$$

In other words, to be off by no more than 10^{-k}, we can use *any* element of the sequence in the k^{th} place or later. This is useful for practical purposes. Certainly the fact that $\sqrt{2}$ can't be measured *exactly* doesn't stop a carpenter from cutting a diagonal brace to a square frame 1 ft. on each side: if we need to specify the length of such a brace, we just run our scheme until the error estimate tells us the error is smaller than we could detect in practice (say, is less than the width of the markers on our ruler) and use that number in place of $\sqrt{2}$.
These ideas are embodied in the following definition.

Definition 2.2.2. *Given a sequence $\{x_k\}$ in $\mathbb{R}$ and numbers $y \in \mathbb{R}$, $\varepsilon > 0$,*

1. *we say the sequence* ***(eventually) approximates*** *y* ***with accuracy*** *ε if there is some value K for the index such that this and every later element of the sequence—that is, every x_k with $k \geq K$—satisfies the error estimate*

$$|x_k - y| < \varepsilon;$$

   See Figure 2.2

2. *we say the sequence* ***converges*** *to y (denoted $x_k \to y$) if it approximates y with* every *accuracy $\varepsilon > 0$.*

See Figure 2.3

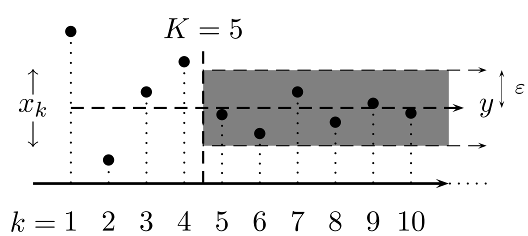


Figure 2.2: The sequence $\{x_k\}$ approximates y with accuracy ε

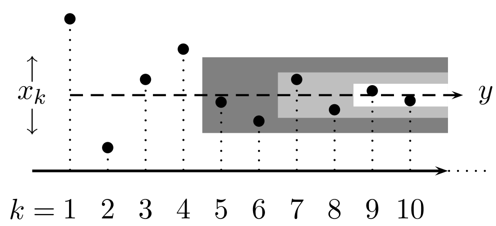


Figure 2.3: The sequence $\{x_k\}$ converges to y

To establish that a particular sequence in $\mathbb{R}$ *approximates* a particular number with a *specific* accuracy $\varepsilon > 0$, we need to name a specific value of K and then show that *it* fits the definition (that every x_k from this place on satisfies the required error estimate). But we can be fairly generous in choosing which K to name. We don't have to name the most efficient one: if some higher value for K makes it easier to establish the estimate, we *use it!* For example, to show that the decimal expansion above approximates $\sqrt{2}$ with accuracy $\varepsilon = \frac{235}{7421}$, we can happily name $K = 4$, and then just use the interval argument above, combined with the very easy observation that $10^{-4} = 1/10000 < \frac{1}{7} < \frac{235}{7421}$ to establish that whenever $k \geq 4$,

$$\left|x_k - \sqrt{2}\right| < 10^{-4} < \frac{235}{7421};$$

it isn't necessary to compute that this fraction is just a bit more than 0.3, so that $K = 2$ already works.

However, to establish that a particular sequence *converges* to a particular number, we need to be more abstract, since we need to *simultaneously* establish approximation with *every* positive accuracy. In general, we can't expect to do this with a single value of K: our choice must depend on the value of ε (presumably getting higher as ε gets smaller, requiring better accuracy), so we have to give some *scheme* for choosing it. But again, this need not be involved. To show that our decimal expressions converge to $\sqrt{2}$, we simply note that, no matter what positive number $\varepsilon > 0$ we are given, there is *some* negative power of 10 which is even smaller. So we say: Given $\varepsilon > 0$, pick K so that $10^{-k} < \varepsilon$. Then for all $k \geq K$, we have

$$\left|x_k - \sqrt{2}\right| < 10^{-4} \leq 10^{-K} < \varepsilon$$

as required.

Our relative indifference to the precise value of K in both definitions is better expressed by the use of the word "eventually": we say that some condition holds **eventually** for a sequence if there is *some* place in the sequence ("...whatever...") after which the condition can be guaranteed to hold. So our definitions could be summarized as: $\{x_k\}$ *approximates* y with given accuracy $\varepsilon > 0$ if eventually

$$|x_k - y| < \varepsilon$$

while it *converges* if *every* such estimate eventually holds.
There are several immediate consequences of these definitions. Note that for any specific accuracy $\varepsilon > 0$, the sequence approximates many different numbers with accuracy ε: for example, the decimal expansion of $\sqrt{2}$ approximates 1.419 with accuracy $\frac{1}{10}$, or even $\frac{1}{200}$. However, it does not *converge* to this number, because it fails to approximate it (even eventually) with accuracy 10^{-3} (or 10^{-100}). It can only converge, if at all, to a *single* number. This allows us to refer to "*the* limit" of a sequence.

Proposition 2.2.3. *A sequence cannot converge to two different numbers. If $\{x_k\}$ is a sequence in $\mathbb{R}$ which converges to both y and z, then $y = z$.*

Proof. The idea is that, for any k, if we have estimates on both $|x_k - y|$ and $|x_k - z|$ then we can use the triangle inequality to say something about $|y - z|$.
Given $\varepsilon > 0$, we can find K_y so that every $k \geq K_y$ satisfies

$$|x_k - y| < \varepsilon$$

and K_z so that every $k \geq K_z$ satisfies

$$|x_k - z| < \varepsilon.$$

Suppose k is chosen higher than both K_y and K_z. Then we have

$$|y - z| \leq |y - x_k| + |x_k - z| < \varepsilon + \varepsilon = 2\varepsilon.$$

But this means, for *every* $\varepsilon > 0$, we have

$$0 \leq |y - z| < 2\varepsilon.$$

Since $|y - z|$ is a set number, it must equal zero, and this means $y = z$. □

In view of this result, we are justified in formulating the following

Definition 2.2.4. *We say y is the* ***limit*** *of the sequence $\{x_k\}$ and write*[6]

$$y = \lim x_k$$

if $x_k \to x$.

A note of caution: not every sequence *has* a limit (Proposition 2.2.3 only says it can't have *more* than one, not that it has *at least* one). A real sequence is **convergent** if it converges to some number, and **divergent** otherwise (we then say that the sequence **diverges**). There are many varieties of divergence, some wilder than others. We illustrate with a few quick examples.
First, the sequence of positive integers

$$\{i\}_{i=1}^{\infty} \quad = \quad 1, 2, 3, \ldots$$

cannot converge to any number, because it gets farther and farther away from (not closer to) any candidate y we could choose for a limit: no matter what accuracy $\varepsilon > 0$ we try, the estimate

$$|i - y| < \varepsilon$$

eventually *fails*—for example, if y is any number below $1,000,000$ then for *all* $i \geq 10^7$ we have

$$|i - y| > 1$$

so no estimate involving any accuracy $0 < \varepsilon < 1$ can hold [7]. Intuitively, this sequence goes to the "right end" of the number line, which we have already referred to as the "virtual number" ∞. *This behavior is a form of divergence*, but has some regularity that makes it a kind of "virtual convergence". In the same way, the sequence of negative integers

$$\{-k\}_{k=1}^{\infty} \quad = \quad -1, -2, -3, \ldots$$

diverges by marching to the "left end" of $\mathbb{R}$, or $-\infty$. It is common to use the convergence arrow in this "virtual" way. Strictly speaking, we shouldn't, but if we're going to sin, at least let us do it with flair[8]. To be precise,

[6]Note that (unlike the notation for limits of functions, later) we do *not* use a subscript like $\lim_{k\to\infty}$ here, since there is no ambiguity about where the index of a sequence is going!

[7]Of course, this is overkill: if $i \geq 10^7$ and $y \leq 10^6$ then $|i - y|$ is a *lot* bigger than 1–but it *does* work!

[8]This kind of naughtiness is often called (with a wink) **abuse of notation**.

Definition 2.2.5. *A real sequence* $\{x_k\}$ ***diverges to (positive) infinity*** *(denoted* $x_k \to \infty$*) if for every* $M \in \mathbb{R}$ *there exists* K *such that, for every* $k \geq K$*,*

$$x_k \geq M.$$

It ***diverges to negative infinity*** *(denoted* $x_k \to -\infty$*) if for every* $m \in \mathbb{R}$ *there exists* K *such that whenever* $k \geq K$*,*

$$x_k \leq m.$$

A second example is the sequence

$$\left\{(-10)_k^k\right\}_{k=0}^{\infty} = 1, -10, 100, -1000, \ldots.$$

This does *not* diverge to ∞ or to $-\infty$, since every other element is on the wrong side of zero. But clearly the elements are getting large: we can say that the *absolute values* give a sequence diverging to ∞:

$$\left|(-10)^k\right| = 10^k \to \infty.$$

Of course, the sequence has no limit.
Intuitively, these examples suggest that when the numbers get large, they can't converge. This can be made precise by extending to sequences the idea of "boundedness" which we formulated earlier for intervals.

Definition 2.2.6. *A real sequence* $\{x_i\}$ *is* ***bounded above*** *if there exists a number* $\beta \in \mathbb{R}$ *such that*

$$x_i \leq \beta \quad \text{for all } i.$$

It is ***bounded below****if there exists a number* $\alpha \in \mathbb{R}$ *such that*

$$\alpha \leq x_i \quad \text{for all } i.$$

It is ***bounded*** *if it is BOTH bounded above AND bounded below, or equivalently, if there exists a number* $\gamma \in \mathbb{R}$ *such that*

$$|x_i| \leq \gamma \quad \text{for all } i.$$

Lemma 2.2.7. *If a sequence is eventually bounded, then it is bounded: that is, to show that a sequence is bounded, we need only find a number* $\gamma \in \mathbb{R}$ *such that the inequality*

$$|x_i| \leq \gamma$$

holds for all $i \geq K$*, for some* K*.*

The proof of this is straightforward, and is left as an exercise (Exercise 32).

Proposition 2.2.8. *Every convergent sequence is bounded.*

Proof. Suppose $x_i \to x$. Pick any $\varepsilon > 0$, and find K so that for all $i \geq K$

$$|x_i - x| < \varepsilon.$$

Now, let

$$\gamma = \varepsilon + |x| .$$

Then for all $i \geq K$,

$$|x_i| \leq |x_i - x| + |x| \leq \varepsilon + |x| = \gamma$$

so that $\{x_i\}$ is *eventually* bounded, and hence bounded, by Lemma 2.2.7. □

Knowledge of an explicit bound on a convergent sequence actually gives us some information about the limit.

Lemma 2.2.9. *Suppose $x_k \to x$ is a convergent sequence for which we can give explicit bounds: for all k,*

$$\alpha \leq x_k \leq \beta.$$

Then these are also bounds on the limit:

$$\alpha \leq x \leq \beta.$$

Proof. (*By contradiction:*) We will show that the inequality $x < \alpha$ is impossible; a similar argument works for the other bound. Suppose $x_k \to x$, $x_k \geq \alpha$ for all k, and $x < \alpha$. Then since

$$\varepsilon = \alpha - x > 0$$

we can find N so that $k \geq N$ guarantees

$$|x_k - x| < \varepsilon.$$

But this implies that

$$\begin{aligned} x_k &= x + (x_k - x) \\ &\leq x + |x_k - x| \\ &< x + \varepsilon \\ &= \alpha \end{aligned}$$

which contradicts the fact that $x_k \geq \alpha$. □

This lemma can be stated in more geometric language:

> *If all elements of a convergent sequence $\{x_k\}$ belong to the closed interval $[\alpha, \beta]$, then so does the limit of the sequence.*

Note that the same is *not* necessarily true of an open (or half-open) interval (See Exercise 31) Thus, even if we have a *strong* inequality involving the *terms* of a sequence, we can only conclude (in general) a *weak* inequality for the *limit*.

How do we tell whether a sequence is bounded, and if so, how do we find a bound? Sometimes a bound is easy to find: for example, consider the sequence defined recursively by

$$x_0 = \frac{1}{3}$$
$$x_{k+1} = \frac{1}{2}(x_k + 1)$$

which we can rewrite as

$$x_k = \frac{1}{2} + \frac{x_k}{2}.$$

It seems pretty clear that 0 is a lower bound, since we deal only with positive numbers, addition, and division by 2:

$$0 \leq x_k \text{ for all } k.$$

But what about an upper bound? If we look at the first few terms

$$\begin{aligned}
x_0 &= \frac{1}{3} \\
x_1 &= \frac{1}{2}\left(\frac{1}{3} + 1\right) \\
&= \frac{2}{3} \\
x_2 &= \frac{1}{2}\left(\frac{2}{3} + 1\right) \\
&= \frac{5}{6} \\
x_2 &= \frac{1}{2}\left(\frac{5}{6} + 1\right) \\
&= \frac{11}{12} \\
&\vdots
\end{aligned}$$

we might suspect that 1 is an upper bound. How can we confirm this?
A technique of proof which is often useful in these kinds of circumstances (particularly given a recursive definition) is a kind of "bootstrap" argument.
We know that the first term is less than 1:

$$x_0 = \frac{1}{3} \leq 1.$$

Now, suppose we knew that at stage k we still had a term less than 1; could we conclude that the *next* term is *also* less than 1? Well, let us do some comparisons. If

$$x_k \leq 1$$

then

$$\begin{aligned} x_{k+1} &= \frac{1}{2} + \frac{x_k}{2} \\ &\leq \frac{1}{2} + \frac{1}{2} \\ &= 1 \end{aligned}$$

so we have established the abstract statement, that *if* $x_k \leq 1$ *then also* $x_{k+1} \leq 1$.
So...? Well, think about it: we know that whenever one of the terms of our sequence is less than 1, the *next* one is, as well. We *also* know that the *first* term is less than 1. But this says they are *all* less than 1: if we wanted to show that $x_{100} \leq 1$, we would argue that since $x_0 \leq 1$, we know that $x_1 \leq 1$; but then since $x_1 \leq 1$, we know that $x_2 \leq 1$; and then since $x_2 \leq 1$, we know that $x_3 \leq 1$;...and so on—after a hundred such mini-arguments, we would arrive at...since $x_{99} \leq 1$, we know that $x_{100} \leq 1$—*so there*!
This reasoning is formalized in **proof by induction**, a technique for proving statements about all natural numbers. It consists of two stages. First, we establish that our statement is true in the **initial case**: in the argument above, we showed (essentially by inspection) that

$$x_0 \leq 1.$$

Then we prove an abstract assertion: that *if* our statement holds for some *particular* (but unspecified) number, *then* it *also* holds for the *next* number. This is known as the **induction step** (and the "if" part is called the **induction hypothesis**). In our argument above, the induction hypothesis was that $x_k \leq 1$, and the induction step was the assertion that in this case also $x_{k+1} \leq 1$.

Sitting in the background, then, is the recognition that these two arguments together tell us that our statement holds for *all* natural numbers, since it is clear that for any *particular* number we could iterate the argument in the induction step, starting from the initial case, going number-by-number until we reached our conclusion for the particular number in question. Hence, it is accepted that we don't need to actually carry out this iterated argument: once we have established the initial case and the induction step, we have proved our statement for all natural numbers.

For a further discussion of proof by induction, and some more examples, see Appendix A.

Note, by the way, that our first observation about the sequence above—that all the terms are positive because we are only adding and dividing by 2, and we start from a positive number—is really an induction argument, as well.

While every convergent sequence must be bounded, it is *not* true that every bounded sequence converges: in other words, there exist *bounded* sequences which *diverge*. For example, consider the sequence

$$\left\{(-1)_k^k\right\}_{k=0}^{\infty} = +1, -1, +1, -1, +1, -1, \ldots$$

which alternates the two values $+1$ and -1. This is *bounded*, since the absolute value of all elements is 1. Nonetheless the sequence *diverges*. To see this, consider the possibilities. It can't converge to $+1$, because every second element (*i.e.*, when k is odd) satisfies

$$\left|(-1)^k - 1\right| = 2 \quad (k \text{ odd})$$

so it cannot approximate $+1$ with accuracy, say, $\varepsilon = 1$. However, if $y \neq 1$, we see that every *even* index term has

$$\left|(-1)^k - y\right| = |y - 1| > 0 \quad (k \text{ even})$$

and so we can't approximate y with accuracy, say, half the distance from y to 1: $\varepsilon = |y - 1| / 2$.

This particular sequence repeats a regular pattern: we call it a **periodic** sequence. The argument here can be adapted to show that a *periodic*

sequence which is not *constant*[9] is divergent. However, there are many sequences which diverge without so regular a pattern. We shall see some examples later.
So in some sense, divergence to infinity is more "like" convergence than other kinds of divergence: the only formal difference is that instead of taking $\varepsilon > 0$ and requiring $|x_k - y| < \varepsilon$ in the case of convergence to y, we take any real number M (or m) and require $x_k > M$ (*resp.* $x_k < m$) for divergence to ∞ (*resp.* $-\infty$). For convergence, "better accuracy" means ε *small*; for divergence to ∞, it means M *large positive* and for divergence to $-\infty$ it means m *large negative*.

Exercises for § 2.2

Answers to Exercises 1-27 (odd only), 28, 31 are given in Appendix B.

Practice problems:

Problems 1-27 are numbered according to the list of sequences on page 19. For each sequence, try to decide whether it is bounded above (and if so, try to give an explicit upper bound) and whether it is bounded below (giving an explicit lower bound if possible). You should be able to justify your assertions. Note: you are *not* expected to give the *best* bound if it exists, only *some* bound that works.
Special Note: For sequences 15-18, 23-24, and 27, the bounds may be harder to find or verify than the others, given your present set of tools; for each of these, see if you can guess whether it is bounded above (*resp.* below) and give a tentative bound if you guess it exists. Try to verify your guess rigorously (*i.e.*, to prove your answer)—but you may find you can't. The attempt to grapple with it, however, should be instructive.

Theory Problems:

28. For each of the following, either give an example or show that no example exists.

    (a) A sequence that diverges to $+\infty$.

    (b) An unbounded sequence that does not diverge to $+\infty$ and does not diverge to $-\infty$.

    (c) An unbounded convergent sequence.

    (d) A bounded divergent sequence.

[9] A **constant sequence** has all elements the same.

(e) A sequence that converges to both -1 and $+1$.

(f) An unbounded sequence for which 3 is an upper bound, but no $\beta < 3$ is an upper bound.

(g) A sequence for which 1 is an upper bound, and no $\beta < 1$ is an upper bound, but for which no term equals 1.

29. Show that there cannot be two different *positive* solutions of the equation

$$x^2 = 2.$$

(*Hint:* Use Exercise 10 in § 1.2.)

30. (a) Show that the square of an *even* integer is *even*.

(b) Show that the square of an *odd* integer is *odd*. (*Hint:* A number q is odd if it can be written in the form $q = 2n + 1$ for some integer n. Show that then q^2 can also be written in this form.)

(c) Show that if m is an integer with m^2 odd (*resp.* even), then m is odd (*resp.* even).

31. Find an example of a convergent sequence $\{x_k\}$ contained in the open interval $(0, 1)$ whose limit is *not* contained in $(0, 1)$. (*Hint:* Look at small fractions.)

32. Prove Lemma 2.2.7.

33. (a) Using the proof of Proposition 2.2.1 as a guide, show that $\sqrt{3}$ is irrational.

(b) Can you extend this to a proof that $\sqrt{p}$ is irrational for every prime number $p \geq 2$?

Challenge problem:

34. (J. Propp) Consider the sequence defined recursively for x_i when $i \geq 2$ by

$$x_i = \frac{x_{i-1} + 1}{x_{i-2}}.$$

(a) Pick several different pairs of positive values for x_0 and x_1, and evaluate the next five terms.

(b) Can you make any general statement about how you think the sequence will behave? Will it converge?

(c) Can you prove your statement?

2.3 Convergence to Unknown Limits

We have encountered some sequences which *seemed* to converge, but we couldn't *prove* convergence because we couldn't get our hands on an explicit limit. In this section, we shall develeop a few techniques for establishing convergence of a sequence without needing to know its limit ahead of time.

These techniques are based on our intuition that there are no "holes" in the number line; this property is sometimes referred to as **completeness** of the number line. Without a more detailed construction for the line, we cannot hope to *prove* completeness. Rather, we shall treat it as an **axiom**; an assertion we accept as an article of faith, from which we can then hope to deduce further properties.

To formulate this axiom, we need some new terminology.

Definition 2.3.1. *A real sequence* $\{x_k\}$ *is*

- ***(strictly) increasing*** *(denoted* $x_k\uparrow$*) if* $x_k < x_{k+1}$ *for every* k*;*
- ***(strictly) decreasing*** *(denoted* $x_k\downarrow$*) if* $x_k > x_{k+1}$ *for every* k*.*

To allow occasional equality between successive terms, we use negative terminology: the sequence is

- ***non-decreasing*** *(*$x_k \uparrow$*) if* $x_k \leq x_{k+1}$ *for every* k*;*
- ***non-increasing*** *(*$x_k \downarrow$*) if* $x_k \geq x_{k+1}$ *for every* k*.*

A sequence which fits one of these descriptions is ***monotone****.*

Monotonicity requires that the *same* inequality holds between *every* pair of successive elements in the sequence. The sequences

$$\{k_k\}_{k=1}^{\infty}, \quad \left\{\frac{1}{k}_k\right\}_{k=1}^{\infty}, \quad \{-k^2_k\}_{k=0}^{\infty}, \quad \left\{\frac{k-1}{k}_k\right\}_{k=2}^{\infty}$$

are all monotone. Both $\{k\}$ and $\{\frac{k-1}{k} = 1 - \frac{1}{k}\}$ are strictly increasing (and hence non-decreasing) while $\{\frac{1}{k}\}$ and $\{-k^2\}$ are strictly decreasing (and hence non-increasing). But the sequences

$$\left\{\frac{(-1)^k}{k}_k\right\}_{k=1}^{\infty}, \quad \{k^3 - 300k_k\}_{k=1}^{\infty}, \quad \{\cos k_k\}_{k=0}^{\infty}$$

are *not* monotone. The first

$$x_k = \frac{(-1)^k}{k}$$

alternates positive and negative values, so that $x_k < x_{k+1}$ for k odd while $x_k > x_{k+1}$ for k even. The second sequence

$$x_k = k(k^2 - 300)$$

starts out decreasing:

$$x_1 = -299,\ x_2 = -592,\ x_3 = -873, \ldots$$

but for example by the time $k = 20$, both factors in x_k are positive and increasing, so the sequence is **eventually increasing** . The third sequence

$$x_k = \cos k$$

is a bit harder to predict in the long run: but certainly for $0 \leq k < \pi$ it decreases, while for $\pi < k < 2\pi$ it increases, so already it is not monotone. We have seen that in general a sequence can be unbounded without necessarily diverging to infinity, or bounded without necessarily converging. However, a *monotone* sequence behaves much more predictably. The behavior of a *bounded, monotone* sequence is the subject of our axiom.

Axiom 2.3.2 (Completeness Axiom)**.** *Every bounded monotone sequence is convergent.*

Note that a monotone sequence is automatically bounded on one side by its initial term: if $\{x_k\}$ is non-decreasing (*resp.* non-increasing), then for all k, $x_0 \leq x_k$ (*resp.* $x_0 \geq x_k$). Thus, to show a *non-decreasing* sequence is bounded, we need only find an *upper* bound for it; for a *non-increasing* sequence, we need only a *lower* bound. Combined with the completeness axiom, this yields a very simple picture of the behavior of those sequences which we know to be monotone.

Proposition 2.3.3. *Suppose $\{x_k\}$ is a monotone sequence.*

1. *If $x_k \uparrow$, then either $\{x_k\}$ is bounded above, and hence converges, or $x_k \to \infty$.*

2. *If $x_k \downarrow$, then either $\{x_k\}$ converges or $x_k \to -\infty$.*

Proof. Suppose $x_k \uparrow$. If β is an upper bound for all x_k, then we have

$$\alpha = x_0 \leq x_k \leq \beta$$

for all k, and hence the sequence converges by the completeness axiom.
If x_k has no upper bound, then for each $M \in \mathbb{R}$ there exists N so that

$$x_N > M.$$

However, since $x_k \uparrow$, we then have, for every $k \geq N$,

$$x_k \geq x_N > M.$$

Thus, for every $M \in \mathbb{R}$, we have found N so that $k \geq N$ guarantees

$$x_k > M,$$

which is the definition of divergence to $+\infty$ ($x_k \to \infty$).
The proof of the second claim is analogous, and is left to you (Exercise 27). □

To understand the completeness axiom in more concrete terms, we examine its implications for decimals. A *finite* string of digits after the decimal denotes a fraction with denominator a power of 10. But an *infinite* string of digits

$$0.d_1 d_2 d_3 \ldots$$

(where each digit is one of the ten integers $\{0, 1, \ldots, 9\}$) must be interpreted as the limit of the sequence of finite decimals

$$\begin{aligned} x_0 &= 0 \\ x_1 &= 0.d_1 \\ x_2 &= 0.d_1 d_2 \\ &\vdots \end{aligned}$$

It is clear that this sequence of fractions is non-decreasing (*right?*), and that for all k,

$$x_0 \leq x_k < x_0 + 1 = 1.0.$$

The Completeness Axiom now guarantees that this sequence of fractions converges. In other words, the number

$$x = 0.d_1 d_2 d_3 \ldots$$

which is supposed to be represented by our infinite decimal actually exists, as the honest limit of the sequence of truncations

$$x = \lim x_k$$

where $x_k = 0.d_1 d_2 \ldots d_k$. We state this formally.

Remark 2.3.4. *Every infinite decimal expression represents a real number, given as the limit of the sequence of finite decimals obtained by truncating the expression.*

How do we tell, in practice, if a sequence is monotone? There are two useful interpretations of the inequality

$$x_k < x_{k+1}.$$

The first is that the **difference** is *positive*:

$$x_{k+1} - x_k > 0.$$

The second applies when *both* elements are *positive*, namely that the **ratio** is greater than 1:

$$\frac{x_{k+1}}{x_k} > 1.$$

These can be used as tricks for establishing (or disproving) monotonicity. As an example, consider the sequence

$$x_k = \frac{k^2}{k+1} \quad k = 1, 2, 3, \ldots.$$

If we examine the difference between x_{k+1} and x_k, we find

$$\begin{aligned} x_{k+1} - x_k &= \frac{(k+1)^2}{(k+1)+1} - \frac{k^2}{k+1} \\ &= \frac{(k+1)^3 - k^2(k+2)}{(k+2)(k+1)} \\ &= \frac{(k^3 + 3k^2 + 3k + 1) - (k^3 + 2k^2)}{(k+2)(k+1)} \\ &= \frac{k^2 + 3k + 1}{(k+2)(k+1)}. \end{aligned}$$

It is clear that for $k \geq 1$ this is a positive number, so

$$x_{k+1} - x_k > 0,$$

or

$$\frac{k^2}{k+1}\uparrow.$$

As another example, consider

$$x_k = \frac{8^k}{k^3} \quad k = 1, 2, ...$$

The ratio x_{k+1}/x_k can be computed as

$$\frac{x_{k+1}}{x_k} = \frac{8^{k+1}}{(k+1)^3} \cdot \frac{k^3}{8^k} = 8\left(\frac{k}{k+1}\right)^3.$$

Now, this is at least 1 precisely if $(k/(k+1))^3$ is at least $1/8$, or equivalently $k/(k+1)$ is at least $1/2$. But if $k \geq 1$, we can write

$$k + 1 \leq k + k = 2k$$

so that

$$\frac{k}{k+1} \geq \frac{k}{2k} = \frac{1}{2}.$$

Thus

$$\frac{8^k}{k^3}\uparrow.$$

An infinite decimal expression is an example of a *series with non-negative terms*. Recall from 18 that using a given sequence of numbers $\{a_i\}_{i=0}^{\infty}$, we can form a new sequence $\{S_k\}_{k=0}^{\infty}$, defined recursively by the rule

$$S_0 = a_0, \quad S_k = S_{k-1} + a_k,\ k = 1, ...$$

This new sequence can also be written using summation notation as

$$S_k = \sum_{i=0}^{k} a_i.$$

This kind of sequence is formally an infinite sum; we refer to it as a **series** whose **terms** are $a_0, a_1, ...$; the finite sums S_k are the **partial sums** of the series. The series **converges** if the sequence of partial sums S_k converges to a limit $S \in \mathbb{R}$, in which case we call S the **sum** of the series:

$$S = \sum_{i=0}^{\infty} := \lim S_k = \lim_k \left(\sum_{i=0}^{k} a_i\right).$$

If the partial sums form a divergent sequence, we say the infinite sum **diverges**.
For example, the infinite decimal expression $d_0.d_1d_2\cdots$ can be rewritten as the infinite sum

$$d_0.d_1d_2\cdots = \sum_{i=0}^{\infty} \frac{d_i}{10^i}.$$

In Remark 2.3.4 we saw that the partial sums of this series (which are the *finite* decimal expressions using these digits) are non-decreasing, so to prove convergence we needed only to establish an upper bound, which in this case was easy.
The "difference" trick can be applied to show that *any* series whose terms are all non-negative gives rise to a non-decreasing sequence of partial sums:

Lemma 2.3.5. *For any series $\sum_{i=0}^{\infty} a_i$ with non-negative terms*

$$a_i \geq 0, \quad i = 1, 2, ...,$$

the partial sums

$$S_k = \sum_{i=0}^{k} a_i, \quad k = 1, 2, ...$$

form a non-decreasing sequence; hence such a sum either converges or diverges to $+\infty$.

Proof. $S_{k+1} - S_k = a_{k+1} \geq 0$, so $S_k \uparrow$. □

A useful corollary of this is

Corollary 2.3.6 (Comparison Test for Positive Series)**.** *If two sequences $\{a_i\}_{i=0}^{\infty}$ and $\{b_i\}_{i=0}^{\infty}$ satisfy $0 \leq a_i \leq b_i$ for all i and the series $\sum_{i=0}^{\infty} b_i$ converges, then the series $\sum_{i=0}^{\infty} a_i$ also converges.*

Proof. Denote the partial sums of these series by

$$A_n := \sum_{i=0}^{n} a_i, \quad B_n := \sum_{i=0}^{n} b_i.$$

We have $\{B_n\}_{n=0}^{\infty}$ nondecreasing with limit B; in particular, each partial sum is at most B. But then $A_n \leq B_n \leq B$; thus $\{A_n\}_{n=0}^{\infty}$ is a bounded, nondecreasing sequence, hence it is convergent, as required. □

This observation is often used to formulate a notation for convergence of positive series: for any series $\sum_{i=0}^{\infty} a_i$ with $a_i \geq 0$, we write

$$\sum_{i=0}^{\infty} a_i = \infty$$

if it diverges (since in that case $S_k = \sum_{i=0}^{k} a_i \to \infty$), and

$$\sum_{i=0}^{\infty} a_i < \infty$$

if it converges. In this notation, Corollary 2.3.6 looks like the obvious statement that if $a_i \leq b_i$ for all i and $\sum_{i=0}^{\infty} b_i < \infty$, then also $\sum_{i=0}^{\infty} a_i < \infty$. The Completeness Axiom (Axiom 2.3.2) requires two conditions to guarantee convergence: the sequence must be *bounded*, and it must be *monotone*. It is clear that an *unbounded* sequence cannot converge, whether or not it is monotone. However, we have some intuitive feeling that a *bounded* sequence should come close to converging, at least in some sense. To clarify this, consider an example. Recall that the sequence

$$x_k = (-1)^k$$

is bounded but divergent, because *even* terms are near $+1$ while *odd* terms are near -1. Technically, the sequence diverges, but we *want* to say that it "goes to" *both* -1 and $+1$. To formulate this urge better, note that if we throw away the odd numbered terms, we get a new sequence which converges to $+1$, whereas if we instead throw away the *even* terms, we get a sequence converging to -1. A sequence obtained from $\{x_k\}$ by "throwing away terms" is called a "subsequence". Suppose we pick an *infinite* subset $\mathcal{K} \subset \mathbb{N}$; then we can write the elements of $\mathcal{K}$ in order (that is, as a strictly increasing sequence of numbers):

$$\mathcal{K} = \{k_1, k_2, ...\}.$$

Then we can define a new sequence by

$$y_i = x_{k_i}, \quad i = 1, 2, ...$$

This new sequence $\{y_i\}_{i=1}^{\infty}$ consists of selected terms of the original sequence $\{x_k\}_{k=1}^{\infty}$ *in the same order* as they appear in the original sequence; we call $\{y_i\}_{i=1}^{\infty}$ a **subsequence** of $\{x_k\}_{k=1}^{\infty}$. Thus, a subsequence

of $\{x_k\}_{k=1}^{\infty}$ is a sequence $\{y_i\}_{i=1}^{\infty}$ such that all the y_i appear, in the same order, as (not necessarily consecutive) elements of $\{x_k\}$.
The following lemma shows that even if a sequence is not monotone, it *does* have a monotone *subsequence*.

Lemma 2.3.7. *Every sequence in* $\mathbb{R}$ *has a monotone subsequence.*

Proof. Call an index k a **peak point**[10] of the sequence $\{x_k\}$ if for every $k' > k$ we have $x_{k'} < x_k$ (see Figure 2.4), and consider two cases:

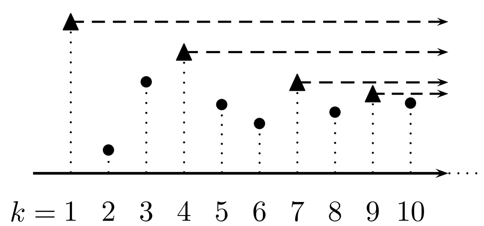


Figure 2.4: Peak points

Case 1: If $\{x_k\}$ has *infinitely many* peak points, let k_1 be some peak point; since there are infinitely many peak points, pick k_2 a later peak point, k_3 an even later one, and so on. We clearly must have

$$x_{k_1} > x_{k_2} > x_{k_3} > ...$$

in other words, we have found a *strictly decreasing* subsequence.

Case 2: If $\{x_k\}$ has only *finitely many* peak points, let k_1 be an index higher than all of them; since it is *not* a peak point, we can find some later index $k_2 > k_1$ for which $x_{k_2} \geq x_{k_1}$; but this is *also* not a peak point, and we can find $k_3 > k_2$ with $x_{k_3} \geq x_{k_2}$, and so on. Then we clearly have

$$x_{k_1} \leq x_{k_2} \leq x_{k_3} \leq ...$$

in other words, we have found a *nondecreasing* subsequence.

In either case, we have produced a monotone subsequence, as required. □

[10] This elegant proof comes from [50, p. 425].

An important corollary of Lemma 2.3.7 is the following result, which extends the applicability of the Completeness Axiom, showing that the phenomenon we observed in the example above is typical for bounded sequences.

Proposition 2.3.8 (Bolzano-Weierstrass Theorem)**.** *Every bounded sequence of numbers has a convergent subsequence.*

This theorem was first formulated and proved by Bernhard Bolzano (1781-1848) in 1817, but his work was not widely known. In the 1860's, Karl Theodor Wilhelm Weierstrass (1815-1897) resurrected some of Bolzano's ideas (with due credit), and demonstrated the power of this result.

Proof. Given a bounded sequence, say $\{x_k\}_{k=1}^{\infty}$, by Lemma 2.3.7, there exists a monotone subsequence $y_i := x_{k_i}$ ($k_{i+1} > k_i$ for all i). Now, the original sequence is bounded, say $\alpha < x_k < \beta$ for all k; since this holds in particular for $k = k_1, k_2, ...$, we have $\alpha < y_i < \beta$ for all i. Thus the subsequence $\{y_i\}$ is monotone and bounded, and therefore convergent by Axiom 2.3.2.

□

There is a name for the kind of point we have found in the preceding result.

Definition 2.3.9. *For any sequence $\{x_k\}$ in $\mathbb{R}$, an* ***accumulation point*** *of the sequence is any point which is the limit of some subsequence $\{y_j\}$ of the sequence $\{x_k\}$.*

Thus, the Bolzano-Weierstrass theorem says that *every bounded sequence has at least one accumulation point.* We note that:

1. if a sequence $\{x_k\}$ has two or more distinct accumulation points, it diverges (Exercise 28);

2. if it has *no* accumulation points, then $|x_k| \to \infty$ (Exercise 29);

3. a *bounded* sequence converges if it has *precisely* one accumulation point, which is then the limit of the sequence (Exercise 30).

We close with an important observation about series. This was stated explicitly by Augustin-Louis Cauchy (1789-1857) in his lectures at the Ecole Polytechnique, published in 1821 [51, pp. 1-3].

Proposition 2.3.10 (Divergence Test for Series)**.** *If the series $\sum_{k=0}^{\infty} a_k$ converges, then the sequence $\{a_k\}$ must converge to zero: $a_k \to 0$.*

Proof. We prove the *contrapositive* statement[11]: If the *terms* a_k do *not* converge to 0, then the *series* $\sum_{k=0}^{\infty} a_k$ *cannot* converge.
To say that $a_k \not\to 0$ means that for some accuracy $\varepsilon > 0$, the sequence $\{a_k\}$ does not eventually approximate 0 with accuracy ε—in other words, for every K there exists $k > K$ for which $|a_k| = |a_k - 0| \geq \varepsilon$.
But for the *series* to converge to anything, say $A = \sum_{k=0}^{\infty} a_k$, we must have that the sequence of partial sums $A_n := \sum_{k=0}^{n} a_k$ converges to A. This means in particular that the sequence $\{A_n\}$ eventually approximates A with accuracy $\frac{\varepsilon}{2}$—that is, for some N, $n > N$ guarantees that

$$|A_n - A| < \frac{\varepsilon}{2}.$$

Now, using N as K in the characterization of the fact that $a_k \not\to 0$, find $n > N$ for which

$$|a_{n+1}| \geq \varepsilon;$$

using the fact that $a_{n+1} = A_{n+1} - A_n$, we then have

$$\varepsilon \leq |a_{n+1}| = |A_{n+1} - A_n| \leq |A_{n+1} - A| + |A_n - A|$$

but since both n and $n+1$ are greater than N, both terms in this last expression are strictly less than $\frac{\varepsilon}{2}$, so their sum is strictly less than ε, a contradiction. □

Be careful about what this theorem does and does *not* say. You can *only* use it to show that a series *diverges* (if the terms don't go to zero). To show that a series *converges* (given that the terms *do* go to zero), you need to use other methods. See Exercise 34.

Exercises for § 2.3

*Answers to Exercises 1-21 (odd only), 25, 33*a *are given in Appendix B.*

Practice problems:

Problems 1-26 are numbered according to the list of sequences on page 19; 23 is omitted (the questions below are difficult to answer for this sequence, at least at this stage).

[11]The *contrapositive* of a statement in the form "if A is true then B is true" ($A \Rightarrow B$) is the statement "if B is false then A is false" (not $B \Rightarrow$ not A). For example, the contrapositive of "if a number equals its square, then it is zero or one" is "if a number is not zero or one, then it does not equal its square". Every statement is logically equivalent to its contrapositive: either both are true, or both are false. See Appendix A, p. 569.

(a) For each of sequences 1-22 and 24-26, decide whether it is

- nondecreasing
- nonincreasing
- eventually nondecreasing
- eventually nonincreasing
- not eventually monotone

(b) For each of sequences 1-15, 20-22, and 25-26, try to decide whether it is convergent (you don't need to find the limit, even if it exists), and if not, try to either find a convergent subsequence or explain why none exists.

Theory Problems:

27. Prove the second part of Proposition 2.3.3

28. Show that a sequence with two distinct accumulation points must diverge. (*Hint:* Look at the proof of divergence for $\{(-1)^k\}$.)

29. (a) Show that if $x_k \to \infty$, then $\{x_k\}$ has *no* accumulation points. (*Hint:* Look at the proof of divergence for $\{k\}$.)

    (b) Show that $\{(-1)^k k\}$ has no accumulation points, even though it does *not* diverge to ∞.

    (c) Show that if $\{x_k\}$ has no accumulation points, then $|x_k| \to \infty$. (*Hint:* If $|x_k|$ does *not* diverge to infinity, find a bounded subsequence, and apply the Bolzano-Weierstrass theorem to it.)

30. (a) Show that if $x_k \to x$ then x is the one and only accumulation point of $\{x_k\}$.

    (b) Find an unbounded sequence with 0 as its *only* accumulation point. (*Hint:* Think about Exercise 29.)

    (c) Show that if $\{x_k\}$ is a *bounded* sequence with x as its one and only accumulation point, then *every* subsequence of $\{x_k\}$ must converge to x.

    (d) Show that if $y \neq \lim x_k$, then *some* subsequence of $\{x_k\}$ does *not* have y as an accumulation point.

    (e) Use the previous two parts of this exercise to show that if $\{x_k\}$ is bounded and x is its one and only accumulation point, then $x_k \to x$.

31. Let $\{x_k\}_{k=1}^{\infty}$ be a sequence of real numbers, and let $S_n := \sum_{k=1}^{n} x_k$ be the sequence of partial sums for the series $\sum_{k=1}^{\infty} x_k$. In this and the next problem, we prove more detailed versions of part of the Divergence Test for series (Proposition 2.3.10).

    (a) Show that if the sequence $\{x_k\}$ is *constant* and *strictly positive* (*i.e.*, there is some $\alpha > 0$ such that $x_k = \alpha$ for all k) then $S_n \to \infty$.

    (b) Show that if the sequence is (not necessarily constant, but) bounded below by some strictly positive number ($x_k \geq \alpha > 0$ for some α), then $S_n \to \infty$.

    (c) Modify your proof above to show that if $x_k \geq \alpha$ *eventually*, then we still have $S_n \to \infty$.

    (d) Use this to show that if the *sequence* $\{x_k\}$ converges to a strictly positive limit $\lim x_k = L > 0$, then $S_n \to \infty$. (*Hint:* What can you say if $|x_k - L| < \frac{L}{2}$?)

32. As in the previous problem, suppose $\{x_k\}_{k=1}^{\infty}$ is a sequence of real numbers and $S_n := \sum_{k=1}^{n} x_k$ are the partial sums of the series $\sum_{k=1}^{\infty} x_k$.

    (a) Show that if $\lim x_k = L < 0$, then S_n cannot converge. (*Hint:* Show that for n large, $|S_n - S_{n+1}| > \frac{|L|}{3}$, and use this to show that the two subsequences $\{S_{2n}\}$ and $\{S_{2n+1}\}$ cannot both converge to the same limit.)

    (b) Modify your proof in (a) to show that if the sequence of absolute values converges to a nonzero value, say $|x_k| \to L > 0$ (even if the terms x_k themselves don't converge) then the sequence S_n of partial sums diverges.

    (c) Does the Divergence Test (Proposition 2.3.10) say that *every* series $\sum_{k=1}^{\infty} x_k$ with $x_k \to 0$ converges?

Challenge problems:

33. Consider the sequence defined recursively by

$$x_1 = \frac{1}{2}, \quad x_{k+1} = x_k + \frac{1}{2}(1 - x_k).$$

    (a) Write down the first three terms of this sequence.

(b) Show by induction[12] that for all $k > 1$,

$$\frac{1}{2} < x_k < 1.$$

(c) Show that this sequence is strictly increasing. (*Hint:* Try a few terms, maybe rewrite the formula, or think geometrically!)

(d) Show that this sequence converges.

(e) You can probably guess what the limit appears to be. Can you *prove* that your guess is correct?

34. Show that the harmonic series $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges to infinity, as follows (a version of this proof was given by Nicole Oresme (1323-1382) in 1350 [20, p. 91] and rediscovered by Jacob Bernoulli (1654-1705) in 1689 [51, pp. 320-4]): let $S_K := \sum_{k=0}^{K} \frac{1}{k}$ be the partial sums of the harmonic series.

(a) Show that for $n = 0, ...,$ the sum $\sigma_n := \sum_{k=2^n}^{2^{n+1}-1} \frac{1}{k}$ is bounded below by $\frac{2^n}{2^{n+1}} = \frac{1}{2}$.

(b) Use Exercise 31b to conclude that the series $\sum_{n=0}^{\infty} \sigma_n$ diverges.

(c) Show that $\sum_{n=0}^{N-1} \sigma_n = S_{2^N-1}$ for $N = 1, ...,$ so that the subsequence S_{2^N} of $\{S_K\}$ diverges to infinity.

(d) Conclude that

$$\sum_{k=1}^{\infty} \frac{1}{k} = \infty.$$

2.4 Finding limits

It is natural to expect that if we perform arithmetic operations on convergent sequences, these should carry over to the limit. We shall verify that this is true and see ways in which it can be used to determine certain limits.

Theorem 2.4.1 (Arithmetic and Limits)**.** *Suppose $x_k \to x$ and $y_k \to y$ in $\mathbb{R}$. Then*

1. $x_k + y_k \to x + y$
2. $x_k - y_k \to x - y$

[12] see Appendix A

3. $x_k y_k \to xy$

4. If $y \neq 0$, $\frac{x_k}{y_k} \to \frac{x}{y}$

5. If $x > 0$, then for every positive integer n, $\sqrt[n]{x_k} \to \sqrt[n]{x}$.

Proof. Proof of (1):
Given $\varepsilon > 0$, we need to find N so that $k \geq N$ guarantees

$$|(x_k + y_k) - (x + y)| < \varepsilon.$$

Let us analyze the quantity we wish to estimate: rearrangement, and Proposition 1.2.1(1) give us

$$|(x_k + y_k) - (x + y)| = |(x_k - x) + (y_k - y)| \leq |x_k - x| + |y_k - y|.$$

One way to make sure that our desired estimate holds is by making each of the terms in this last expression less than $\varepsilon/2$. But we can do this: since $x_k \to x$, we can find N_1 so that $k \geq N_1$ guarantees

$$|x_k - x| < \frac{\varepsilon}{2}$$

and since $y_k \to y$, we can also find N_2 so that $k \geq N_2$ guarantees

$$|y_k - y| < \frac{\varepsilon}{2}.$$

Now, let

$$N \geq \max\{N_1, N_2\}.$$

Then $k \geq N$ guarantees

$$|(x_k + y_k) - (x + y)| \leq |x_k - x| + |y_k - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

as required. ♢

Proof of (2):
This is just like the preceding proof; we leave it as an exercise (Exercise 25). ♢

Proof of (3):
Again, given $\varepsilon > 0$ we need to find N so that $k \geq N$ guarantees

$$|x_k y_k - xy| < \varepsilon.$$

This doesn't break up as nicely as the expression in the first part, but a trick helps us considerably: we subtract and add the quantity xy_k in the middle:

$$\begin{aligned} |x_k y_k - xy| &= |x_k y_k - xy_k + xy_k - xy| \\ &\leq |x_k y_k - xy_k| + |xy_k - xy| . \end{aligned}$$

Now we can factor each term to get

$$|x_k y_k - xy| \leq |x_k - x| \, |y_k| + |x| \, |y_k - y| .$$

In the spirit of the proof of the first part, we wish to make each term less than $\varepsilon/2$. For the second term, this is relatively easy: assuming $x \neq 0$, we can change the *desired* inequality

$$|x| \, |y_k - y| < \frac{\varepsilon}{2}$$

into

$$|y_k - y| < \frac{\varepsilon}{2\,|x|}$$

which we can guarantee (since $\varepsilon/2\,|x| > 0$) for all $k \geq N_1$, for some choice of N_1 (because $y_k \to y$)[13]. The *desired* inequality for the *first* term

$$|x_k - x| \, |y_k| < \frac{\varepsilon}{2}$$

is a bit more subtle, because the factor $|y_k|$ changes with k. However, we know by Proposition 2.2.3 that any convergent sequence is bounded, so it is possible to find $M \in \mathbb{R}$ such that

$$|y_k| \leq M \quad \text{for all } k.$$

Because we can always increase M (this only makes the inequality easier to attain), we can assume $M > 0$. Then our *desired* inequality translates into

$$|x_k - x| < \frac{\varepsilon}{2M}$$

which we can guarantee via $k \geq N_2$ for some choice of N_2, since $x_k \to x$. Now, let us write out a FORMAL PROOF of convergence. Given $\varepsilon > 0$, use $y_k \to y$ to pick N_1 such that $k \geq N_1$ guarantees

$$|y_k - y| < \frac{\varepsilon}{2\,|x|}$$

[13] If $x = 0$, it is even easier to make the second term small (Exercise 26)

(if $|x| = 0$, use $|x| + 1 = 1$ in place of $|x|$ in the denominator). Also, let $M > 0$ be an upper bound for $\{|y_k|\}$ (which exists because the y_k are convergent, hence bounded), and use $x_k \to x$ to pick N_2 so that $k \geq N_2$ guarantees

$$|x_k - x| < \frac{\varepsilon}{2M}.$$

Now, let

$$N \geq \max\{N_1, N_2\}.$$

Then $k \geq N$ guarantees

$$\begin{aligned} |x_k y_k - xy| &\leq |x_k - x|\,|y_k| + |x|\,|y_k - y| \\ &< \frac{\varepsilon}{2M} \cdot M + |x|\,\frac{\varepsilon}{2\,|x|} \\ &= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

as required. $\Diamond$

Proof of (4):
This is much like (3), but has a few new twists. Again, given $\varepsilon > 0$ we need to find N so that $k \geq N$ guarantees

$$\left|\frac{x_k}{y_k} - \frac{x}{y}\right| < \varepsilon.$$

As in (3), we subtract-and-add a term in the middle that allows us (via Proposition 1.2.1(1)) to work with two expressions that factor nicely.

$$\begin{aligned} \left|\frac{x_k}{y_k} - \frac{x}{y}\right| &= \left|\frac{x_k}{y_k} - \frac{x}{y_k} + \frac{x}{y_k} - \frac{x}{y}\right| \\ &\leq \left|\frac{x_k}{y_k} - \frac{x}{y_k}\right| + \left|\frac{x}{y_k} - \frac{x}{y}\right| \\ &= \frac{|x_k - x|}{|y_k|} + \frac{|x|\,|y_k - y|}{|y_k|\,|y|}. \end{aligned}$$

Again, we want to force each term to be less than $\frac{\varepsilon}{2}$. But now we have the variable quantity $|y_k|$ in *both* terms. We cannot use the boundedness of $\{y_k\}$ to save us, because an *upper* bound on the denominator can only give a *lower* bound on the fraction. However, we use a different trick: since $y_k \to y$ and $|y| > 0$, we can find N_1 so that $k \geq N_1$ guarantees

$$|y_k - y| < \frac{|y|}{2}.$$

Notice (draw a picture!) that this implies

$$|y_k| > |y| - |y_k - y| > |y| - \frac{|y|}{2} = \frac{|y|}{2} > 0$$

and this *lower* bound on the denominator helps lead to an *upper* bound for the fraction. If we replace $|y_k|$ with the lower bound $\frac{|y|}{2}$ in both terms above, we see that it is enough to force

$$\frac{|x_k - x|}{|y|/2} = \frac{2}{|y|}|x_k - x| < \frac{\varepsilon}{2} \tag{2.1}$$

and

$$\frac{|x|\,|y - y_k|}{\frac{|y|}{2}|y|} = \frac{2|x|}{|y|^2}|y - y_k| < \frac{\varepsilon}{2}. \tag{2.2}$$

But the first estimate (2.1) is accomplished once

$$|x_k - x| < \frac{\varepsilon |y|}{4}$$

which is guaranteed by $k \geq N_2$ for some N_2.
The second estimate (2.2) holds once

$$|y - y_k| < \frac{\varepsilon |y|^2}{4|x|}$$

which can be arranged via $k \geq N_3$ for some N_3.
Let's turn this into a FORMAL PROOF:
Given $\varepsilon > 0$, pick

- N_1 so that $k \geq N_1$ guarantees (via $y_k \to y$)

  $$|y_k - y| < \frac{|y|}{2}, \quad \text{hence} \quad |y_k| > \frac{|y|}{2}.$$

- N_2 so that $k \geq N_2$ guarantees (via $x_k \to x$)

  $$|x_k - x| < \frac{\varepsilon |y|}{4}$$

- N_3 so that $k \geq N_3$ guarantees

  $$|y - y_k| < \frac{\varepsilon |y|^2}{4|x|}.$$

Now set

$$N \geq \max\{N_1, N_2, N_3\};$$

for $k \geq N$, we have

$$\begin{aligned} \left|\frac{x_k}{y_k} - \frac{x}{y}\right| &\leq \frac{|x_{+k} - x|}{|y_k|} + \frac{|x|}{|y_k|\,|y|}\,|y - y_k| \\ &\leq \frac{2}{|y|} \cdot \frac{\varepsilon\,|y|}{4} + \frac{|x|}{|y|^2/2} \cdot \frac{\varepsilon\,|y|^2}{4\,|x|} \\ &= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

$\Diamond$

Proof of (5):
Note that if we divide the terms by x, we have

$$y_k = \frac{x_k}{x} \to 1$$

and the result we want would say $\sqrt[n]{y_k} \to 1$. Let us first concentrate on this special case:

> **Claim:** *if $y_k \to 1$, then $\sqrt[n]{y_k} \to 1$.*

Suppose we have a number $b > 1$; then any (positive integer) power or root of b is also > 1 (Exercise 24). In particular, multiplying both sides of

$$\sqrt[n]{b^{n-1}} > 1$$

by $\sqrt[n]{b}$, we get (for any $b > 1$)

$$b = \left(\sqrt[n]{b}\right)\left(\sqrt[n]{b^{n-1}}\right) > \left(\sqrt[n]{b}\right)(1) = \sqrt[n]{b}.$$

So, for every $b > 1$,

$$b > \sqrt[n]{b} > 1.$$

Now, if $0 < a < 1$, then $b = \frac{1}{a} > 1$, and we have

$$\frac{1}{a} > \frac{1}{\sqrt[n]{a}} > 1$$

or (taking reciprocals)

$$a < \sqrt[n]{a} < 1.$$

In either case, $\sqrt[n]{a}$ is between a and 1. It follows that for every number $y > 0$,

$$|1 - \sqrt[n]{y}| \leq |1 - y| .$$

But then if $y_k \to 1$, given $\varepsilon > 0$, we can find K so that $k \geq K$ guarantees $|1 - y_k| < \varepsilon$, and then

$$|1 - \sqrt[n]{y_k}| < |1 - y_k| < \varepsilon.$$

This shows our claim to be true: whenever $y_k \to 1$, we also have $\sqrt[n]{y_k} \to 1$. $\diamondsuit$

Now, look at the general case. Suppose $x_k \to x > 0$. Let

$$y_k = \frac{x_k}{x} \to 1,$$

so by the claim

$$\sqrt[n]{y_k} = \frac{\sqrt[n]{x_k}}{\sqrt[n]{x}} \to 1.$$

But then

$$\sqrt[n]{x_k} = \sqrt[n]{x}\,\sqrt[n]{y_k} \to \left(\sqrt[n]{x}\right)(1) = \sqrt[n]{x}$$

by part (3), proving part (5). $\diamondsuit$ $\square$

As an immediate application of Theorem 2.4.1, we can evaluate any limit of the form $\lim \frac{p(n)}{q(n)}$, where $p(x)$ and $q(x)$ are polynomials. We illustrate with three examples.

Consider first the limit

$$\lim \frac{3n^2 - 2n + 5}{2n^2 - 3n - 1}.$$

In this form, it is not obvious what is going on: each of the terms $3n^2$ and $2n$ in the numerator is getting large, as are $2n^2$ and $3n$ in the denominator; however, in each case they are subtracted from one another, so it is not *a priori* clear even whether the numerator and/or denominator is increasing or decreasing with n. However, we might have the intuition that n^2 will become large much faster than n, and might be led to concentrate on the leading terms of the numerator and denominator. This can be made precise by the trick of dividing both the numerator and the denominator by n^2 (and hence not changing the value of the quotient); this leads to the limit

$$\lim \frac{3 - \frac{2}{n} + \frac{5}{n^2}}{2 - \frac{3}{n} - \frac{1}{n^2}}$$

in which every individual term in the numerator and the denominator, with the exception of the constants 3 and 2, is clearly going to zero. Thus, we can conclude that

$$\lim \frac{3n^2 - 2n + 5}{2n^2 - 3n - 1} = \lim \frac{3 - \frac{2}{n} + \frac{5}{n^2}}{2 - \frac{3}{n} - \frac{1}{n^2}} = \frac{3 - 0 + 0}{2 - 0 - 0} = \frac{3}{2}.$$

As a second example, we consider

$$\lim \frac{3n - 2}{2n^2 - 3n + 1}.$$

Again, several individual terms in this grow very large, but we expect the leading term in the denominator, $2n^2$, to grow faster than the rest. To analyze this, we again divide the numerator and the denominator by n^2 to obtain

$$\lim \frac{3n - 2}{2n^2 - 3n + 1} = \lim \frac{\frac{3}{n} - \frac{2}{n^2}}{2 - \frac{3}{n} + \frac{1}{n^2}} = \frac{0 - 0}{2 - 0 + 0} = 0.$$

As a final example, consider the limit of the sequence

$$\lim \frac{n^3 - 3n^2 + 2n}{100n^2 + 200}$$

and again divide the numerator and denominator by n^3, the highest power of n that appears anywhere in our expression, leading to

$$\lim \frac{n^3 - 3n^2 + 2n}{100n^2 + 200} = \lim \frac{1 - \frac{3}{n} + \frac{2}{n^2}}{\frac{100}{n} + \frac{200}{n^3}}.$$

This is a little more subtle: the numerator of our new expression goes to 1 while the denominator goes to 0. One reasonable intuitive argument is that a number close to 1 divided by a very small quantity (which in our case is clearly positive) must get very large positive. However, if the signs in the denominator were not all positive, we could be uncertain about the sign of the ratio. A better approach in this case would have been to divide the numerator and denominator by a lower power of n, designed to turn the denominator into a single term. For example, if we divided the original numerator and denominator by n^2, we would have

$$\lim \frac{n^3 - 3n^2 + 2n}{100n^2 + 200} = \lim \frac{n - 3 + \frac{2}{n^2}}{100 + \frac{200}{n^2}}$$

in which the denominator clearly goes to 100, while the numerator becomes $n - 3$ (which gets large positive) added to $\frac{2}{n}$ which clearly goes to zero; we conclude that

$$\lim \frac{n^3 - 3n^2 + 2n}{100n^2 + 200} = \infty.$$

With a little thought, we see that these three examples illustrate the three cases of the following

Remark 2.4.2. *Suppose $p(x) = a_k x^k + a_{k-1}x^{k-1} + ... + a_0$ ($a_k \neq 0$) and $q(x) = b_m x^m + b_{m-1}x^{m-1} + ... + b_0$ ($b_m \neq 0$) are polynomials of degree k and m, respectively, and consider the limit*

$$\lim \frac{p(n)}{q(n)}.$$

Then

1. *If $k < m =$, then $\frac{p(x)}{q(x)} \to 0$;*
2. *If $k > m$, then $\frac{p(x)}{q(x)} \to \pm\infty$, where the sign is the same as the sign of the ratio of the leading terms, a_k/b_m;*
3. *If $k = m$, then $\frac{p(x)}{q(x)} \to \frac{a_k}{b_m}$.*

Theorem 2.4.1 can also be used to *evaluate* certain limits whose *existence* we can establish via the completeness axiom. For example, the repeating decimal expression

$$x = 0.153153153\ldots$$

identifies x as the limit of the sequence of decimal fractions x_k obtained by truncating the above after k digits:

$$x_k = 0.\underbrace{153153\ldots}_{k \text{ digits}}.$$

We know that the x_k converge, by Remark 2.3.4. Now note that if we divide x_k by 1000, we shift all digits to the right 3 spaces:

$$\frac{x_k}{1000} = 0.000\underbrace{153153\ldots}_{k \text{ digits}}.$$

Compare this with the term three places further along the sequence:

$$\begin{aligned} x_{k+3} &= 0.\underbrace{153153\ldots}_{k+3 \text{ digits}} \\ &= 0.153\underbrace{153153\ldots}_{k \text{ digits}}. \end{aligned}$$

Because of the repetition, all digits from the fourth through the k^{th} in the two decimals match up. This means we can write

$$\begin{aligned} x_{k+3} - \frac{x_k}{1000} &= 0.153\ \underbrace{000\ldots0}_{(k \text{ zeroes})} \\ &= 0.153. \end{aligned}$$

We can conclude that

$$\lim \left(x_{k+3} - \frac{x_k}{1000} \right) = 0.153 = \frac{153}{1000}.$$

But from the results above, we also know that

$$\lim \left(x_{k+3} - \frac{x_k}{1000} \right) = x - \frac{x}{1000} = \frac{999x}{1000}.$$

From this we conclude

$$\frac{999x}{1000} = \frac{153}{1000}$$

or

$$x = \frac{153}{999} = \frac{17}{111}.$$

It should be clear that a similar calculation shows that *any* (eventually) repeating decimal expression represents a rational number, whose value (as a ratio of integers) can be determined explicitly.
As another example, consider the sequence

$$x_k = r^k, \quad k = 1, 2, ...$$

where $0 < r < 1$. We note first that this sequence is strictly decreasing (use ratios!) and bounded below by zero, hence it converges. Let

$$x = \lim x_k.$$

Our intuition says that $x = 0$. We shall verify this.

Note that

$$x_{k+1} = rx_k$$

so we can conclude (from Theorem 2.4.1) that

$$\lim x_{k+1} = r \lim x_k = rx.$$

But of course, also

$$\lim x_{k+1} = \lim x_k = x$$

so we have

$$x = rx$$

or

$$(1 - r)x = 0.$$

Since $r \neq 1$, we conclude that

$$x = 0.$$

Pursuing this further, we note that when $-1 < r < 0$, we can look at

$$|x_k| = |r|^k,$$

and $0 < |r| < 1$, so $|x_k| \to 0$, hence $x_k \to 0$. We have shown [14]

Lemma 2.4.3. *If* $|r| < 1$*, then* $\lim r^k = 0$.

Now, look at the same sequence when $r > 1$. Our whole calculation carries through without a hitch (we *only* used that $r \neq 1$: if $x = \lim r^k$, then $x = 0$). But of course this is nonsense! We know if $r > 1$ then $r^k > 1$ for all k, so x must be at least 1.

What's wrong? *We have forgotten to check convergence.* In fact, the sequence r^k is strictly increasing for $r > 1$ (can you prove this?). If it were bounded above, it would have to converge. On one hand, we know from Lemma 2.2.9 that $x = \lim r^k$ would have to be at least 1 ($x \geq 1$), since $x_k = r^k > 1$ for all k. But on the other hand, we have just established that x must satisfy $(1 - r)x = 0$, and since $r \neq 1$, this means $x = 0$. The contradiction shows that we must have r^k *unbounded* (and hence by Proposition 2.3.3, diverging to ∞) if $r > 1$.

Lemma 2.4.4. *If* $r > 1$*, then* $r^k \to \infty$.

[14] with one slight gap—what is it, and how do we fill it?

An immediate consequence of this lemma is that if $r < -1$, then the sequence r^k is unbounded (in fact, $|r^k| \to \infty$), and hence diverges (but, since the signs alternate, we don't have divergence to infinity).
Using these lemmas, we can investigate a particularly well-behaved family of series.

Definition 2.4.5. *A series $\sum_{i=0}^{\infty} a_i$ is a **geometric series** with **ratio** r if all terms are nonzero and*

$$\frac{a_{i+1}}{a_i} = r$$

independent of i for $i = 0, 1, \ldots$.

It is, of course, possible to obtain an explicit formula for a_i from the recursive condition

$$a_{i+1} = ra_i$$

namely,

$$a_i = a_0 r^i.$$

From this, we see that the partial sums are

$$\begin{aligned} S_k = \sum_{i=0}^{k} a_i &= a_0 + \ldots + a_k \\ &= a_0 + a_0 r + +a_0 r^2 + \ldots + a_o r^k \\ &= a_0(1 + r + r^2 + \ldots + r^k). \end{aligned}$$

Now, a trick: if we multiply S_k by r, we get all but the first term of this last sum, together with an extra term $a_0 r^{k+1}$. Subtracting, this gives

$$(1 - r)S_k = a_0 - a_0 r^{k+1},$$

or (if $r \neq 1$)

$$S_k = \frac{a_0(1 - r^{k+1})}{1 - r}.$$

We have seen that if $|r| < 1$, then the term r^{k+1} in this expression goes to zero, so that

if $|r| < 1$, then

$$S_k \to \frac{a_0(1 - 0)}{1 - r} = \frac{a_0}{1 - r}.$$

Also, if $r > 1$, then $r^{k+1} \to \infty$, and $S_k \to \infty$ (see Exercise 29). This is the essential part of the proof of the following

Proposition 2.4.6 (Geometric Series). *If $\sum_{i=0}^{\infty} a_i$ is a geometric series with ratio r, then*

- *if $|r| \geq 1$, the series diverges;*
- *if $|r| < 1$, then the series converges, and*

$$\sum_{i=0}^{\infty} a_i = \frac{a_0}{1-r}.$$

It should be noted that geometric series had been dealt with by the Greeks; for example Archimedes of Syracuse (*ca.*212-287 BC) in *The Quadrature of the Parabola* used the special case of Proposition 2.4.6 with $r = \frac{1}{4}$; however, he did not use the notion of a limit per se, but rather showed that an infinite sum of areas equivalent to this series sums to $\frac{4}{3}$ times the initial area via a proof by contradiction[15] [31, pp. 249-50].
For an elegant geometric proof of Proposition 2.4.6 for $0 < r < 1$, due to Leibniz, see Exercise 35.
Besides Theorem 2.4.1, the other important tool for calculating limits is the following result[16].

Theorem 2.4.7 (Squeeze Theorem). *Suppose $\{a_k\}$, $\{b_k\}$ and $\{c_k\}$ are three real sequences such that*

1. *eventually $a_k \leq c_k \leq b_k$*
2. *$\{a_k\}$ and $\{b_k\}$ both converge to the same limit*

$$\lim a_k = L = \lim b_k.$$

Then

$$c_k \to L.$$

Proof. Since (eventually) c_k is always between a_k and b_k, we know that c_k is closer to L than the farther of a_k and b_k:

$$|c_k - L| \leq \max\{|a_k - L|, |b_k - L|\}.$$

But for k large, both of the distances on the right are small, and hence so is the distance $|c_k - L|$ on the left.
More precisely, given $\varepsilon > 0$ pick

[15] See the proof of Proposition 2.2.1 for a discussion of proof by contradiction.

[16] This is also often referred to as the "Sandwich Theorem" or the "Pinching Theorem". Notice the relation of this to the method of compression of Archimedes (see Exercise 11 in § 5.1).

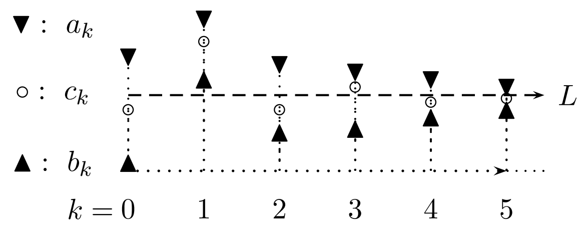


Figure 2.5: Theorem 2.4.7 (Squeeze Theorem)

- N_1 so that $k \geq N_1$ guarantees

$$|a_k - L| < \varepsilon$$

- N_2 so that $k \geq N_2$ guarantees

$$|b_k - L| < \varepsilon$$

- N_3 so that $k \geq N_3$ guarantees

$$a_k \leq c_k \leq b_k.$$

Then $k \geq N = \max\{N_1, N_2, N_3\}$ guarantees

$$|c_k - L| \leq \max\{|a_k - L|, |b_k - L|\} < \varepsilon.$$

□

We illustrate two uses of this result.
First, consider the sequence

$$x_k = \frac{1}{k}\cos k, \quad k = 0, 1, 2, ...$$

We know that $1/k \to 0$, but we cannot use Theorem 2.4.1 to conclude $x_k \to 0$, because we don't know that $\cos k$ converges[17]. However, we *do* know that for any angle θ, $|\cos\theta| \leq 1$, so for $k = 1, 2, ...$

$$-\frac{1}{k} \leq \frac{1}{k}\cos k \leq \frac{1}{k}.$$

[17] In fact, $\cos k$ *diverges* (see Exercise 33, as well as Exercise 13 in § 3.1).

Now, since

$$\lim \left(-\frac{1}{k}\right) = 0 = \lim \frac{1}{k},$$

we can use the squeeze theorem with $a_k = -\frac{1}{k}$, $b_k = \frac{1}{k}$, and $c_k = x_k = \frac{1}{k}\cos k$ to conclude that

$$x_k \to 0.$$

As a second application, we consider another class of series. We will say a series of nonzero terms is an **alternating series** if the signs of the terms strictly alternate between positive and negative. One example of an alternating series is a geometric series with ratio $r < 0$. Another important example is the **alternating harmonic series**

$$\sum_{i=0}^{\infty} \frac{(-1)^i}{i+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \dots$$

The **harmonic series**

$$\sum_{i=0}^{\infty} \frac{1}{i+1} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots$$

diverges (see Exercise 34 in § 2.3). However, the following result shows that the *alternating* harmonic series *converges.*

Proposition 2.4.8 (Alternating Series Test)**.** *Suppose a series* $\sum_{i=0}^{\infty} a_i$ *of nonzero terms satisfies*

- *for* $i = 0, 1, ...$
$$-1 < \frac{a_{i+1}}{a_i} < 0$$
- $|a_i| \to 0.$

Then the series converges.

We comment that the first condition says that the summands alternate sign, and have decreasing absolute values; the second condition is that they go to zero.

Proof. We will assume the first term is positive

$$a_0 > 0;$$

the analogous argument if $a_0 < 0$ is left to you (Exercise 28).
Let us consider the relative positions of the various partial sums. We have

$$S_0 = a_0$$

and, since $a_1 < 0$,

$$S_1 = S_0 + a_1 = S_0 - |a_1| < S_0.$$

Now

$$S_2 = S_1 + a_2$$

and

$$0 < a_2 < |a_1|$$

so

$$S_1 < S_2 < S_0.$$

Then, since $a_3 < 0$,

$$S_3 = S_2 + a_3 = S_2 - |a_3| < S_2$$

but also, since $|a_3| < a_2$,

$$S_3 = S_1 + a_2 - |a_3| > S_1$$

so

$$S_1 < S_3 < S_2 < S_0.$$

(see Figure 2.6)

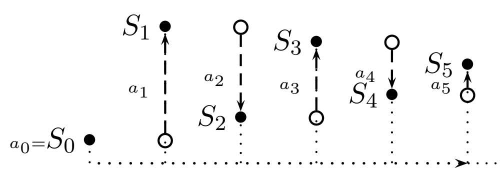


Figure 2.6: Alternating Series

At each successive stage, we move in the direction *opposite* to our previous step, but a *shorter* distance, so that S_{k+1} is always between S_k and S_{k-1}. Put more precisely, for each k the two differences

$$S_k - S_{k-1} = a_k$$
$$S_{k+1} - S_{k-1} = a_k + a_{k+1}$$

have the same sign (as a_k), since a_k and a_{k+1} have opposite sign, with $|a_{k+1}| < |a_k|$. Furthermore,

$$|S_{k+1} - S_{k-1}| = |a_k| - |a_{k+1}| < |a_k| - |S_k - S_{k-1}|$$

so that for all k, S_{k+1} lies between S_k and S_{k-1}.
In the case we are considering ($a_0 > 0$) the *odd*-numbered partial sums S_{2j+1}, $j = 0, 1, ...$ form an *increasing* sequence bounded *above* by S_0 while the *even*-numbered partial sums S_{2j}, $j = 0, 1, ...$ are *decreasing* and bounded *below* by S_1. It follows that each of these two sequences converges, and since

$$|S_{2j+1} - S_{2j}| = |a_{2j+1}| \to 0$$

we see that they have the same limit. Now, form two new sequences:

$$\alpha_i = \begin{cases} S_i & \text{if } i \text{ is odd} \\ S_{i+1} & \text{if } i \text{ is even} \end{cases}$$

$$\beta_i = \begin{cases} S_i & \text{if } i \text{ is even} \\ S_{i-1} & \text{if } i \text{ is odd.} \end{cases}$$

That is, the α's are the odd-numbered partial sums, each repeated once, while the β's are the even-numbered ones (with repetition). Then we also have

$$\lim \alpha_i = \lim \beta_i$$

and furthermore,

$$\alpha_k \le S_k \le \beta_k \quad \text{for } k = 0, 1, ...$$

Hence the S_k converge by the Squeeze Theorem. □

Exercises for § 2.4

Answers to Exercises 1-15 (odd only), 19, 20-21aceg, 22 are given in Appendix B.

Practice problems:

Problems 1-19 are numbered according to the list of sequences on p. 19; problems 10 and 17-18 are omitted.
For each of sequences 1-9, 11-16, and 19 in the list on p. 19, either find the limit or show that none exists.

20. Express the number given by each (eventually repeating) decimal expansion below as a ratio of two integers:

    (a) $0.111\ldots$ (b) $0.1010\ldots$ (c) $0.0101\ldots$ (d) $1.0101\ldots$
    (e) $3.123123\ldots$ (f) $3.1212\ldots$ (g) $1.0123123\ldots$

21. Give the decimal expansion of each rational number below

    (a) $\frac{3}{4}$ (b) $\frac{2}{3}$ (c) $\frac{5}{6}$ (d) $\frac{2}{9}$
    (e) $\frac{3}{7}$ (f) $\frac{7}{11}$ (g) $\frac{11}{7}$

22. Consider a sequence of the form

$$x_n = \frac{an^p + 12n - 5}{n^{12} + 3n^2}.$$

    For each of the following, find values of a and p for which the given phenomenon occurs. The answer to some parts need not be unique.

    (a) x_n diverges to ∞.
    (b) $\lim x_n = 10$.
    (c) $\lim x_n = 0$.
    (d) x_n diverges to $-\infty$.

Theory problems:

23. Suppose we have $\lim x_k = L$. Show that $\lim |x_k| = |L|$ as follows:

    (a) If $L > 0$, show that eventually $x_k > 0$. How does this prove our result in this case?
    (b) Similarly, if $L < 0$, eventually $x_k < 0$. Use this to prove the result in this case.
    (c) What about if $L = 0$?
    (d) Now consider the converse statement. Suppose we know that the sequence $|x_k|$ converges. Does it follow that x_k converges? Prove that it does, or give a counterexample.

24. (a) Show that if $0 < a < b$ then $0 < a^n < b^n$ for any integer $n > 0$.
    (b) Use this to show that if $0 < A < B$ then $0 < \sqrt[n]{A} < \sqrt[n]{B}$. (*Hint:* Consider the alternatives.)

25. Give a formal proof that, if $x_k \to x$ and $y_k \to y$, then $x_k - y_k \to x - y$. (*Hint:* Mimic the proof for sums, with appropriate modifications.)

26. In the proof of Theorem 2.4.1 part 3, what do you do if $x = 0$?

27. Prove Remark 2.4.2 as follows:

    (a) If the degree of the denominator, m, is at least equal to the degree k of the numerator, divide both by n^m; then the denominator consists of b_m plus terms that go to zero, while every term in the numerator except possibly the one coming from $a_m x^m$, if $k = m$, goes to zero. Then use Theorem 2.4.1 to show that $p(n)/q(n) \to 0$ if $k < m$ and $p(n)/q(n) \to a_k/b_k$ if $k = m$.

    (b) If $k > m$, use the above to show that $p(n)/a_k n^k \to 1$. Then show that $a_k n^k/q(n) \to \pm\infty$, where the sign is that of a_k/b_m, by writing this fraction as n^{k-m} times a fraction which goes to a_k/b_m.

28. Give the proof of the Alternating Series Test when $a_0 < 0$. (*Hint:* What changes from the proof given for $a_0 > 0$?)

29. In the proof of divergence for a geometric series with ratio $r > 1$, we have $1 - r^{k+1} \to -\infty$. So why does $S_k \to \infty$, not $-\infty$?

Challenge problems:

30. (a) For sequence 26 in the list on page 19, (i) Look at the first few terms and try to identify this as a geometric series. (ii) Prove your guess. (*Hint:* Try proof by induction[18].) (iii) Use this to find the limit of the sequence or else to show divergence.

    (b) Relate sequence 27 in the same list to sequence 26, and use this information to determine the limit or show that the sequence diverges.

31. (a) Show that every repeating decimal expansion converges to a rational number.

[18] See Appendix A.

(b) Is it true that the decimal expansion of any rational number must be (eventually) repeating? (Give some reasons to support your conjecture, and give a proof if you can.)

(c) Give an example of a fraction (as a ratio of integers) for which the decimal expansion starts repeating only after the fourth position to the right of the decimal point.

32. A **binary expansion** for a real number $x \in (0, 1)$ is a sequence $\{b_k\}$, where each b_k is either 0 or 1, such that

$$\sum_{k=1}^{\infty} \frac{b_k}{2^k} = x.$$

(a) Find a binary expansion for each of the numbers $\frac{1}{3}$ and $\frac{5}{6}$.

(b) Here is a systematic procedure for finding a binary expansion of an arbitrary rational number $\frac{p}{q}$, $0 < p < q$:

i. Consider the multiples of $\frac{p}{q}$ by powers of 2, writing each as a mixed fraction

$$2^k \frac{p}{q} = M_n + \frac{r_n}{q},$$

where M_n and r_n are non-negative integers with $0 \leq r_n < q$. **Show** that $M_n < 2^n$.

ii. Since there are only finitely many possibilities for r_n, they must repeat: for some n and $k > 0$,

$$r_n = r_{n+k}.$$

Show that we can assume $n + k \leq q$.

iii. **Show** that if the expansion in powers of 2 of $2^{n+k}\frac{p}{q}$ is

$$2^{n+k}\frac{p}{q} = b_1 \cdot 2^{n+k-1} + b_2 \cdot 2^{n+k-2} + \cdots$$
$$\cdots + b_{n+k} \cdot 2^0 + b_{n+k+1} \cdot 2^{-1} + b_{n+k+2} \cdot 2^{-2} + \cdots$$

with each $b_i = 0$ or 1, then

•

$$M_{n+k} = b_1 \cdot 2^{n+k-1} + b_2 \cdot 2^{n+k-2} + ... + b_{n+k} \cdot 2^0$$

and

$$M_n = b_1 \cdot 2^{n-1} + b_2 \cdot 2^{n-2} + ... + b_n \cdot 2^0.$$

- $b_{n+k+j} = b_{n+j}$ for $j = 1, 2, ...;$
- A binary expansion of $\frac{p}{q}$ is

$$\frac{p}{q} = \sum_{i=1}^{\infty} \frac{b_i}{2^i}.$$

(c) Use this to find the binary expansion of $\frac{1}{5}$, $\frac{1}{25}$ and $\frac{1}{100}$.

(d) Which numbers have a finite binary expansion (*i.e.*, all terms after some place are zeroes)?

(e) Which numbers have more than one binary expansion?

(f) In general, show that for any integer $\beta > 1$, every $x \in (0,1)$ has an expansion in powers of β

$$x = \sum_{k=1}^{\infty} \frac{b_k}{\beta^k}$$

where each b_k is an integer satisfying $0 \leq b_k < \beta$. What about if $\beta > 1$ is *not* an integer?

33. Give a proof by contradiction showing that the sequence $\{\cos k\}$ diverges, as follows: suppose it converges to $c := \lim \cos k$.

(a) Show that $\{\sin k\}$ also converges, where $s := \lim \sin k$ satisfies

$$s \cdot \sin 1 = c \cdot (\cos 1 - 1).$$

(*Hint:* Use the angle summation formula for $\cos\theta$ to express $\sin k$ in terms of $\sin 1$, $\cos k$ and $\cos(k+1)$.)

(b) Use this to show that either s and c are both zero, or neither is.

(c) Show that we can't have $s = 0 = c$, so by the previous part, each of s and c is nonzero. (*Hint:* Use the identity $\sin^2\theta + \cos^2\theta = 1$.)

(d) Use the angle summation formula for $\sin\theta$ to show that

$$s \cdot (1 - \cos 1) = c \cdot \sin 1.$$

(e) Conclude that

$$\frac{c}{s} = -\frac{s}{c}.$$

But this implies

$$s^2 = -c^2$$

which is impossible. The contradiction proves the result.[19]

History note:

34. Prop. 1 in Book X of Euclid's *Elements* (ca. 300 BC) reads:

    Two unequal magnitudes being set out, if from the greater there be subtracted a magnitude greater than its half, and from that which is left a magnitude greater than its half, and if the process be repeated continually, there will be left some magnitude which will be less than the lesser magnitude set out.

    Reformulate this as a statement about a (recursively defined) sequence, and relate it to results in this section (but note that the notion of limit does not appear in this or any other statement in Euclid).

    For more on how the Greeks handled notions that we attack with the limit concept, see Exercises 7-13 in § 5.1.

35. [24, p. 26] **Leibniz sums a geometric series:** In a treatise written in 1675-6 but only published in 1993 [39], Leibniz formulated the sum of a geometric series as follows:

    The greatest term of an infinite geometric series is the mean proportional between the greatest sum and the greatest difference.

    Here, he is thinking of a series of *decreasing, positive* terms $\sum_{i=0}^{\infty} a_i$ for which the ratio $a_i : a_{i+1}$ is constant. His statement then is

    $$\frac{\sum_{i=0}^{\infty} a_i}{a_0} = \frac{a_0}{a_0 - a_1}.$$

    (a) Show that this is equivalent to the formula in Proposition 2.4.6. Here is his proof of the formula (see Figure 2.7)

[19] See Exercise 13 in § 3.1 for an even stronger statement.

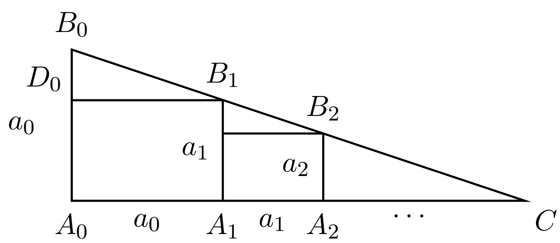


Figure 2.7: Leibniz' summation of a geometric series

(b) Construct the vertical line segment A_0B_0 and horizontal line segment A_0A_1, each of length a_0; then A_1B_1 (vertical) and A_1A_2 (horizontal) of length a_1, and so on. Show that the points B_j, $j = 0, 1, 2, ...$ lie on a line, and let C be where this intersects the horizontal line through A_0.

(c) Show that the points A_j converge to C, and conclude that the length of A_0C is $\sum_{i=0}^{\infty} a_i$.

(d) Let D_0 be the intersection of A_0B_0 with the horizontal through B_1. Show that the triangles $\triangle A_0B_0C$ and $\triangle D_0B_0B_1$ are similar.

(e) Conclude that A_0C is to A_0B_0 as D_0B_1 is to D_0B_0.

(f) Show that this is precisely Leibniz' formula.

(g) Can you see how to handle the case $-1 < r < 0$ by these methods?

2.5 Bounded Sets

So far we have dealt primarily with intervals on the real line. However, it is often useful to be able to speak about more general kinds of sets. In general, a **set** is simply a collection of objects, called the **elements** of the set. A set can be specified in one of several ways. One way is to simply list the elements of the set: **list notation** consists of just such a list, enclosed in braces. For example, the collection S of all integers between 2 and 50 which are perfect squares could be written

$$S := \{4, 9, 16, 25, 36, 49\}.$$

A number belongs to this set precisely if it equals one of the six numbers on this list. Note that in list notation, order *doesn't* matter: the set

$$\{4, 16, 36, 49, 25, 9\}$$

has precisely the same elements as the one specified above, and hence it is the *same* set. List notation is most useful for sets with finitely many elements, but sometimes it is possible to indicate an infinite list, like the positive integers

$$\{1, 2, 3, 4, 5, \ldots\}$$

where, once the pattern is clear, we write "…" to read "*and so on*". Of course, list notation is not feasible for sets like intervals, since there is no systematic way to list *all* the real numbers between two endpoints.
Another way to specify a set is to give a criterion for deciding whether or not a given object belongs to the set. The format for this is **set-builder notation**. In this notation, the set S of perfect squares between 2 and 50 would be written

$$S := \{x \in \mathbb{R} \,|\, 2 \leq x \leq 50 \text{ and } x = y^2 \text{ for some integer } y\}.$$

The general form of set-builder notation is

$$\{\mathit{objects} \,|\, \mathit{property}\}$$

where the expression "objects" to the left of the vertical bar specifies the general kinds of objects that are considered potential members of the set, while to the right is an "admission test": a list of properties which an object must exhibit in order to belong to the set. Every object (of the sort specified on the left) which passes the test(s) on the right belongs to this set. In our example, the expression for S would be pronounced "*S is the set of numbers x such that x is at least 2 and at most 50 and equals y^2 for some integer y*".
We extend the notation for membership which we used earlier in connection with intervals to indicate membership in any set. Thus, for example, we could assert that

$$25 \in S$$

where S is the set specified above (this would be pronounced "25 *belongs to (or is a member of) S*". We also use the notion of a *subset*: a set A is a **subset** of the set B (denoted $A \subset B$) if every element of A is also an element of B. For example, the collection $\{4, 16, 36\}$ is a subset of the set S described above:

$$\{4, 16, 36\} \subset S = \{4, 9, 16, 25, 36, 49\}.$$

We have already spoken of bounded intervals and bounded sequences; this notion extends naturally to any set of real numbers.

Definition 2.5.1. *Suppose S is a set of real numbers.*

1. *A **lower bound** (resp. **upper bound**) for S is any number $\alpha \in \mathbb{R}$ (resp. $\beta \in \mathbb{R}$) such that $\alpha \leq s$ (resp. $s \leq \beta$) for every $s \in S$; we will sometimes write these inequalities as $\alpha \leq S$ (resp. $S \leq \beta$).*
2. *S is **bounded below** (resp. **bounded above**) if there exists a lower (resp. upper) bound for S.*
3. *S is **bounded** if it is bounded above and also bounded below.*

The following are pretty obvious observations about bounded sets (Exercise 6).

Remark 2.5.2. *Suppose $S \subset \mathbb{R}$ is a set of real numbers.*

1. *If S is finite, it is bounded.*
2. *S is bounded precisely[20] if it is contained in some bounded interval.*
3. *S is bounded precisely if the set $\{|s| \mid s \in S\}$ of absolute values of elements of S is bounded* above.
4. *If S is bounded, so is any subset of S.*
5. *S is **unbounded** precisely if it contains some sequence which is unbounded.*

While a bound for S need not itself belong to S, the lowest (*resp.* highest) number in S is an element of S for which the condition $\alpha \leq S$ (*resp.* $S \leq \beta$) holds. Such numbers are called, respectively, the *minimum* (*resp.* *maximum*) of S.[21]

[20]Note the meaning of the word "precisely" in parts 2, 3, and 5: for example, in part 2, there are two assertions being made: first, that every bounded set is contained in some bounded interval, and second, that any set which is contained in some bounded interval is itself bounded. Our use of the word "precisely" is often rendered via the phrase "if and only if" (sometimes abbreviated "iff"); in my experience, this phrase often proves confusing to the novice, so I have avoided it in this book.

[21]Are we justified in referring to *the* minimum?

Definition 2.5.3. *Given a set $S \subset \mathbb{R}$,*

1. *$\alpha \in \mathbb{R}$ is the* ***minimum*** *of S, denoted*

$$\alpha = \min S = \min_{x \in S} x$$

if α is a lower bound for S which itself belongs to S:

   (a) $\alpha \leq S$

   (b) $\alpha \in S$.

2. *$\beta \in \mathbb{R}$ is the* ***maximum*** *of S*

$$\beta = \max S = \max_{x \in S} x$$

if β is an upper bound for S which itself belongs to S:

   (a) $S \leq \beta$

   (b) $\beta \in S$.

A straightforward process of elimination can be used to find the minimum and maximum of a *finite* set. For example, if

$$S = \{-3, 1, -\pi, 0, \sqrt{2}\}$$

we can start out with the first two numbers as tentative candidates for the minimum and maximum

$$a_0 = -3, \quad b_0 = 1$$

then compare the next number, $-\pi$, to both: since $-\pi < a_0$ we replace the lower bound a_0 with $-\pi$ and leave the upper bound alone:

$$a_1 = -\pi, \quad b_1 = 1.$$

Since the next number, 0, lies between a_1 and b_1, we leave both unchanged at the next step:

$$a_2 = -\pi, \quad b_2 = 1.$$

Finally, since $\sqrt{2} > b_2$, we replace b_2 with $\sqrt{2}$:

$$a_3 = -\pi, \quad b_3 = \sqrt{2}.$$

Since all numbers in the list have been checked, we have

$$\min S = -\pi, \quad \max S = \sqrt{2}.$$

However, such a process is unending when the set $S \subset \mathbb{R}$ is *infinite*. In fact, an infinite set $S \subset \mathbb{R}$ may *fail to have* a minimum and/or maximum element for either of two reasons:

- if S is not bounded below (*resp.* not bounded above) then *no* number $\alpha \in \mathbb{R}$ (*resp.* $\beta \in \mathbb{R}$) can satisfy the first condition, $\alpha \leq S$ (*resp.* $S \leq \beta$);
- even if S *is* bounded (above and/or below), it may fail to have a minimum and/or maximum because no *element of* S is a lower (*resp.* upper) bound for S: that is, no number satisfying the first condition can also satisfy the second.

For example, the collection of positive integers

$$S = \{1, 2, 3, \dots\}$$

has $\min S = 1$, but it is not bounded *above*, so it has no maximum. The second situation holds for the open interval

$$S = (0, 1)$$

which is clearly bounded (above *and* below) but has neither a maximum nor a minimum. To see this, note that if $\alpha \in S$, then since $\alpha > 0$, the number $s = \frac{\alpha}{2}$ is an element of S for which the condition $\alpha \leq s$ *fails*, so α cannot be a *lower* bound for S (and a similar argument shows that no element of S can be an *upper* bound for S; see Exercise 4).
By contrast, the *closed* interval $[0, 1]$ has

$$\min_{x \in [0,1]} x = 0, \quad \max_{x \in [0,1]} x = 1.$$

The distinction between the closed and open interval is precisely whether the endpoints do or don't belong to the set, and affects whether the endpoints are the minimum and maximum of the set. However, for *both* sets $[0, 1]$ and $(0, 1)$, the endpoint 0 (*resp.* 1) is the *best* lower (*resp.* upper) bound for the set, in a sense made precise by the following definition.

Definition 2.5.4. *Given a set $S \subset \mathbb{R}$ of numbers,*

1. *$\alpha \in \mathbb{R}$ is the* ***infimum*** *of S*

$$\alpha = \inf S = \inf_{x \in S} x$$

if

- *α is a lower bound for S:*

$$\alpha \leq S$$

- *any* other *lower bound for S is itself lower than α:*

$$\alpha' \leq S \Rightarrow \alpha' \leq \alpha.$$

2. *$\beta \in \mathbb{R}$ is the* ***supremum*** *of S*

$$\beta = \sup S = \sup_{x \in S} x$$

if

- *β is an upper bound for S:*

$$S \leq \beta$$

- *any* other *upper bound for S is itself higher than β:*

$$S \leq \beta' \Rightarrow \beta \leq \beta'.$$

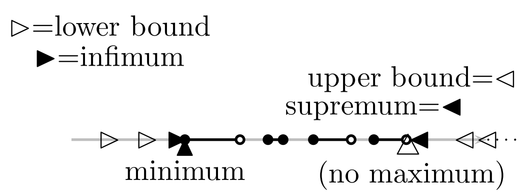


Figure 2.8: Bounds for a set

The infimum (*resp.* supremum) of S is sometimes called the **greatest lower bound** (*resp.* **least upper bound**) for S:

$$\text{glb } S = \inf S, \quad \text{lub } S = \sup S.$$

The supremum and/or infimum of a set S need not themselves belong to the set S. For example, it is easy to see that the infimum (*resp.* supremum) of *any* bounded interval—open, closed, or half-open—is its left (*resp.* right) endpoint (see exercise Exercise 3). For more complicated sets, the determination of the infimum and supremum may be more involved. However, a few observations can be useful in this context.

Lemma 2.5.5. *Given $S \subset \mathbb{R}$,*

- *The minimum (resp. maximum) of S, if it exists, is automatically the infimum (resp. supremum) of S.*
- *If the infimum (resp. supremum) of S itself belongs to the set S, then it is automatically the minimum (resp. maximum) of S.*
- *A lower (resp. upper) bound for S is the infimum (resp. supremum) of S precisely if it is the limit of some sequence in S.*

Proof. The first two statements are a simple matter of chasing definitions (Exercise 5). In proving the third statement, we shall concentrate on *lower* bounds.

Suppose α is a lower bound for S and $s_k \in S$ is a sequence in S converging to α. We need to show $\alpha = \inf S$. Since α is already known to be a lower bound for S, we merely need to show that no *higher* number $\alpha' > \alpha$ can be a lower bound for S. Given $\alpha' > \alpha$, since $s_k \to \alpha$, eventually $|s_k - \alpha| < \alpha' - \alpha$, which guarantees (eventually)

$$s_k < \alpha'$$

and this prevents α' from being a lower bound for S. Since this argument works for *every* $\alpha' > \alpha$, we have shown that α is the *greatest* lower bound for S:

$$\alpha = \inf S.$$

Conversely, suppose $\alpha = \inf S$. We need to find a sequence s_k in S with $s_k \to \alpha$. If it happens that $\alpha \in S$ (that is, if $\alpha = \min S$), then the constant sequence $s_k = \alpha$ for all k works. So assume

$$\alpha = \inf S \notin S.$$

We can pick $s_0 \in S$ arbitrarily, and of course

$$\alpha < s_0.$$

Now, consider the point

$$c_0 = \frac{1}{2}(\alpha + s_0)$$

halfway between α and s_0. Since $c_0 > \alpha$, c_0 is *not* a lower bound for S, and so there must exist $s_1 \in S$ with

$$\alpha < s_1 < c_0.$$

We proceed recursively: having found $s_k \in S$ (so $\alpha < s_k$), we consider the point

$$c_k = \frac{1}{2}(\alpha + s_k)$$

halfway between: since $c_k > \alpha$, c_k is not a lower bound for S, and hence we can find some element $s_{k+1} \in S$ with

$$\alpha < s_{k+1} < c_k.$$

Note that, since s_{k+1} is closer to α than the point c_k halfway between α and s_k, we have

$$|s_{k+1} - \alpha| < \frac{1}{2}|s_k - \alpha|$$

so

$$|s_k - \alpha| \leq \frac{1}{2^k}|s_0 - \alpha| \to 0$$

and hence $\{s_k\}$ is a sequence in S converging to α, as required[22]. □

With the tools provided by Lemma 2.5.5 let us examine a few examples. First, for any *finite* set, we know that the minimum and maximum always exist, and hence agree with the infimum and supremum.
Consider now the set $S := \{\frac{n+1}{n} \mid n \text{ a positive integer}\}$. We see immediately that every element is greater than 1, and that $\frac{n+1}{n} \to 1$ as a (decreasing) sequence. Thus $\inf S = 1$, and since no element equals 1, the minimum does not exist. Since the sequence $\frac{n+1}{n}$ is decreasing, the highest value will be achieved when $n = 1$, so $\max S = \sup S = \frac{1+1}{1} = 2$.
Next, consider $S = \{(1 - 2^p)/3^q \mid p, q \text{ integers with } 0 < p < q\}$. Since $p \geq 1$, the numerator is always negative, so the fraction is negative, and since $q \geq p + 1$, we know that for p fixed, $(1 - 2^p)/3^q$ is a sequence in S starting with $(1 - 2^p)/3^{p+1}$ and increasing to 0. Thus $\sup S = 0$ (and S has no maximum). Furthermore, since the lowest value of $(1 - 2^p)/3^q$ for any particular p is $(1 - 2^p)/3^{p+1}$, we can consider the possible values of this last fraction, which can be rewritten as

$$\frac{1 - 2^p}{3^{p+1}} = \frac{1}{3}\left(\frac{1}{3^p} - \left(\frac{2}{3}\right)^p\right)$$

The value of this fraction is $-\frac{1}{9}$ for $p = 1, 2$ and after that it increases toward zero (to see this, consider the values for p and $p + 1$, and compare by cross-multiplying and some algebra). Thus, $\inf S = \min S = -\frac{1}{9}$.

[22] This argument is essentially an application of the bisection algorithm, given in general in § 2.6.

We now formulate a basic fact about bounded sets, which follows from the completeness axiom.[23]

Theorem 2.5.6. *If $S \subset \mathbb{R}$ is nonempty and bounded below (resp. bounded above), then* $\inf S$ *(resp.* $\sup S$*) exists.*

Proof. We will deal with the case that S is bounded below. Suppose we have a lower bound for S:

$$\alpha \leq s \text{ for every } s \in S.$$

We will use a variant of the bisection algorithm (Proposition 2.6.3) together with the Nested Interval Property (Lemma 2.6.1) to locate $\inf S$. We start with

$$a_0 = \alpha$$

and, since S is assumed nonempty, pick b_0 to be any element of S: we have our initial interval

$$I_0 = [a_0, b_0]$$

with

$$a_0 \leq S \text{ and } b_0 \in S.$$

Now consider the midpoint of I_0,

$$c_0 = \frac{1}{2}(a_0 + b_0).$$

If c_0 is a lower bound for S, we take

$$a_1 = c_0, \quad b_1 = b_0$$

so that $I_1 = [a_1, b_1]$ has

$$a_1 \leq S \text{ and } b_1 \in S.$$

If c_0 is *not* a lower bound for S, there must be some element of S, which we call b_1, with $b_1 < c_0$. Since a_0 *is* a lower bound for S, we must have

$$a_0 < b_1 < c_0.$$

In this case, we set $a_1 = a_0$ and form

$$I_1 = [a_1, b_1]$$

[23] In fact, the statement of Theorem 2.5.6 is often used as a formulation of the completeness axiom: it is possible to show that our statement of the completeness axiom follows from this statement: see Exercise 8.

again we can guarantee

$$a_1 \leq S, \quad b_1 \in S.$$

Note also that

$$\|I_1\| = b_1 - a_1 \leq c_0 - a_0 = \frac{1}{2}\|I_0\|$$

Proceeding recursively, we obtain a nested sequence of intervals

$$I_0 \supset I_1 \supset I_2 \supset \dots$$

where $I_k = [a_k, b_k]$ with

$$\|I_k\| = b_k - a_k \leq \frac{1}{2}(b_{k-1} - a_{k-1}) \leq \dots \leq \frac{1}{2^k}(a_0 - b_0) \to 0.$$

As before, we see that the two sequences $\{a_k\}$ and $\{b_k\}$ are monotone and bounded, and converge to the same point,

$$\gamma = \lim a_k = \lim b_k.$$

Since $a_k \leq S$ for each k, it follows from Lemma 2.2.9 that

$$\gamma \leq S.$$

However, since $\{b_k\}$ constitutes a sequence *in* S converging to γ, it follows by Lemma 2.5.5 that

$$\gamma = \inf S$$

as required. □

Finally, we note that if a set is *not* bounded above (*resp.* below), then the supremum (*resp.* infimum) as defined here does not exist. However, it is standard practice, similar to writing $\lim a_k = \infty$ when a sequence diverges to infinity, to indulge in the abuse of notation

$$\sup S = \infty \qquad \text{when } S \text{ has no upper bound}$$

and

$$\inf S = -\infty \qquad \text{when } S \text{ has no lower bound.}$$

Exercises for § 2.5

Answers to Exercises 1acegikm, 2ace, 7a(i,iii,v) are given in Appendix B.

Practice problems:

1. For each set S below, find each of the following or else show that it does not exist: (i) $\sup S$; (ii) $\inf S$; (iii) $\max S$; (iv) $\min S$.

   (a) $S = \{1, 2, -\frac{1}{2}, -3, \frac{2}{3}, \frac{1}{3}\}$

   (b) $S = \{\sqrt{3} - \sqrt{2}, \sqrt{13} - \sqrt{12}, \sqrt{3}, \sqrt{2}, \sqrt{12}, \sqrt{13}\}$

   (c) $S = (-1, 1)$

   (d) $S = (-1, 1]$

   (e) $S = \{x \mid 0 \leq x \leq 1 \text{ and } x \text{ rational}\}$

   (f) $S = \{x \mid 0 \leq x \leq 1 \text{ and } x \text{ irrational}\}$

   (g) $S = \{\frac{p}{q} \mid 0 < p < q\}$

   (h) $S = \{\frac{p}{q} \mid q \neq 0 \text{ and } |p| \leq |q|\}$

   (i) $S = \{\frac{p}{q} \mid p^2 > q^2 > 0\}$

   (j) $S = \{\frac{p}{q} \mid p, q > 0 \text{ and } p^2 < 3q^2\}$

   (k) $S = \{x > 0 \mid \sin \frac{1}{x} = 0\}$

   (l) $S = \{x \mid \tan x = 0\}$

   (m) $S = \{x \mid |x - 1| < \left|\frac{1}{x} - 1\right|\}$

   (n) $S = \{x \mid x > \frac{1}{x}\}$

2. For each equation below, find an explicit upper bound and an explicit lower bound for the set of solutions of the equation, *if one exists*. Justify your answers. Note: (i) You don't need to give the *least* upper or *greatest* lower bound, just some bound which you can prove works; (ii) You may assume without proof that each equation has at least one solution.

   (a) $x^3 - 100x^2 + 35x - 2 = 0$

   (b) $x^5 - 4x^3 + 5 = 0$

   (c) $x = \cos x$

(d) $x = a \sin x$ (your answer will depend on a)

(e) $e^x = \cos x$

(f) $x \cos x = \sin x$

Theory problems:

3. Suppose $a < b$ are real numbers. Show that if S is any one of the intervals (a, b), $[a, b]$, $(a, b]$, $[a, b)$, then

$$a = \inf S \text{ and } b = \sup S.$$

(*Hint:* (1) show that a (*resp.* b) is a lower (*resp.* upper) bound for S, and then (2) show that no *higher* (*resp. lower*) number can be a lower (*resp.* upper) bound for S.)

4. Show that the interval $S = (0, 1)$ is bounded above, but has no maximum. (*Hint:* Mimic the strategy for the lower bound on page 73.)

5. Let $S \subset \mathbb{R}$ be any set of real numbers which is bounded below.

   (a) Show that if S has a minimum, say $x = \min S$, then $x = \inf S$.

   (b) Show that if $\inf S$ belongs to S then it is the minimum of S.

6. Prove Remark 2.5.2.

7. Suppose $I = [a, b]$ is a closed interval with $a \neq b$, and let $A = \{|x| \mid x \in I\}$ be the set of absolute values of points in I.

   (a) For each phenomenon below, either give an example or show that no such example exists:

      i. $\max A = |a|$ and $\min A = |b|$
      ii. $\max A = |b|$ and $\min A = |a|$
      iii. $\max A = |a|$ and $\min A \neq |b|$
      iv. $\max A = |b|$ and $\min A \neq |a|$
      v. $\max A \neq |a|, |b|$

   (b) Show that for any bounded set $S \subset \mathbb{R}$,

$$\sup\{|s| \mid s \in S\} - \inf\{|s| \mid s \in S\} \leq |\sup S - \inf S| . \qquad (2.3)$$

   (*Hint:* Use Exercise 6f in § 1.2.)

Challenge problems:

8. Show that Theorem 2.5.6 implies Axiom 2.3.2; that is, give an argument that starts from the assumption that the statement of Theorem 2.5.6 is true and concludes, without using Axiom 2.3.2 or any of its consequences that every monotone bounded sequence converges. (*Hint:* If $\{x_n\}_{n=1}^{\infty}$ is a monotone bounded sequence, let $S = \{x_n \mid n = 1, ...\}$ be the collection of all numbers that appear in the sequence. Show that S is bounded, and hence by assumption has a supremum and infimum. Now relate $\inf S$ or $\sup S$ to $\lim x_n$.)

9. A sequence $\{x_i\}_{i=1}^{\infty}$ is a **Cauchy sequence** if the points are getting close to *each other*—that is, given $\varepsilon > 0$ there exists a place K such that any two points occuring in the sequence beyond K are at most ε apart:
$$|x_{k_1} - x_{k_2}| < \varepsilon \text{ whenever } k_1, k_2 \geq K.$$
   (a) Show that any convergent sequence must be Cauchy. (*Hint:* use the triangle inequality, and the fact that the terms are eventually all within $\frac{\varepsilon}{2}$ of the limit.)
   (b) Show that any Cauchy sequence is bounded. (*Hint:* use the definition of Cauchy, with $k_1 = K$.)
   (c) Show that a Cauchy sequence cannot have two distinct accumulation points.
   (d) Conclude that every Cauchy sequence is convergent.

   This criterion for convergence was understood by both Cauchy and Bolzano; while Bolzano's publication predates Cauchy's, this like much of his other work was not widely known at the time, so that the influential version was Cauchy's.

10. Use Exercise 9 to prove the **Absolute Convergence Test:** Given any series $\sum_{i=0}^{\infty} x_i$, if the series of absolute values $\sum_{i=0}^{\infty} |x_i|$ converges, then so does the original series. (*Hint:* The sequence of partial sums for the absolute series $\sum_{i=0}^{n} |x_i|$ is Cauchy by Exercise 9a; use the triangle inequality to show that the partial sums for the original series are also a Cauchy sequence, and apply Exercise 9d.)

History note:

11. **Bolzano's Lemma:** In his proof of the Intermediate Value Theorem (see Exercise 14 in § 3.2), a crucial lemma is as follows [20, p. 308], [6, p. 269]:

> If M is a property of real numbers which does not hold for all x, and there exists a number u such that all numbers $x < u$ have property M, then there exists a largest U such that all numbers $x < U$ have property M.

Restate this lemma in the language of this section. (*Hint:* If S is the collection of numbers with the property M, then the set of numbers for which the property does *not* hold is its complement, S^c.) Bolzano's proof of the lemma was based on creating a Cauchy sequence (see Exercise 9) converging to the desired number U via the Bisection Algorithm (see Proposition 2.6.3).

2.6 The Bisection Algorithm (Optional)

In this section we consider some consequences of the Completeness Axiom (Axiom 2.3.2) which will prove useful in locating numbers specified by certain properties.
The first verifies a statement that most of us would readily accept on on intuitive basis. One can think of this as another way of saying that the number line has no "holes".

Lemma 2.6.1 (Nested Interval Property)**.** *Suppose we have a family of closed intervals*

$$I_j = [a_j, b_j]\,, \quad j = 0, 1, 2, ...$$

which are nested:

$$I_0 \supset I_1 \supset I_2 \supset I_3 \supset \ldots.$$

Then there is at least one number that belongs to ALL *of the intervals—in other words, their intersection is nonempty:*

$$\bigcap_{j=0}^{\infty} I_j \neq \emptyset.$$

Proof. We need to produce some point belonging to *all* the intervals I_j; we will do this by finding a convergent sequence whose limit lies in $\bigcap_{j=0}^{\infty} I_j$.

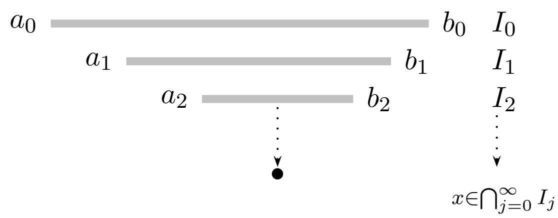


Figure 2.9: Nested Interval Property

The nested property

$$I_k \supset I_{k+1} \quad \text{for all } k$$

means in particular that

$$a_k \leq a_{k+1} \leq b_{k+1} \leq b_k,$$

so that the *left* endpoints of the I_k form a *non-decreasing* sequence

$$a_k \uparrow .$$

Also, the *right* endpoint of any one of these intervals (for example, b_0) is an upper bound for the sequence $\{a_k\}$:

$$a_k \leq b_k \leq b_0 \quad \text{for all } k.$$

It follows that the a_k converge.
Let

$$a = \lim a_k.$$

Pick an interval I_j, (that is, pick $j \in \{0, 1, 2, ...\}$); eventually, the sequence $\{a_k\}$ is trapped in this interval:

$$a_j \leq a_k \leq b_j, \quad \text{or } a_k \in I_j, \text{ for } k \geq j$$

from which it follows by Lemma 2.2.9 that

$$a = \lim a_k \in I_j.$$

Since we have shown $a \in I_j$ for *all* j, we have

$$a \in \bigcap_{j=0}^{\infty} I_j$$

as required. □

When the nested intervals in Lemma 2.6.1 also have length shrinking to zero, we get a powerful tool for specifiying a real number.

Corollary 2.6.2. *Suppose we have a family of closed intervals*

$$I_j = [a_j, b_j], \quad j = 0, 1, 2, ...$$

which are nested and also have length shrinking to zero:

$$\|I_j\| \to 0.$$

Then their intersection consists of precisely one point:

$$\bigcap_{j=0}^{\infty} I_j = \{x_0\} \text{ for some } x_0 \in \mathbb{R}$$

and in fact x_0 is the common limit of the sequence of left (resp. right) endpoints of the intervals I_j:

$$x_0 = \lim a_j = \lim b_j.$$

Proof. We know that the intersection has *at least* one point; in fact, we saw in the proof of Lemma 2.6.1 that the *left* endpoints $\{a_j\}$ of I_j converge to a point $a \in \bigcap_{j=0}^{\infty} I_j$; a similar argument (Exercise 2) shows that the *right* endpoints $\{b_j\}$ also converge to a point $b \in \bigcap_{j=0}^{\infty} I_j$.
We need to show that the intersection cannot contain two *distinct* points, say $x, x' \in \bigcap_0^{\infty} I_j$. We do this by contradiction: if $x \neq x'$, then their distance is positive: $|x - x'| > 0$. However, we have assumed that $\|I_j\| \to 0$, so for j sufficiently large, $\|I_j\| < |x - x'|$. But then x and x' can't *both* belong to I_j; in particular, they can't both belong to the intersection of *all* the I_j's. This shows that there is a unique point x_0 in the intersection. But then in particular, we must have $a = b = x_0$, giving the limit statement. □

A special case of this situation is the **bisection algorithm**, which is a technique for (abstractly) finding a real number by successive subdivision of an interval. Start with a closed interval $I_0 = [a_0, b_0]$; let $c_0 := \frac{1}{2}(a_0 + b_0)$ be its midpoint. Now pick I_1 to be one of the two halves into which c_0 divides I_0: that is, choose either $I_1 = [a_0, c_0]$ or $I_1 = [c_0, b_0]$. In either case, set a_1 (*resp.* b_1) the left (*resp.* right) endpoint, so $I_1 = [a_1, b_1]$. Now repeat the game, using the midpoint c_1 of I_1 to divide *it* into two equal closed intervals, and pick one of them to be I_2, and so on. The results above imply that no matter how we pick the next "half" at each stage, the intersection of all the intervals will specify a unique point $x_0 \in \mathbb{R}$.

Proposition 2.6.3 (Bisection Algorithm)**.** *Suppose $I_j = [a_j, b_j]$, $j = 0, ...,$ is a sequence of closed intervals such that for each j, I_{j+1} is one of the "halves" of I_j (that is, one of the endpoints of I_{j+1} is an endpoint of I_j, and the other is the midpoint c_j of I_j).*
Then their intersection is a single point

$$\bigcap_{j=0}^{\infty} I_j = \{x_0\}$$

and in fact

$$\lim a_j = \lim c_j = \lim b_j.$$

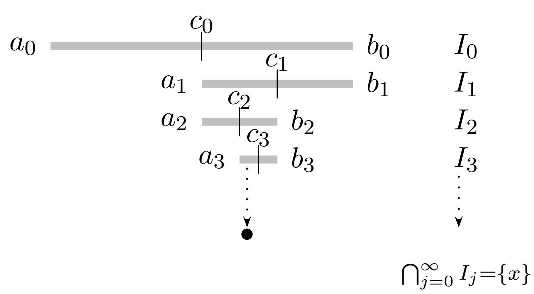


Figure 2.10: Bisection Algorithm

Proof. The basic observation is that the lengths of the intervals are going to zero:

$$\|I_{j+1}\| = \frac{1}{2}\|I_j\|$$

so that

$$\|I_j\| = 2^{-j}\|I_0\| \to 0$$

by Lemma 2.4.3. Then everything follows from Corollary 2.6.2, except the fact that the midpoints also converge to x_0; but this follows from the Squeeze Theorem (Theorem 2.4.7), since

$$a_j \leq c_j \leq b_j, \ j = 0, ...$$

and we already know that the outside sequences have the same limit. □

Exercises for § 2.6

Answers to Exercises 1ac are given in Appendix B.

Practice problems:

1. (a) Give an example of a family of nested closed intervals whose intersection contains *more than one* point.

   (b) Show that the intersection of the nested open intervals $I'_j := \left(0, \frac{1}{j}\right)$ has *no* points.

   (c) Can the intersection of a family of nested intervals contain *exactly two* points?

Theory problems:

2. Modify the argument in the proof of Lemma 2.6.1 to show that if $I_j = [a_j, b_j]$ are nested intervals then the sequence $\{b_j\}$ of *right* endpoints also converges to a point in the intersection $\bigcap I_j$.

3. Suppose $I_j = [a_j, b_j]$ is a family of nested closed intervals with length going to zero, as in Corollary 2.6.2. Use the fact that $\lim a_j = \lim b_j = x_0$ to show that *any* sequence $\{x_j\}_{j=0}^{\infty}$ with $x_j \in I_j$ for $j = 0, ...$ converges to the unique point $x_0 = \bigcap_{j=0}^{\infty} I_j$

4. Use Corollary 2.6.2 to prove the Alternating Series Test (Proposition 2.4.8).

Challenge problem:

5. Fix $x \in [0, 1]$. In the bisection algorithm (Proposition 2.6.3) , choose I_{j+1} so as to contain x at each stage.

   Define $d_j \in \{0, 1\}$ for $j = 1, ...$ by:

$$d_j = \begin{cases} 0 & \text{if } a_j = a_{j-1} \text{ (so } b_j = c_{j-1}), \\ 1 & \text{if } b_j = b_{j-1} \text{ (so } a_j = c_{j-1}). \end{cases}$$

   Show that $\sum_{j=1}^{\infty} \frac{d_j}{2^j}$ is a binary expansion for x.

I don't here consider Mathematical Quantities as composed of Parts extreamly small, but as generated by a continual motion.

Isaac Newton, *Quadrature of Curves*
in John Harris, *Lexicon Technicum* (1710) [55, vol. 1, p. 141]

A function of a variable quantity is an analytic expression composed in any way whatsoever of the variable quantity and numbers or constant quantities.

Leonhard Euler
Introductio in Analysin Infinitorum (1748) [22, p. 3]

3 Continuity

The modern notion of a continuous function is the result of a long evolution of ideas. Leibniz was the first to use the word "function" in the way we do today; of course interrelated quantities were nothing new even to the Greeks. But the idea of looking at the nature of the interrelationship itself can perhaps be traced back to the Merton scholars and Nicole Oresme in the mid-1300's. Nicole Oresme (1323-1382) was the first to formulate the idea of plotting a graph: for example, he would represent the position of a moving body at different times by thinking of the time in a "representative interval", which was horizontal, and placing a vertical line segment whose length corresponded to the displacement over each point of the interval. However, for him the object of interest was the geometric figure formed by the curves so defined, in effect the area under the graph. Three hundred years later, René Descartes (1596-1650) and Pierre de Fermat (1601-1665) rediscovered these ideas, but used them in a different way, stressing the two-way connection between a "locus" (curve) in the plane and a formula (in their case, usually an equation involving powers of x and y) [9].
For Leibniz, a function was what many beginning students think of as a function today: it was given by a formula, which could be a finite or infinite expression. This point of view persisted and in fact was further developed in the eighteenth century; Euler tacitly assumed that all functions are given by power series, perhaps in pieces, and built very powerful analytic tools on this basis; in the late eighteenth century, Joseph Louis Lagrange

(1736-1813) tried to use this idea as the basis for a rigorous development of calculus in his *Théorie des Fonctions Analytiques* of 1797 [37].

Curiously, it was Euler who laid the seeds of the replacement of this view. A topic of central interest in this period was the wave equation, which describes the vibration of a string; part of the mathematical construction is to give *initial conditions*—that is, to specify the shape of the string at the outset. Euler suggested that *any* shape that one could draw freehand should be admitted.

The major instigators of a serious study of continuity were Bernhard Bolzano (1781-1848) and Augustin-Louis Cauchy (1789-1857). Bolzano gave a proof of the Intermediate Value Theorem (Theorem 3.2.1) in a privately published pamphlet in 1817 [6]; this had little impact, as it was not widely disseminated. Cauchy wrote a series of papers and lectures, notably his *Cours d'analyse* of 1821, which were highly influential and set forth careful definitions of the concepts of limit, continuity, derivative and integral [26]. One impetus for this careful re-evaluation was the work of Jean le Rond D'Alembert (1717-1783) and Jean-Baptiste-Joseph de Fourier (1768-1830) on solutions of the wave and heat equations. d'Alembert showed that "any" function of the form $y(x,t) = \frac{1}{2}f(x+at) + \frac{1}{2}f(x-at)$, where t is time and x represents the position along the string (and a is a constant) would be a solution to the wave equation—that is, the shape of the string at time t would be the graph of y versus x. Fourier developed a representation of solutions to the heat equation (which describes the distribution of temperature in a body over time) in infinite series of trigonometric functions (now known as **Fourier series**). In the first case, it became important to understand what "any" function could mean, and in the latter convergence questions came to the fore.

The study of these questions throughout the nineteenth century, notably by Johann Peter Gustav Lejeune Dirichlet (1805-1859) and Bernhard Georg Friedrich Riemann (1826-1866) led to both a more careful and a more general notion of functions, continuity and integrals. The ultimate step in abstracting these ideas is associated to Karl Theodor Wilhelm Weierstrass (1815-1897) and his colleagues. These later investigators produced many counter-intuitive examples which forced a rethinking of the concepts of the calculus which relied less on geometric intuition than had been common in earlier times. This process of abstraction and refinement is sometimes referred to as the "arithmetization of analysis".

3.1 Continuous Functions

Basically, a "function" is a *calculation*: given a number (often denoted x) we perform some calculation that yields a new number (often denoted $f(x)$). For example

$$f(x) = x^2 + 1$$

denotes a function which accepts as input value a real number, then squares it and adds 1 to obtain the output value; some specific values are

$$\begin{aligned} f(1) &= 1^2 + 1 &&= 2 \\ f(2) &= 2^2 + 1 &&= 5 \\ f(\sqrt{2}) &= (\sqrt{2})^2 + 1 = 3. \end{aligned}$$

The notion of "calculation", however, is not limited to what can be done with addition, multiplication, or any other arithmetic or algebraic operations. Any process which tells us how to determine an output value $f(x)$ from an input value x should be acceptable. When we broaden our idea of calculation in this way, we face two issues, illustrated by the following "definition":

> $f(x)$ is the real number y which satisfies
>
> $$y^2 = x.$$

- The *first* problem is that when $x < 0$, this rule cannot apply, since there are *no* numbers of the kind specified by our definition. We need to identify the **domain** of the function, the set of input numbers to which our calculation (or process) can be applied. In the case of our definition above, the domain is

  $$\operatorname{dom}(f) = \{x \mid x \geq 0\} = [0, \infty)\,.$$

- The *second* problem is that, even when the rule can be applied to a given input value of x, there might be some ambiguity about which output value is specified. In our example, the input value $x = 4$ leads to the condition

  $$y^2 = 4$$

  which is satisfied by both $y = -2$ and $y = 2$. In a well-defined function, the output value $y = f(x)$ associated to any particular input value x (from the domain) must be specified *unambiguously*. In the case of the square root functions, this can be accomplished by specifiying the sign of y.

When both of these features are present, we have a **well-defined function**: for example

For $x \geq 0$, $y = f(x)$ is the non-negative solution of the equation

$$y^2 = x.$$

defines the standard square-root function

$$f(x) = \sqrt{x}.$$

We incorporate these two considerations in a formal definition.

Definition 3.1.1. *A **function** with **domain** $D \subset \mathbb{R}$ is a rule which, in an unambiguous way, assigns to each number $x \in D$ in its domain another number $f(x)$.*

The (output) number $f(x)$ (pronounced "f of x") which the function[1] f assigns to the (input) number x is often called the **value** of f at x. The domain of a function may be given explicitly, or it may be implicit. It is always stated explicitly for a **function defined in pieces**, such as the absolute value, whose definition we recall from § 1.2:

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

(Here the domain is $(-\infty, \infty)$.) Sometimes, though, we just write down a calculation without specifying which input values are acceptable. In this case, it is conventional to assume the **natural domain**, the set of *all* numbers for which the calculation makes sense. For the most part, this involves observing two prohibitions:

- **Thou Shalt Not** divide by zero.
- **Thou Shalt Not** take even-order roots of negative numbers.

[1] When speaking of a function in the abstract, it is customary to use a letter as its name; then its value is denoted by appending a parenthesis enclosing the input. Thus, we speak of the *function* f, whose *value* corresponding to a given input x is $f(x)$. However, it is not uncommon to also use the expression $f(x)$ to specify a function: for example, we might speak of "the function x^2."

For example, the formula

$$f(x) = \frac{x^2}{x^2 - 1}$$

is subject to the first prohibition: do not attempt this calculation at home if $x = \pm 1$, since it involves division by $x^2 - 1 = 0$. The natural domain of this function is

$$\mathrm{dom}(f) = (-\infty, -1) \cup (-1, 1) \cup (1, \infty).$$

However, the formula

$$g(x) = \sqrt{\frac{x^2}{x^2 - 1}}$$

in addition to avoiding division by 0 when $x = \pm 1$, also involves the second prohibition. When is $x^2/(x^2 - 1)$ non-negative? Since the numerator x^2 is never negative, the sign of the fraction is determined by the sign of the denominator, which is negative if $|x| < 1$. Thus, we need to avoid all values of x for which $x^2 - 1 \leq 0$ and $x^2 > 0$—which means we avoid all of the closed interval $[-1, 1]$ (*except* $x = 0$, where clearly $g(0) = 0$). Thus, the natural domain here is

$$\mathrm{dom}(g) = (-\infty, -1) \cup \{0\} \cup (1, \infty).$$

The definition we have formulated allows some very wild functions, not necessarily given by a single formula. Consider for example the following function, first suggested as a "pathological" example by Johann Peter Gustav Lejeune Dirichlet (1805-1859) in 1829 (see [27, pp. 94-110]):

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if} x \text{ is irrational.} \end{cases}$$

This is well-defined: for any particular number x, we need only decide whether x is rational or not: only one of these possibilities holds, so there is no ambiguity. Of course, this is only true in principle: in practice, the decision may be difficult to make. While we know that

$$f(0) = 1$$
$$f(1) = 1$$
$$f(\sqrt{2}) = 0$$

it could take a while [2] to decide $f(x)$ for

$$x = \sqrt{\frac{721}{935}}.$$

Since simply "plugging in" a given input value may not be as easy as it seems, it might be useful to understand what happens if, instead of using the *precise* value $x = p$, we used a reasonable *approximation* to it. For a "reasonable" function f, we would expect that if x is an approximation to p, then $f(x)$ should be an approximation to $f(p)$. We make this precise by means of sequences.

Definition 3.1.2. *Suppose f is a function defined on the interval I. We say that f is* ***continuous on I*** *if, whenever $\{x_n\}$ is a sequence in I which converges to a point*

$$x_0 = \lim x_n$$

with $x_0 \in I$, the corresponding sequence of values of f at x_n converges to the value at the limit:

$$\lim f(x_n) = f(x_0)\,.$$

Basically, a function is continuous if it respects limits—in other words, we can find the limit of the sequence of output values $f(x_n)$ by applying f to the limit of the input sequence $\{x_n\}$.
There are two kinds of functions whose continuity is automatic: **constant functions** like $f(x) := 5$ and the **identity function** $f(x) = x$; it is almost too silly for words to remark that if $f(x) = 5$ (*resp.* $f(x) = x$) for all x and we have a sequence $x_n \to x_0$ then $\lim f(x_n) = 5 = f(x_0)$ (*resp.* $\lim f(x_n) = \lim x_n = x_0 = f(x_0)$). Our earlier results on the arithmetic of limits for *sequences* (Theorem 2.4.1) give us basic ways of building new continuous functions from these and other old ones.

Theorem 3.1.3 (Arithmetic and Continuous Functions)**.** *Suppose the functions $f(x)$ and $g(x)$ are both continuous on the interval I. Then so are the functions*

1. $(f + g)(x) := f(x) + g(x)$
2. $(f - g)(x) := f(x) - g(x)$
3. $(fg)(x) := f(x)\,g(x)$

[2] and if you think *this* is easy, replace the fraction with the ratio of two randomly chosen odd numbers with 25 digits each!

4. $\left(\frac{f}{g}\right)(x) := \frac{f(x)}{g(x)}$, *provided* $g(x) \neq 0$ *for* $x \in I$
5. $\big(f(x)\big)^n$ *for any integer* n *(if* $n < 0$*, provided* $f(x) \neq 0$ *for* $x \in I$*)*
6. $\sqrt[n]{f(x)}$ *(if* n *is even, provided* $f(x) \geq 0$ *for* $x \in I$*).*

Proof. This is a matter of writing out definitions and plugging into Theorem 2.4.1. We do this for (1) and leave the rest to you.
Suppose $x_0 = \lim x_n$ (all in I). By assumption, f and g are continuous on I so $f(x_0) = \lim f(x_n)$ and $g(x_0) = \lim g(x_n)$. then by Theorem 2.4.1(1) we have

$$\begin{aligned} \lim(f+g)(x_n) &:= \lim[f(x_n) + g(x_n)] \\ &= \lim f(x_n) + \lim g(x_n) \\ &= f(x_0) + g(x_0) \\ &= (f+g)(x_0). \end{aligned}$$

□

Starting from constants and the identity, we can use Theorem 3.1.3 to establish continuity of many "everyday" functions.

Definition 3.1.4. 1. *A* ***polynomial*** *in* x *is any function which can be obtained from constants and the identity function by means of addition, subtraction and multiplication. Any polynomial can be expressed in the form*

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0.$$

The numbers $a_0, \ldots, a_n$ *are the* ***coefficients*** *of* P*, and the highest power* x^n *of* x *appearing in* $P(x)$ *(with nonzero coefficient) is the* ***degree*** *of* P*.*

2. *A* ***rational function*** *is one which can be obtained from constants and the identity function using division as well as addition, multiplication and subtraction*[3]*. Any rational function can be expressed as a fraction whose numerator and denominator are both polynomials in* x*:*

$$f(x) := \frac{P(x)}{Q(x)}, \quad P,\ Q \text{ polynomials.}$$

[3] Or, as the Mock Turtle said to Alice, "Ambition, Distraction, Uglification, and Derision."[17]

Sometimes the maximum[4] *of the degrees of P and Q (i.e., the highest power of x appearing anywhere in the expression above) is called the* ***degree*** *of $f(x)$.*

3. *More generally, any function which can be obtained from constants and the identity function by any algebraic operations, including powers and roots in addition to the arithmetic operations is an* ***algebraic function****.*

Examples of algebraic functions include the polynomials, like

$$3x^2 - 2x + 5, \quad x^3 - \pi x + \sqrt{2}$$

the rational functions, like

$$\frac{x+1}{x^2+1}, \quad x - \frac{1}{x}, \quad 3x^{-5} + 2x^{-3} + 5x + 4$$

and other functions like

$$\sqrt{\frac{x+1}{x^2+1}}, \quad \sqrt[3]{x^2 + \sqrt{x}}, \quad x^{2/3} - 5x^{-1/5}.$$

An immediate corollary of Theorem 3.1.3 is

Corollary 3.1.5 (Continuity of Algebraic Functions)**.** *Any algebraic function is continuous on any interval where it is defined.*

Another class of "everyday" functions are the trigonometric functions. You are probably used to taking their continuity for granted, but it is best to establish it carefully.

First, let us recall the geometric definition of the sine and cosine functions[5]. Picture a line segment of length 1, with one end at the origin, rotated counterclockwise through an angle of θ **radians** (which is to say, the moving end of the segment travels counterclockwise a distance of θ units of length along the circumference of the circle of radius 1 centered at the origin). Then by definition, the cartesian coordinates of the endpoint are $(\cos\theta, \sin\theta)$.

We wish to show that these functions are continuous, which means that for any convergent sequence $\theta_n \to \theta_0$, we have $\sin\theta_n \to \sin\theta_0$ and

[4] This usage is common, but not universal. The degree sometimes means the degree of the numerator minus the degree of the denominator.

[5] This definition originated with Euler in 1748. See Exercise 14 for what the definition was before then.

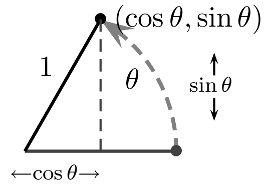


Figure 3.1: Definition of $\cos\theta$ and $\sin\theta$

$\cos\theta_n \to \cos\theta_0$. It will turn out that the basic case we need to show is when $\theta_0 = 0$: that $\theta_n \to 0$ implies $\sin\theta_n \to 0$ and $\cos\theta_n \to 1$. Again, this may seem geometrically obvious, but let us formulate this geometric intuition as a careful proof. The basis of our argument is a pair of inequalities.

Lemma 3.1.6. *For any angle (in radians) satisfying* $|\theta| < \frac{\pi}{2}$,

$$|\sin\theta| \le |\theta| \tag{3.1}$$

$$0 \le 1 - \cos\theta \le |\theta|. \tag{3.2}$$

Proof. We begin with the case $0 < \theta < \frac{\pi}{2}$. Form a triangle $\triangle OAB$ (see Figure 3.2) using the line segment from the origin (O with coordinates

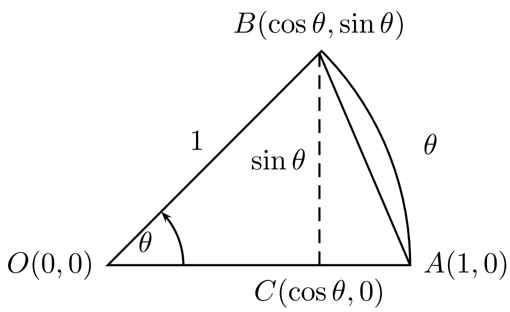


Figure 3.2: Lemma 3.1.6

$(0, 0)$) one unit long along the positive x-axis to A (coordinates $(1, 0)$), and the line segment OB obtained by rotating OA counterclockwise θ radians, so that B has coordinates $(\cos\theta, \sin\theta)$ (as per the definition above). We also drop a vertical segment from B to the x-axis, hitting it at C (coordinates $(\cos\theta, 0)$).
Now compare the following distances:

- the *arc* $\overset{\frown}{AB}$ of the circle joining A to B has length θ units (because of the definition of radians) and is longer than the *straight-line* distance $\overline{AB}$ between these points:
$$\overline{AB} \leq \theta$$
- the line segment $\overline{AB}$ is the hypotenuse of the right triangle $\triangle ABC$, one of whose legs is $\overline{BC}$, of length $\sin\theta$; since the hypotenuse is longer than either leg,
$$\sin\theta \leq \overline{AB}.$$

Combining these (and noting that all quantities are positive), we have

$$0 \leq \sin\theta \leq \theta \quad \text{for } 0 < \theta < \frac{\pi}{2}. \tag{3.3}$$

To obtain a corresponding inequality for $\cos\theta$, replace the leg $\overline{BC}$ with $\overline{AC}$, which has length $1 - \cos\theta$ (why?) in the argument above to obtain the inequality

$$0 \leq 1 - \cos\theta \leq \theta \quad \text{for } 0 < \theta < \frac{\pi}{2}. \tag{3.4}$$

Now when θ is nonpositive (but still between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$), we use the fact that $\sin\theta$ (*resp.* $\cos\theta$) is an *odd* (*resp. even*) function to express these quantities in terms of a positive angle. This, together with the fact that $\cos\theta \leq 1$ for *all* θ gives the desired inequality; we leave the details to you (Exercise 10). □

The inequalities (3.1),(3.2) easily give us the convergence statements we want when the angle goes to zero; to take care of other limit angles we recall the angle-summation formulas

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta \tag{3.5}$$
$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta. \tag{3.6}$$

(See Exercise 16 in § 4.2 for a proof.)

Proposition 3.1.7. *The functions* $\sin\theta$ *and* $\cos\theta$ *are continuous everywhere.*

Proof. Suppose we have a convergent sequence of angles $\theta_n \to \theta_0$; we need to show that $\sin\theta_n \to \sin\theta_0$ and $\cos\theta_n \to \cos\theta_0$. We will concentrate on the first statement and leave the second to you.

It will be convenient to rewrite the sequence of angles in terms of their difference from θ_0:

$$\theta_n = \theta_0 + t_n;$$

then the assumption $\theta_n \to \theta_0$ translates to

$$t_n \to 0.$$

To show that $\sin\theta_n \to \sin\theta_0$, we use the angle-sum formulas (3.5),(3.6) to write

$$\sin\theta_n = \sin(\theta_0 + t_n) = \sin\theta_0 \cos t_n + \cos\theta_0 \sin t_n$$

and analyze each term separately.
In the first term, the factor $\sin\theta_0$ is constant, so we focus on the second factor. From (3.2) (and since eventually $|t_n| < \frac{\pi}{2}$) we conclude that

$$0 \le 1 - \cos t_n \le |t_n|$$

and since $t_n \to 0$ the Squeeze Theorem (Theorem 2.4.7) implies that $1 - \cos t_n \to 0$ or

$$\cos t_n \to 1.$$

It follows that

$$\lim(\sin\theta_0 \cos t_n) = \sin\theta_0.$$

Similarly in the second term, the first factor is constant, and the second satisfies

$$|\sin t_n| \le |t_n|$$

which together with $t_n \to 0$ implies that

$$\sin t_n \to 0$$

so that (using Theorem 3.1.3(3))

$$\lim(\cos\theta_0 \sin t_n) = (\cos\theta_0) \cdot 0 = 0.$$

Finally, combining these two limits using Theorem 3.1.3(1) we have

$$\begin{aligned} \lim\theta_n &= \lim(\sin\theta_0 \cos t_n + \cos\theta_n \sin t_n) \\ &= \sin\theta_0 + 0 \\ &= \sin\theta_0. \end{aligned}$$

The proof of continuity for $\cos\theta$ is entirely analogous, using parts of the above together with the second angle-sum formula: the details are left to you (Exercise 11). □

Recall that the rest of the trigonometric functions are defined in terms of the sine and cosine using ratios:

$$\tan\theta := \frac{\sin\theta}{\cos\theta} \quad \sec\theta := \frac{1}{\cos\theta}$$
$$\cot\theta := \frac{\cos\theta}{\sin\theta} \quad \csc\theta := \frac{1}{\sin\theta}.$$

Note that $\tan\theta$ and $\sec\theta$ are undefined when $\cos\theta = 0$ (*i.e.*, for θ an *odd* multiple of $\frac{\pi}{2}$) while $\cot\theta$ and $\csc\theta$ are undefined when $\sin\theta = 0$ (*i.e.*, for θ an *even* multiple of $\frac{\pi}{2}$). Otherwise, they are automatically continuous by Theorem 3.1.3(4) and Proposition 3.1.7:

Corollary 3.1.8 (Continuity of Trig Functions)**.** *The trigonometric functions are continuous on any interval where they are defined: for the tangent and secant (resp. cotangent and cosecant) functions, the interval must not include any* odd *(resp.* even*) multiple of* $\frac{\pi}{2}$.

The graphs of the four basic trig functions are shown in figures Figure 3.3 and Figure 3.4.

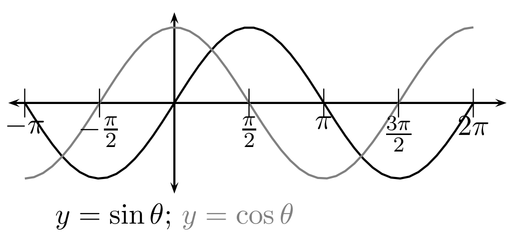


Figure 3.3: Graphs of $\cos\theta$ and $\sin\theta$

Another very important way to build new functions from old ones is via composition. The notion of a *composition* of functions is just that of following one computation with another. Here is a general definition.

Definition 3.1.9. *Suppose f and g are functions. The* ***composition of*** g ***with*** f *is the function*

$$g \circ f$$

defined by

$$(g \circ f)(x) = g(f(x)).$$

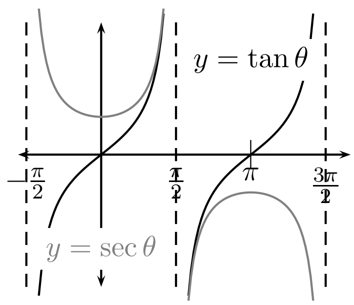


Figure 3.4: Graphs of $\tan\theta$ and $\sec\theta$

Notice that, in order to calculate $(g\circ f)(x)$, we need to have $x \in \text{dom}(f)$ *and* $f(x) \in \text{dom}(g)$.
Consider, for example, f and g defined by

$$f(x) = x^2 + 1, \quad g(x) = \sqrt{x}.$$

Then

$$(g\circ f)(x) = g(f(x)) = g\big(x^2 + 1\big) = \sqrt{x^2 + 1}$$

while

$$(f\circ g)(x) = f(g(x)) = f\big(\sqrt{x}\big) = (\sqrt{x})^2 + 1 = x + 1.$$

However, in the latter case the calculation of $(f\circ g)(x)$ *only* makes sense if $x \in \text{dom}(g)$, which is to say $x \geq 0$; thus the domain of $f\circ g$ is $[0, \infty)$, not $(-\infty, \infty)$, even though the formula $x + 1$ makes sense for *all* x. Notice in particular that the *order* of composition *dramatically* affects the outcome. Fortunately, composition respects continuity.

Theorem 3.1.10 (Composition of Continuous Functions)**.** *Suppose*

- *f is continuous on the interval J*
- *g is continuous on the interval I*

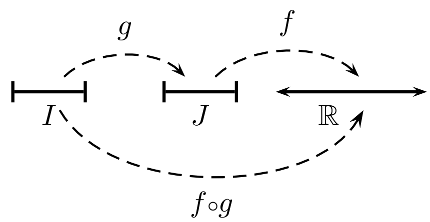


Figure 3.5: Composition

- *For every x in I, $g(x)$ is in J.*

Then the composition

$$(f\circ g)(x) := f(g(x))$$

is continuous on I.

Proof. Suppose $\{x_n\}$ is a convergent sequence in I whose limit $x_0 := \lim x_n$ also belongs to I. Denote the corresponding g-images by $y_n := g(x_n)$ (including, when $n = 0$, $y_0 = g(x_0)$). By assumption, each y_n belongs to J. Furthermore, the continuity of g says that

$$\lim y_n := \lim g(x_n) = g(x_0) := y_0$$

Now, continuity of f on J insures that

$$\lim f(y_n) = f(y_0)\,.$$

But by definition, this is the same as

$$\lim(f\circ g)(x_n) := \lim f(g(x_n)) = f(g(x_0)) = (f\circ g)(x_0).$$

□

This result, combined with Theorem 3.1.3 and Corollary 3.1.8, gives us many further examples of continuous functions, for example

$$\sin(x^2+1), \quad \sqrt{1+\cos^2 x},$$
$$\tan\sqrt{\frac{x+1}{x^2+1}}, \quad \frac{\sin x + \cos 2x}{\tan(3x+2)}.$$

Exercises for § 3.1

Answers to Exercises 1-4ace, 5ac, 7ac, 8ace,12b are given in Appendix B.

Practice problems:

1. What is the natural domain of each function below?

   (a) $f(x) = \dfrac{x^2-4}{x^2-9}$ (b) $f(x) = \dfrac{x^2-9}{x^2-4}$ (c) $f(x) = \sqrt{\dfrac{x^2-4}{x^2-9}}$

   (d) $f(x) = \sqrt{\cos x}$ (e) $f(x) = \sqrt{\tan x}$ (f) $f(x) = \tan \dfrac{1}{x}$

   (g) $f(x) = \sqrt{1-|x|}$ (h) $f(x) = \sqrt{|x|-1}$

2. In each part below, you are given a verbal description of two inter-related variables; give an explicit definition of a function (*e.g.*, via a formula) expressing the relation as indicated. Be sure to specify the domain of your function.

   (a) Express the circumference C of a circle as a function of its area A.

   (b) Express the area A of a square as a function of its perimeter P.

   (c) Express the height h of an equilateral triangle in terms of the length s of its side.

   (d) Express the area A of an equilateral triangle in terms of the length s of its side.

   (e) Express the volume V of a sphere in terms of its surface area S.

   (f) Express the diagonal d of a cube in terms of the length s of its side.

   (g) The postage rate for first-class flat mail up to 3.5 ounces as of May 14, 2007 is given in Table 3.1. Express the postage P as a function of the weight w in ounces. *Note:* this will be a function defined in pieces.

   (h) The cost of local priority mail for items up to 5 pounds is given in Table 3.2. Express the cost C of sending an item via local priority mail as a function of the weight w in pounds.

3. Let $f(x) = 1 - x^2$. Evaluate each expression below:

   (a) $f(0)$ (b) $f(1)$ (c) $f(-1)$

   (d) $f(x+1)$ (e) $f(2x)$ (f) $f(f(x))$

Table 3.1: First Class (Flat) Mail Rates

weight (oz.)	cost (¢)
1	41
2	58
3	75
3.5	92

Table 3.2: Local Priority Mail Rates

weight (lb.)	cost ($)
1	4.60
2	4.60
3	5.05
4	5.70
5	6.30

4. Given $f(x) = 2x^2 - 1$, $g(x) = |x|$, $h(x) = \sqrt{1-x}$, find each composite function below and specify its domain:

   (a) $f \circ g$ (b) $g \circ f$ (c) $f \circ h$

   (d) $h \circ f$ (e) $h \circ g$ (f) $f \circ h \circ g$

5. For each function defined in pieces below, decide which value(s) (if any) of the parameter α (and β, if indicated) make the function continuous on $(-\infty, \infty)$:

   (a) $f(x) = \begin{cases} x+1 & \text{for } x \leq 1, \\ \alpha - x & \text{for } x > 1 \end{cases}.$

   (b) $f(x) = \begin{cases} (x+\alpha)^2 & \text{for } x < 0, \\ x^2 + \alpha & \text{for } x \geq 0 \end{cases}.$

   (c) $f(x) = \begin{cases} (x-\alpha)^2 & \text{for } x \leq 0, \\ (x+\beta)^2 & \text{for } x > 0 \end{cases}.$

   (d) $f(x) = \begin{cases} (x-\alpha)^2 & \text{for } x \leq -1, \\ (x+\beta)^2 & \text{for } x > -1 \end{cases}.$

6. Let $f(x) = \sin \frac{2}{x}$ for $x \neq 0$.

(a) Compute the limit of the sequence $f(x_n)$, where $x_n = \frac{1}{n\pi}$;

(b) Compute the limit of the sequence $f(y_n)$, where $y_n = \frac{2}{(4n+1)\pi}$;

(c) Compute the limit of the sequence $f(z_n)$, where $z_n = \frac{2}{(4n+3)\pi}$;

(d) Conclude that there is no value that can be assigned to $f(0)$ to make $f(x)$ continuous on $(-\infty, \infty)$.

7. In each part below, you are given a function f and an interval I. Decide whether f is continuous on I and if not, give an explicit sequence $\{x_k\}$ converging in I for which the sequence $\{f(x_k)\}$ does not converge to $f(\lim x_k)$:

(a) $f(x) = \begin{cases} x^2 & \text{if } x < 1, \\ x+1 & \text{if } x \geq 1, \end{cases} \quad I = [0, 2)$

(b) $f(x) = \begin{cases} x \sin \frac{1}{x} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0, \end{cases} \quad I = [-1, 1]$

(c) $f(x) = \begin{cases} 0 & \text{if } x \text{ is rational}, \\ 1 & \text{if } x \text{ is irrational}, \end{cases} \quad I = [-1, 1]$

(d) $f(x) = \begin{cases} \frac{1}{q} & \text{if } 0 \neq x = \frac{p}{q} \text{ in lowest terms}, \\ 0 & \text{if } x = 0 \text{ or } x \text{ is irrational}, \end{cases} \quad I = [-1, 1]$

8. Give an example of each kind of function below, or explain why none exists.

(a) A function with domain $(0, 1)$ whose values include all positive real numbers.

(b) A function with domain $(-\infty, \infty)$ whose values are all between 0 and 1.

(c) A function on $[-1, 1]$ with value -1 at $x = 1$, 1 at $x = -1$, which never takes the value 0.

(d) A polynomial of degree n which never takes the value 0; how does the answer depend on the degree?

(e) A function whose graph is the lower half of the curve $x^2 + y^2 = 1$.

Theory problems:

9. Show that the absolute value function $|x|$ is continuous on $\mathbb{R}$.

10. Complete the proof of Lemma 3.1.6 by showing that when $-\frac{\pi}{2} < \theta < 0$,

$$\theta < \sin\theta < 0$$
$$0 \leq 1 - \cos\theta \leq -\theta.$$

(*Hint:* Replace θ with $-\theta$ in the inequalities Equation (3.3) and Equation (3.4), and use properties of the sine and cosine functions.)

11. Use the angle-sum formula

$$\cos(\theta_0 + t_n) = \cos\theta_0 \cos t_n - \sin\theta_0 \sin t_0$$

to complete the proof of Proposition 3.1.7 by showing that $\cos\theta$ is continuous on $(-\infty, \infty)$.

Challenge problems:

12. Suppose that f is a quadratic function of the form $f(x) = ax^2 + b$ with $a \neq 0$.

    (a) Show that if $g(x) = cx^2 + d$, $c \neq 0$ is another quadratic function *commuting* with f (*i.e.*, $(f \circ g)(x) = (g \circ f)(x)$ for all x), then $f(x) = g(x)$ for all x.

    (b) Suppose instead that $g(x) = Ax^3 + Bx^2 + Cx + D$, where $A \neq 0$. Under what conditions on a, b, A, B, C, D do f and g commute? (*Hint:* Note that $g \circ f$ is an *even* function; this forces certain coefficients to be zero. Then look at the rest.)

    You may use the fact that two polynomials are equal as functions precisely if all corresponding coefficients are equal.

13. In this problem, we will use the continuity of $f(x) = \sin x$ to show that every $L \in [-1, 1]$ is an accumulation point of the sequence $x_k = \sin k$. This shows that the divergence of the sequences $\sin k$ and $\cos k$, shown in Exercise 33 (§ 2.4), is a much wilder phenomenon than we might have imagined.

(a) Let p_k, $k = 0, 1, 2, ...$, denote the point $p_k = (\sin k, \cos k)$ on the unit circle obtained by rotating the point $p_0 = (1, 0)$ counterclockwise by k radians. Note that since π is irrational (which we use without proof), these points are all distinct, and once $k > 2\pi$, they start to intertwine on the circle.

We want to show that *for any point p on the unit circle, there is a sequence of integers $k_i \to \infty$ for which the points p_{k_i} converge to p.*[6]

Suppose that this is not true for some point p, so that some open arc A containing p contains none of the points p_k, $k = 0, 1, ...$ Extend the arc as far as you can in both directions preserving this property; this means that each endpoint of the arc is either itself one of the p_k or is a limit of some sequence of p_k's (from outside the arc).

Now consider all the arcs A_n obtained by rotating our given arc *clockwise* an integer number $n > 0$ of radians. Show that if $p_k \in A_n$ then $p_{k+n} \in A$, contrary to our assumption. Thus, *none* of the arcs A_n can contain *any* of the points p_k, $k = 0, 1, ...$

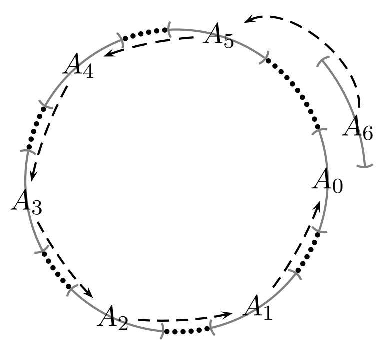


Figure 3.6: Density of $\sin n$

This means that two of these arcs cannot overlap, since then one of them contains an endpoint of the other, and this endpoint is a limit of p_k's. Thus, the arcs A_n are all disjoint.

But this is absurd, since the disjointness of the arcs means that the circumference of the circle must be at least the sum of their

[6] The official terminology for this is: the points p_{k_i} are **dense** in the unit circle.

lengths, if we let $\varepsilon > 0$ denote the length of A, then each of the rotated arcs A_n also has length ε, which means the total of their lengths is infinite. This contradiction proves the italicized statement.

(b) Now, pick any $L \in [-1, 1]$, and let p be the point $(L, \sqrt{1 - L^2})$ on the circle lying $\arcsin L$ radians counterclockwise from (1,0). By the previous argument, there is a sequence of points p_{k_i} converging to p. But $p_{k_i} = (\sin k_i, \cos k_i)$, so we have $\sin k_i \to L$.

History note:

14. The definition of the trigonometric functions on p. 94 originated with Euler in [22, Chap. VIII]. Originally, sines and cosines were defined not for angles, but for arcs in a given circle [20, p. 275]: the sine of an arc was half the length of the chord subtended by an arc of twice the length, and its cosine was the length of the perpendicular from this chord to the center of the circle.

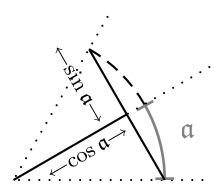


Figure 3.7: Old definition of $\sin \mathfrak{a}$ and $\cos \mathfrak{a}$, for an *arc* $\mathfrak{a}$

(a) Show that the definition gives quantities that depend on the radius of the circle used, but two equal arcs (in the same circle) have the same sine and cosine.

(b) The change in point of view to thinking in terms of the angle is due to George Joachim Rheticus (1514-1576), a student of Copernicus [36, p. 239]. Show that if we measure these lengths using the radius of the circle as the unit of length, the resulting definition agrees with the one on p. 94 for angles of between zero and $\frac{\pi}{2}$ radians.

(c) What happens for angles greater than $\frac{\pi}{2}$?

3.2 The Intermediate Value Theorem and Implicit Functions

In § 2.4, we calculated a decimal expression for $\sqrt{2}$, leaving unjustified the assumption that this number, defined as the positive solution of the equation $x^2 = 2$, actually exists. Of course, we saw later that the decimal expression we obtained describes an actual number, and we could (with a slight effort) show that it does, indeed, solve the equation. However, a more primitive approach, which we also used in an intuitive way, was to notice that since $1^2 < 2 < 2^2$, *some* number between 1 and 2 satisfies the equation $x^2 = 2$. In this section, we establish a wide-ranging generalization of this reasoning, which applies to *any* continuous function.

Theorem 3.2.1 (Intermediate Value Theorem)**.** *Suppose f is continuous on the closed interval $[a, b]$ and C is any number between the endpoint values $f(a)$ and $f(b)$.*
Then there exists at least one point $t \in [a, b]$ with

$$f(t) = C.$$

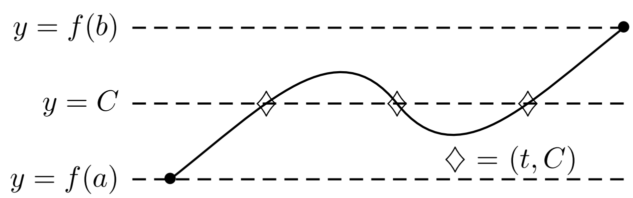


Figure 3.8: Theorem 3.2.1 (Intermediate Value Theorem)

During the eighteenth century, the word "continuous" as a description of a function was associated with being able to express it by a formula, and a function was "discontinuous" at a point where one had to switch formulas, like the point $x = 0$ for the absolute value. In a prize paper written in 1791, Louis François Antoine Arbogast (1753-1803) first made a distinction between this kind of "discontinuity" and the notion of *discontiguity*—failure of the intermediate value property.[7] The earliest careful proof that a continuous function (in our sense) has the intermediate value property was contained in a pamphlet published privately in 1817 by Bernhard Bolzano (1781-1848) [6], who lived and worked in Prague. A

[7] See also the discussion following Corollary 4.9.5 (Darboux's Theorem).

selection from this is reprinted in [5, pp. 15-16] and discussed in some detail in [8, pp.96-101]; a very recent translation of the full pamphlet is given by [46] (see Exercise 14 for a working through of Bolzano's proof of the Intermediate Value Theorem). This paper contained also a formulation of what we call the Completeness Axiom (Axiom 2.3.2), and a version of the Bolzano-Weierstrass Theorem (Proposition 2.3.8). However, Bolzano worked far from the center of scientific discourse, and so his work was not widely known. The other early contributor to this area was Augustin-Louis Cauchy (1789-1857), who in 1821 published his *Cours d'analyse* [12] in which he set down a rigorous foundation for the notions of limits, continuity, and derivatives, and in particular gave his proof of the Intermediate Value Theorem (see Exercise 15).

Proof. Note that the condition that C is between $f(a)$ and $f(b)$ means one of three sets of inequalities:

1. if $f(a) = f(b)$, it means $f(a) = C = f(b)$;
2. if $f(a) < f(b)$, it means $f(a) \leq C \leq f(b)$;
3. if $f(a) > f(b)$, it means $f(a) \geq C \geq f(b)$.

If C equals one of the two numbers $f(a)$ or $f(b)$, there is nothing to prove (why?), so we can ignore the first case and assume strict inequalities in the other two. We shall give the proof for the second case; it is easy to deduce the result in the third case if it is known to always hold for the second (Exercise 10).
We consider the set $S := \{x \in [a, b] \mid f(x) > C\}$; this is nonempty, because it contains $x = b$, and is bounded below by a. Therefore, it has an infimum, by Theorem 2.5.6, and we can set

$$t := \inf S = \{x \in [a, b] \mid f(x) > C\}.$$

By Lemma 2.5.5(3), there exists a sequence $t_i \in S$ converging to t. Since $f(x)$ is continuous, we have $f(t) = \lim f(t_i)$, and since $f(t_i) > C$, Lemma 2.2.9 says that

$$f(t) \geq C.$$

Since $f(a) < C$, the same is true of points slightly to the right of a, and so $a < t$: it follows that there exists a sequence $t_i^* \in [a, t)$ also converging to t. But note that necessarily $f(t_i^*) \leq C$ for all i, and so

$$f(t) = \lim f(t_i^*) \leq C.$$

Since we have shown that $f(t)$ is both at least and at most C, it must equal C and we are done.

$\square$

The Intermediate Value Theorem is an *existence* theorem: it tells us that a given equation has *at least one* solution. In specific cases, there may be more than one. For example, the function

$$f(x) = x^3 - x$$

is continuous on the interval $[-2, 2]$, and

$$f(-2) = -6 < 0 < 6 = f(2)\,.$$

The Intermediate Value Theorem (with $C = 0$) tells us that the equation

$$f(x) = 0$$

has at least one solution in the interval $[-2, 2]$. In fact, it has three: $x = -1$, $x = 0$ and $x = 1$ all satisfy

$$x^3 - x = 0.$$

However, if we also happen to know that f is *strictly monotone* on $[a, b]$, then uniqueness follows as well. We shall use this to justify the definition of several basic functions.

Definition 3.2.2. *Suppose $S \subset \mathbb{R}$ is any subset of the domain of the function f. We say f is*

- ***strictly increasing*** *on S (denoted $f\uparrow$ on S) if for any two points $x, y \in S$,*
$$x < y \text{ guarantees } f(x) < f(y)$$
- ***strictly decreasing*** *on S (denoted $f\downarrow$ on S) if for any two points $x, y \in S$,*
$$x < y \text{ guarantees } f(x) > f(y)$$

f is ***strictly monotone*** *on S if it satisfies one of these conditions on S.*

Remark 3.2.3. *If f is strictly monotone on $S \subset \mathbb{R}$, then for any number C, the equation*

$$f(x) = C$$

cannot have two distinct solutions which both belong to S.

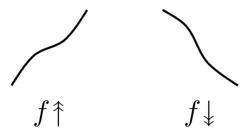


Figure 3.9: Strictly monotone functions

We shall consider several instances of these conditions, exploiting monotonicity together with the Intermediate Value Theorem to justify certain important, but *implicitly defined* functions.

Proposition 3.2.4. *For each positive integer n, the function*

$$g(x) = \sqrt[n]{x}$$

is well-defined for all $x \geq 0$ by

> *$g(x)$ is the number $y \geq 0$ satisfying*

$$y^n = x.$$

Proof. Let $f(x) = x^n$. We know that f is continuous and strictly increasing on $[0, \infty)$ (see Exercise 10 in § 1.2). Given $x > 0$, let $a = 0$ and pick b to be any number which is greater than both x and 1. Since $1 < b$, we have $1 < b^{n-1}$, and multiplying both sides by b,

$$b < b^n.$$

Since $x < b$ by construction, we can conclude

$$f(a) = 0 < x < b^n = f(b).$$

By the Intermediate Value Theorem on $[a, b]$ (with $C = x$), we know there exists at least one number $y \in [a, b]$ with

$$y^n = f(y) = x.$$

Furthermore, since f is strictly monotone, this number is defined unambiguously, and the definition above for

$$g(x) = \sqrt[n]{x}$$

gives a good function. □

Before looking at our final examples, we make some general observations. Suppose $f(x)$ is continuous and strictly increasing on an interval $[a, b]$. Set $A = f(a)$ and $B = f(b)$. The Intermediate Value Theorem guarantees that for each number $C \in [A, B]$ the equation

$$f(y) = C$$

has a solution $y \in [a, b]$; since f is strictly monotone, this solution is unique. In other words, we can always define a function g with domain $[A, B]$ by the rule

$$\text{for } x \in [A, B]\,,\ y = g(x) \Leftrightarrow y \in [a, b] \text{ and } f(y) = x.$$

This function g is called the **inverse** of f. One can check that

$$(f \circ g)(x) = f(g(x)) = x \text{ for all } x \in [A, B]$$
$$(g \circ f)(y) = g(f(y)) = y \text{ for all } y \in [a, b]\,.$$

The inverse of f is often denoted f^{-1}.
There is a general terminology for the properties entering into this reasoning. A function is **one-to-one** on a subset A of its domain if different inputs yield different outputs: that is,

$$x \neq x' \Rightarrow f(x) \neq f\left(x'\right),$$

in other words, for each choice of y the equation

$$y = f(x)$$

has *at most one* solution x in A. Remark 3.2.3 says that any strictly monotone function is one-to-one. (When f is continuous and A is an interval, the converse is also true: a continuous function which is one-to-one on an interval must be strictly monotone on that interval. See Exercise 11.) A function maps a subset A of its domain **onto** a set B if every element of B is the image of some element of A; in other words, for every y in B the equation

$$y = f(x)$$

has *at least one* solution x in A. Theorem 3.2.1 says that a function which is continuous on a closed interval $[a, b]$ maps this interval onto the closed interval with endpoints $f(a)$ and $f(b)$. Sometimes we use sloppy language and simply say that f is "onto" when the identity of the sets A and B is

clear. The argument above showed that a continuous strictly monotone function has both properties, and as a result it has an inverse.[8]
We have already defined $\sqrt[n]{x}$ as the inverse of x^n and found a special argument to show that it is continuous. This is always possible in the general situation sketched above.

Proposition 3.2.5 (Continuous Inverse Function Theorem[9])**.** *Suppose $f(x)$ is continuous and strictly increasing on the closed interval $[a,b]$: then the inverse function*

$$g = f^{-1}$$

can be defined on $[A,B]$, where $A = f(a)$ and $B = f(b)$, by

$$y = g(x) \ \ (x \in [A,B]) \Leftrightarrow y \in [a,b] \ \text{ and } f(y) = x.$$

Under these assumptions, f^{-1} is continuous and strictly increasing on $[A,B]$.

Of course, there is an analogous statement for strictly decreasing continuous functions.

Proof. We need to show that if $x_i \to x$ in $[A,B]$, then $y_i = g(x_i) \to y = g(x)$ in $[a,b]$. There are two parts to this:

1. if x_i converge then $y_i = g(x_i)$ converge
2. if $x_i \to x$ and $y_i = g(x_i) \to y$, then $y = g(x)$.

We prove the second statement first. Suppose $x_i \to x$ in $[A,B]$ and $y_i = g(x_i) \to y$ in $[a,b]$. We need to show that $y = g(x)$. But since $y_i \to y$, we know that $f(y_i) \to f(y)$, and since $f(y_i) = x_i \to x$, this forces $f(y) = x$, which is the same as $y = g(x)$.
Now, to show the first statement, we will use Exercise 30c in § 2.3, which says that if a bounded sequence has precisely one accumulation point, then it converges to that point. Since the sequence $y_i := g(x_i)$ is bounded, it has at least one accumulation point by the Bolzano-Weierstrass Theorem (Proposition 2.3.8). But any accumulation point is the limit of some

[8] There is a parallel second terminology, derived from French, for these properties. A one-to-one function is called an **injection**, and an onto function is called a **surjection**; in this terminology a function which is one-to-one on A and maps A onto B is called a **bijection** between A and B.

[9] I have borrowed the name "Inverse Function Theorem" which usually refers to functions of several variables and involves derivatives, but this can be viewed as a one-variable analogue.

subsequence, say y_{i_k}. The corresponding points x_{i_k} are a subsequence of x_i, so converge to x, and by the preceding paragraph, this implies that $y_{i_k} \to y = g(x)$. This shows that $y = g(x)$ is the *only* accumulation point of the sequence $\{y_i\}$, and thus $y_i \to y$, proving continuity of $g(x)$.
Finally, to show that $g(x)$ is strictly increasing on $[A, B]$, suppose $A \leq x < x' \leq B$ and let $y = g(x)$, $y' = g(x')$. If $y \geq y'$, then since $f(x) \uparrow$ by hypothesis, $x = f(y) \geq f(y') = x'$, contradicting our assumption that $x < x'$. Hence, $y < y'$, proving that $g(x)$ is strictly increasing. □

We shall apply this to define the **inverse trigonometric functions**.
If $f(\theta) = \sin\theta$, we would like to define a function which assigns to a number x the angle θ for which $\sin\theta = x$. Since $|\sin\theta| \leq 1$, we need to require that $-1 \leq x \leq 1$. Furthermore, there are many values of θ that give the same sine. However, they represent only two possible *geometric* angles, corresponding to the two points at the same height (x) on the unit circle. We note that if we stay in the right semicircle, there is no problem.

Definition 3.2.6. *For $x \in [1-, 1]$, the **inverse sine** (or **arcsine**[10]) of x is defined by*

$$\theta = \arcsin x \Leftrightarrow -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2} \text{ and } \sin\theta = x.$$

We note that $\sin\theta$ is strictly increasing and continuous on the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, so $\arcsin x$ is well-defined, strictly increasing and continuous on $[-1, 1]$. Its graph is given in Figure 3.10.

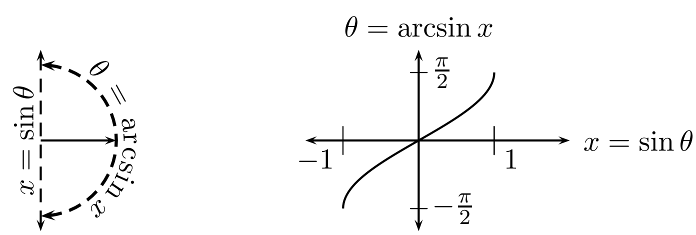


Figure 3.10: $y = \arcsin x$

For the cosine, we need to stay in the *upper* semicircle.

[10]Sometimes the notation $\sin^{-1} x$ is used for the arcsine. However, I have avoided this since it can easily be confused with the cosecant.

Definition 3.2.7. *For $x \in [1-, 1]$, the* ***inverse cosine*** *(or* ***arccosine****) of x is defined by*

$$\theta = \arccos x \Leftrightarrow 0 \le \theta \le \pi \text{ and } \cos\theta = x.$$

Again, $\cos\theta$ is strictly monotone and continuous on $[0, \pi]$, so $\arccos x$ is also well-defined, strictly monotone and continuous on $[-1, 1]$. Its graph is given in Figure 3.11

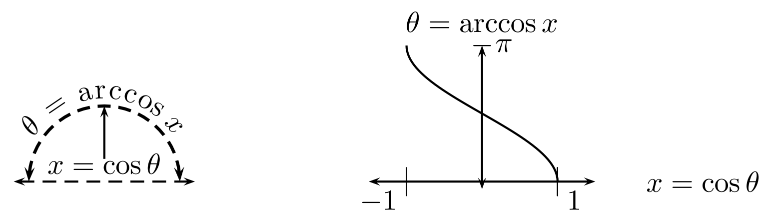


Figure 3.11: $y = \arccos x$

We know that for any angle θ,

$$\sin^2\theta + \cos^2\theta = 1,$$

so that

$$\sin\theta = \pm\sqrt{1-\cos^2\theta}, \quad \cos\theta = \pm\sqrt{1-\sin^2\theta}$$

and vice-versa. However, one can check that if θ equals the arccosine (*resp.* arcsine) of some number, then its sine (*resp.* cosine) is actually nonnegative (Exercise 9) so that in those cases we can ignore the "minus" possibility in these identities.
The arctangent is defined similarly. Note that $\tan\theta$ can take any *real* value (not just values in $[-1, 1]$, as is the case for sine and cosine) and that it blows up at $\theta = \pm\frac{\pi}{2}$. We therefore take the value to lie in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

Definition 3.2.8. *For any $x \in \mathbb{R}$, the* ***arctangent*** *is defined by*

$$\theta = \arctan x \Leftrightarrow x = \tan\theta \text{ and } -\frac{\pi}{2} < \theta < \frac{\pi}{2}.$$

The graph of this function is given in Figure 3.12.

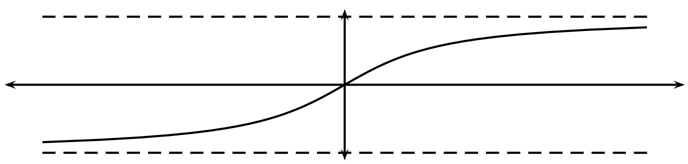


Figure 3.12: $y = \arctan x$

It is also possible to define inverses for the other trigonometric functions, but these are less standard (see Exercise 6 and Exercise 7).

Exercises for § 3.2

Answers to Exercises 1ac, 2ace, 4, 5ace, 6de, 7bc, 8ace, 9bc are given in Appendix B.

Practice problems:

1. For each function below, use the Intermediate Value Theorem to show that the equation $f(x) = 0$ has at least one solution in the indicated interval.

   (a) $f(x) = x^3 - 3x + 1$, $[1, 2]$.
   (b) $f(x) = 10x^3 - 10x + 1$, $[-2, 2]$.
   (c) $f(x) = 2\cos x - 1$, $[0, 2]$.
   (d) $f(x) = \dfrac{x^2 - 12x + 8}{x - 2}$, $[0, 1]$.

2. For each function below, find an integer N so that $f(x) = 0$ has a solution in the interval $[N, N + 1]$:

   (a) $f(x) = x^3 - x + 3$
   (b) $f(x) = x^5 + 5x^4 + 2x + 1$
   (c) $f(x) = x^5 + x + 1$
   (d) $f(x) = 4x^2 - 4x + 1$
   (e) $f(x) = \cot x$
   (f) $f(x) = \dfrac{4x^3 - x^2 - 2x + 2}{x^2 - 2}$

3. Verify that the rational function

$$f(x) = \frac{x^2 + 1}{x - 1}$$

takes both positive and negative values, but is never zero. Why doesn't this contradict the Intermediate Value Theorem?

4. Does the equation

$$\cos x = x$$

have any solutions? If so, can you specify an interval on which it has precisely one solution? Justify your answers.

5. For each function $f(x)$ below, decide whether or not it is monotone on the given interval I. Be specific (increasing? decreasing? nonincreasing? nondecreasing?). Justify your answers.

   (a) $f(x) = x^2 - 2x, \quad I = [0, 2]$

   (b) $f(x) = x^2 - 2x + 1, \quad I = [1, 2]$

   (c) $f(x) = \dfrac{x}{x+1}, \quad I = [0, 1]$

   (d) $f(x) = \lfloor x \rfloor$ (this denotes the largest integer $\leq x$), $I = (-\infty, \infty)$

   (e) $f(x) = \begin{cases} \sin x & \text{for } x \leq \frac{\pi}{2}, \\ 1 - \cos x & \text{for } x > \frac{\pi}{2}, \end{cases} \quad I = [0, \pi]$.

6. (a) **Show** that $\tan x$ is strictly increasing on $\left[0, \frac{\pi}{2}\right)$. (*Hint:* Consider the numerator and denominator.)

   (b) **Show** that for any x, $-\tan(-x) = \tan x$, and use this together with part (a) to show that $\tan x$ is strictly increasing on $\left(-\frac{\pi}{2}, 0\right]$.

   (c) How do (a) and (b) together prove that $\tan x$ is strictly increasing on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$?

   (d) Give an interval including $\left(0, \frac{\pi}{2}\right)$ on which $f(x) = \cot x$ is defined, strictly monotone, and takes negative as well as positive values.

   (e) Give a reasonable definition of the inverse cotangent function.

Theory problems:

7. (a) Show that there is no interval on which $f(x) = \sec x$ is monotone and takes both positive and negative values.

   (b) Give a reasonable definition of the function $\operatorname{arcsec} x$.

   (c) Do the same for the inverse cosecant.

8. True or False? Justify.

(a) An even degree polynomial is either constant or monotonic.

(b) Every odd degree polynomial is monotonic on its natural domain.

(c) The functions $f(x) = x^2$ and $g(x) = \sqrt{x}$ are inverses of each other on their natural domains.

(d) The functions $f(x) = x^3$ and $g(x) = x^{\frac{1}{3}}$ are inverses of each other on their natural domains.

(e) If $\sin\theta = x$ then $\arcsin x = \theta$.

(f) If $\arcsin x = \theta$ then $\sin\theta = x$.

9. (a) Show that if $\theta = \arcsin x$ for some x then

$$\cos\theta \geq 0.$$

(b) Use this to give a simplified expression for

$$\cos(\arcsin x).$$

(c) Do the analogous calculation for $\arccos x$.

10. Complete the proof of the Intermediate Value Theorem by showing that the third case of the proof follows from the second. That is, if we know that for every function satisfying $f(a) < f(b)$ the conclusion of the theorem holds, then it also holds for every functions satisfying $f(a) > f(b)$. (*Hint:* If $f(x)$ satisfies the second condition, find a related function $g(x)$ that satisfies the first, then see what equation in $g(x)$ will give a solution of $f(x) = C$.)

Challenge problem:

11. Use the Intermediate Value Theorem to show that a continuous function which is one-to-one on an interval must be strictly monotone on that interval.

12. (a) Show that a function which is continuous on an interval I and which takes only integer values on I must be constant on I.

(b) What can you say about a continuous function which takes only *rational* values on I?

13. Suppose $f(x)$ is a polynomial with at least two nonzero coefficients, and the coefficient of the highest power of x that appears has a sign opposite that of the lowest power of x which appears. Show that the equation $f(x) = 0$ has at least one solution.

History notes:

14. **Bolzano's Proof of the Intermediate Value Theorem:** Bolzano (1817) [6, p. 273] formulated the Intermediate Value Theorem in the following form:

    > *Suppose two functions $f(x)$ and $g(x)$ are continuous on $[a, b]$ with $f(a) < g(a)$ and $f(b) > g(b)$. Then for some number $\bar{x}$ between a and b, $f(\bar{x}) = g(\bar{x})$.*

    (a) Show that this statement is equivalent to Theorem 3.2.1.

    (b) Bolzano's proof uses the idea of the supremum; he had already established (in his own terms) that every nonempty set that is bounded above has a least upper bound (see Exercise 11 in § 2.5). Now, he lets M denote the set of numbers $r < b - a$ such that *either* $r \leq 0$ or $f(a + r) < g(a + r)$. Clearly, M is bounded above, and is nonempty (why?). Thus we can let U be the least upper bound of M; show that $0 < U < b - a$. Then an argument like the one we have for Theorem 3.2.1 shows that there are sequences $\{u_k\}$ and $\{v_k\}$ with $u_k \in M$ and increasing to U, and v_k decreasing to U, and so $\bar{x} = a + U$ works to give our conclusion.

15. **Cauchy's Proof of the Intermediate Value Theorem:** Cauchy (1821)[12] proves the Intermediate Value Theorem as follows:

    (a) Show that it suffices to prove Theorem 3.2.1 assuming that $C = 0$.

    (b) Given an integer $n > 1$, we divide $[a, b]$ into n subintervals of equal length, using the intermediate points $a = x_0 < x_1 < ... < x_n = b$. Since $f(a)$ and $f(b)$ have opposite sign (why?), there is at least one subinterval across which $f(x)$ changes sign (*i.e.*, the signs of $f(x)$ are opposite at the endpoints). Call this interval $[a_1, b_1]$. Now repeat the process, starting with $[a_1, b_1]$. Repeating the process, we get two sequences a nondecreasing sequence $\{a_k\}$ and a nonincreasing sequence $\{b_k\}$; since $|b_k - a_k|$ goes to zero (why?) the two sequences converge to the same point, but there f must equal zero (why?).

3.3 Extreme Values and Bounds for Functions

In this section, we extend to functions some of the notions discussed in § 2.5 for sets, and establish another important property of continuous functions.

Definition 3.3.1. *Given a function f and a subset $S \subset \mathrm{dom}(f)$ of its domain,*

1. *$\alpha \in \mathbb{R}$ is a **lower bound for f on S** if*

$$\alpha \leq f(s) \text{ for every } s \in S;$$

2. *$\beta \in \mathbb{R}$ is an **upper bound for f on S** if*

$$f(s) \leq \beta \text{ for every } s \in S.$$

*The function f is **bounded on S** if it has both a lower bound and an upper bound on S. If the phrase "on S" is omitted in any of these definitions, it is understood that we take S to be the (natural) domain of f.*

Note that these definitions are just the old definitions for bounds on a set of numbers, except that the set we apply them to is not S, but the set of *output values* which occur for f when the input values come from S. This set is sometimes called the **image of S under f**

$$f(S) = \{y \in \mathbb{R} \,|\, y = f(s) \text{ for some } s \in S\}.$$

(see Exercise 4).
The notions of minimum, maximum, infimum and supremum carry over to functions in a similar way.

Definition 3.3.2. *Suppose f is a function and $S \subset \mathrm{dom}(f)$ is a subset of its domain.*

1. *A number $\alpha \in \mathbb{R}$ is the **infimum of f on S***

$$\alpha = \inf_{x \in S} f(x)$$

if it is the greatest lower bound for f on S:

(a) *for every* $s \in S$,

$$\alpha \leq f(s)\,,$$

and

(b) *if* α' *is a lower bound for* S, *then*

$$\alpha' \leq \alpha.$$

2. *A number* $\alpha \in \mathbb{R}$ *is the* ***minimum value of*** f ***on*** S

$$\alpha = \min_{x \in S} f(x)$$

if it is a lower bound for f *on* S *that equals the value of* f *at some point of* S:

(a) *for every* $s \in S$,

$$\alpha \leq f(s)\,,$$

and

(b) *for* some $s_{min} \in S$,

$$\alpha = f(s_{min})\,.$$

We say f ***achieves its minimum on*** S *if it has a minimum on* S.

3. *A number* $\beta \in \mathbb{R}$ *is the* ***supremum of*** f ***on*** S

$$\beta = \sup_{x \in S} f(x)$$

if it is the least upper bound for f *on* S:

(a) *for every* $s \in S$,

$$f(s) \leq \beta,$$

and

(b) *if* β' *is an upper bound for* S, *then*

$$\beta \leq \beta'.$$

4. *A number* $\beta \in \mathbb{R}$ *is the* ***maximum value of*** f ***on*** S

$$\beta = \max_{x \in S} f(x)$$

if it is an upper bound for f *on* S *that equals the value of* f *at some point of* S:

(a) for every $s \in S$,

$$f(s) \leq \beta,$$

and

(b) for some $s_{max} \in S$,

$$\beta = f(s_{max})\,.$$

*We say f **achieves its maximum on** S if it has a maximum on S.*

By applying Theorem 2.5.6 to the image of S under f, we easily see the following (Exercise 5)

Remark 3.3.3. *If f is bounded below (resp. above) on $S \subset \operatorname{dom}(f)$, then $\inf_{x \in S} f(x)$ (resp. $\sup_{x \in s} f(x)$) exists.*
f achieves its minimum value (resp. maximum value) precisely if the infimum (resp. supremum) exists and belongs to the image of S—that is, if for some $s_{min} \in S$ (resp. $s_{max} \in S$)

$$f(s_{min}) = \inf_{x \in S} f(x) \ \ (resp.\ f(s_{max}) = \sup_{x \in S} f(x)).$$

When a set or function is not bounded above (*resp.* below) we sometimes abuse notation (as we did with divergence to infinity) by writing

$$\sup_{x \in s} f(x) = \infty$$

if f is not bounded above on S and

$$\inf_{x \in S} f(x) = -\infty$$

if f is not bounded below on S.
We consider a few examples of some of these phenomena.
First, the function

$$f(x) = x^2$$

is bounded below, but *not* above, on $(-\infty, \infty)$, so the supremum does not exist:

$$\sup_{x \in \mathbb{R}} f(x) = \infty,$$

but it achieves its minimum at $s_{min} = 0$:

$$\inf_{x \in \mathbb{R}} f(x) = \min_{x \in \mathbb{R}} f(x) = f(0) = 0.$$

On the interval

$$S = [2, \infty)\,,$$

a similar situation occurs: f is bounded below, but not above, so the supremum does not exist:

$$\sup_{2 \leq x} f(x) = \infty,$$

but the minimum is achieved at $s_{min} = 2$:

$$\inf_{2 \leq x} f(x) = \min_{2 \leq x} f(x) = f(2) = 4.$$

On the interval

$$S = (2, \infty)$$

we have the same bounds:

$$\sup_{x>2} f(x) = \infty, \quad \inf_{x>2} f(x) = 4$$

but since *no* $x > 2$ achieves $f(x) = 4$, f fails to achieve its minimum (or maximum) value on $(2, \infty)$.
Finally, on the interval

$$S = (-3, 2)$$

the function is bounded, and in fact achieves its minimum value:

$$\inf_{-3<x<2} f(x) = \min_{-3<x<2} f(x) = f(0) = 0.$$

It has a supremum

$$\sup_{-3<x<2} f(x) = 9$$

but fails to achieve this value for $-3 < x < 2$, and hence has no maximum value on $(-3, 2)$. See Figure 3.13.
As another example, the function

$$f(x) = \frac{1}{x}$$

is bounded below, but not above, on

$$S = (0, \infty)\,;$$

since $\frac{1}{n} \in (0, \infty)$ and $\lim f\left(\frac{1}{n}\right) = +\infty$, we have

$$\sup_{x>0} \frac{1}{x} = \infty.$$

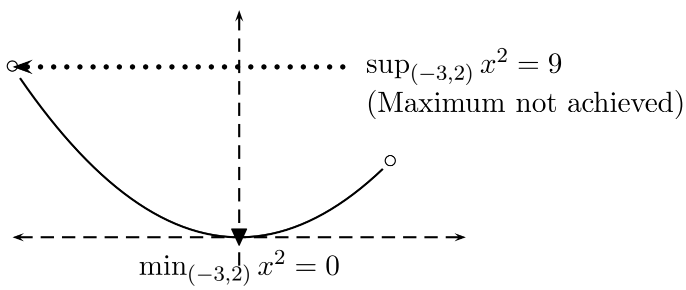


Figure 3.13: $y = x^2$ on $(-3, 2)$

Also, since $\frac{1}{x} > 0$ for $x > 0$, and $n \in (0, \infty)$ with $\lim \frac{1}{n} = 0$, we have

$$\inf_{x>0} \frac{1}{x} = 0$$

but no x achieves the value 0.

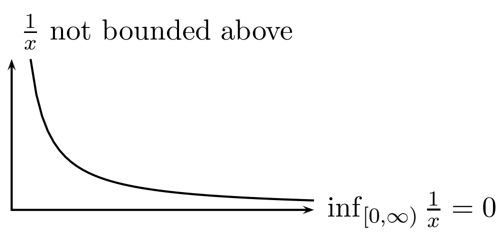


Figure 3.14: $y = \frac{1}{x}$ on $[0, \infty)$

As a final example, consider

$$f(x) = \frac{1}{x^2 + 1}.$$

This is defined for all x, and (since $x^2 \geq 0$)

$$0 < f(x) \leq 1$$

for all $x \in \mathbb{R}$. Thus, the function is bounded. We can check that

$$\sup \frac{1}{x^2+1} = \max \frac{1}{x^2+1} = f(0) = 1$$

but, since $\lim f(n) = 0$ and $f(x) > 0$ for all x,

$$\inf \frac{1}{x^2+1} = 0$$

but $1/(x^2+1)$ does not achieve a minimum value.

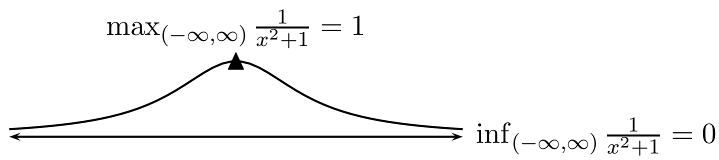


Figure 3.15: $y = \frac{1}{x^2+1}$ on $(-\infty, \infty)$

The following result states two basic properties of continuous functions on closed intervals: first, on any closed interval $[a,b]$, a continuous function is *bounded* (so $\inf_{[a,b]} f(x)$ and $\sup_{[a,b]} f(x)$ both exist), and furthermore it achieves its extreme values, so $\min_{[a,b]} f(x)$ and $\max_{[a,b]} f(x)$ both exist, as well.

Theorem 3.3.4 (Extreme Value Theorem)**.** *Suppose f is continuous on the closed interval $[a,b]$. Then f achieves its minimum and maximum on $[a,b]$: there exist points $s_{min}, s_{max} \in [a,b]$ such that*

$$f(s_{min}) \le f(s) \le f(s_{max})$$

for all $s \in [a,b]$.

The elegant proof we give here is due to Daniel Reem [43][11] We will concentrate on upper bounds; see Exercise 3 for lower bounds. Also see Exercise 7 for an alternative proof (the one found in many texts).

Proof. For $n = 1, 2, ...$ let $S_n \subset [a,b]$ consist of all points in $[a,b]$ with a finite decimal expansion of the form $...a_0.a_{-1} \cdots a_n$ (that is, all digits more than n places to the right of the decimal point are zero). Note that $S_n \subset S_m$ for $n \le m$.

[11] He calls it the "programmer's proof"—can you see why this name is appropriate?

Note first that S_n is finite (because elements of S_n are spaced at least 10^{-n} apart, so S_n has no more than $10^n/(b-a)$ points) , and hence so is the set $f(S_n) := \{f(x) \mid x \in S_n\}$ of values taken on by $f(x)$ among points of S_n. As we saw on p. 72, this set has a maximum; that is, there is some $x_n \in S_n$ such that for all $x \in S_n$, $f(x) \leq y_n = f(x_n)$.
Now consider the sequence of points $x_n \in [a,b]$. By Proposition 2.3.8 (the Bolzano-Weierstrass Theorem) the sequence x_n has a convergent subsequence; say $x_{n_k} \to x_\infty$. Since $[a,b]$ is closed, $x_\infty \in [a,b]$. Furthermore, since f is continuous, $f(x_\infty) = \lim f(x_{n_k}) = \lim y_{n_k}$. Call this limit y_∞. Since $f(S_n) \subset f(S_{n+1})$ for all n, we see that the sequence of values y_n is non-decreasing: $y_n \leq y_{n+1}$. It follows that the whole sequence of values converges: $y_n \to y_\infty$.
We want to show that for any $s \in [a,b]$, $f(s) \leq y_\infty = f(x_\infty)$. for $n = 1, 2, \ldots$, let s_n be its decimal expansion, truncated after n digits past the decimal point. Then $s_n \to s$, so since $f(x)$ is continuous, $f(x) = \lim f(s_n)$.
However, $s_n \in S_n$, so $f(s_n) \leq y_n$, and hence $f(x) = \lim f(s_n) \leq \lim y_n = f(x_\infty)$. Thus, x_∞ is precisely the point called s_{max} in the statement of the theorem.

□

We stress that, like the Intermediate Value Theorem, this result only asserts the *existence* of a maximum and minimum for any continuous function on any closed interval. It does *not* tell us how to find it. Methods for locating extreme values of a function involve the investigation of derivatives, which we handle in the next chapter.

Exercises for § 3.3

Answers to Exercises 1ace, 2adef, 8 are given in Appendix B.

Practice problems:

1. For each function below, decide whether it is bounded above and/or below on the given interval I. If so, decide whether it achieves its maximum (*resp.* minimum) on I.

   (a) $f(x) = x^2, \quad I = (-1, 1)$ (b) $f(x) = x^3, \quad I = (-1, 1)$

   (c) $f(x) = x^2, \quad I = (-\infty, \infty)$ (d) $f(x) = x^2, \quad I = (0, \infty)$

   (e) $f(x) = \begin{cases} \frac{1}{q} & \text{if } 0 \neq x = \frac{p}{q} \text{ in lowest terms,} \\ 0 & \text{if } x = 0 \text{ or } x \text{ is irrational,} \end{cases} \quad I = [0, 1]$

(f) $f(x) = \begin{cases} \frac{1}{q} & \text{if } 0 \neq x = \frac{p}{q} \text{ in lowest terms,} \\ 1 & \text{if } x = 0 \text{ or } x \text{ is irrational,} \end{cases}$ $\quad I = [0, 1]$

2. Give an example of each item below, or else explain why no such example exists. Give a continuous example where possible, otherwise explain why no such example exists. Similarly, examine whether it makes a difference if the interval includes all its (finite) endpoints.

   (a) A bounded function with domain an unbounded interval.

   (b) An unbounded function with domain an unbounded interval.

   (c) A bounded function with domain a bounded interval.

   (d) An unbounded function with domain a bounded interval.

   (e) A bounded function on a bounded interval I which does not achieve its maximum on I.

   (f) A bounded function on an unbounded interval I which *does* achieve its maximum and minimum on I.

   (g) A bounded function on a bounded interval I which does not achieve its maximum on I, but achieves its minimum on I..

Theory problems:

3. Complete the proof of Theorem 3.3.4 by showing that a function f which is continuous on the closed interval $[a, b]$ achieves its *minimum* on $[a, b]$.

   (*Hint:* There are two possible proofs of this. One mimics the *proof* for upper bounds on p. 124 while the other cleverly uses the *result* about *maxima.*)

4. Show that α (*resp.* β) is a lower (*resp.* upper) bound for the function f on the set S (in the sense of Definition 3.3.1) precisely if it is a lower (*resp.* upper) bound (in the sense of Definition 2.5.1) for the image $f(S)$ of S under f.

5. Show how Remark 3.3.3 follows from Theorem 2.5.6.

6. Show that for any polynomial f there exists a minimum for $|f(x)|$ on $(-\infty, \infty)$.

7. The "usual" proof of the Extreme Value Theorem given in many texts works in two stages.

(a) First, we prove that every function which is continuous on $[a,b]$ is bounded; the proof is by contradiction. Suppose f were *not* bounded above on $[a,b]$. By definition, this would mean that for every number—in particular for every positive integer k—we could find a point $s_k \in [a,b]$ where

$$f(s_k) \geq k. \tag{3.7}$$

Show that the sequence of points $\{s_k\}$ is bounded, and that some subsequence $\{t_j\}$

$$t_j = s_{k_j}, \ k_1 < k_2 < \ldots$$

must converge, say

$$t_j \to t.$$

Show that $t \in [a,b]$ and that

$$\lim f(t_j) = f(t) \in \mathbb{R}.$$

Use Equation (3.7) to show that $f(t)$ must exceed every integer, a contradiction which proves that f must be bounded on $[a,b]$.

(b) Now, to show that f achieves its maximum on $[a,b]$, let

$$\beta = \sup_{s \in [a,b]} f(s) \in \mathbb{R}.$$

Since β is the supremum of the set of values of f on $[a,b]$, it must be the limit of some sequence of values:

$$\beta = \lim f(s_k), \quad s_k \in [a,b].$$

Show that some subsequence of $\{s_k\}$

$$t_j = s_{k_j} \quad k_1 < k_2 < \ldots$$

converges:

$$t_j \to t \in [a,b],$$

and that the corresponding subsequence of values converges to β:

$$f(t_j) \to \beta.$$

Show that

$$f(t) = f(\lim t_j) = \lim f(t_j) = \beta$$

so that

$$\max_{x \in [a,b]} f(x) = f(t) = \beta.$$

Challenge problem:

8. For each function below, decide whether it is bounded above and/or below on the given interval I. If so, decide whether it achieves its maximum (*resp.* minimum) on I.

   (a) $f(x) = \begin{cases} x^2 & \text{for } x \leq a, \\ a+2 & \text{for } x > a, \end{cases}$
   where I is the *closed* interval whose endpoints are $-a-1$ and $a+1$. (Of course, your answer will depend on a, and may involve several cases.)

   (b) $f(x) = \begin{cases} x^2 & \text{for } x \leq a, \\ a+2 & \text{for } x > a, \end{cases}$
   where I is the *open* interval whose endpoints are $-a-1$ and $a+1$. (Again, your answer will depend on a, and may involve several cases.)

3.4 Limits of Functions

We have seen many functions which are continuous wherever they are defined, but how can a function *fail* to be continuous? Well, we need to have a convergent sequence $x_n \to x_0$ for which the sequence of values $f(x_n)$ does *not* converge to $f(x_0)$. Of course, for the sequence of values to make sense, the terms x_n, $n = 1, ...$ must belong to the domain of f. Since any term which happens to coincide with x_0 automatically has value $f(x_0)$, we can discard these terms, since they will not contribute to the problem. For any set $X \subset \mathbb{R}$, we write $X \setminus \{x_0\}$ to denote the set of points in X distinct from x_0

$$X \setminus \{x_0\} := \{x \in X \,|\, x \neq x_0\}.$$

So, to study failures of continuity, we want to consider convergent sequences $\{x_n\}$ in $\mathrm{dom}(f) \setminus \{x_0 := \lim x_n\}$. This situation has a name.

Definition 3.4.1. *A point $x_0 \in \mathbb{R}$ is an* ***accumulation point*** *of the set $X \subset \mathbb{R}$ if it is the limit of some sequence in $X \setminus \{x_0\}$.*

The collection of all accumulation points of a set $X \subset \mathbb{R}$ is often denoted $\mathbf{acc(}\boldsymbol{X}\mathbf{)}$. Note that our definition does *not* require that x_0 itself belong to X in order to be an accumulation point. For example, if X is the open interval $(0,1)$, then $\mathrm{acc}(X)$ includes, in addition to all the points in X, the endpoints 0 and 1:

$$\mathrm{acc}((0,1)) = [0,1]\,.$$

This will turn out to be useful. But also, a point which belongs to X need not be an accumulation point: for example, the domain of the function

$$g(x) := \sqrt{\frac{x^2}{x^2-1}}$$

is

$$\mathrm{dom}(g) = (-\infty,-1) \cup \{0\} \cup (1,\infty)\,;$$

it contains the number 0, but no *other* numbers nearby, so 0 is *not* an accumulation point of this set. A point which belongs to the set $X \subset \mathbb{R}$ but is *not* an accumulation point of X is called an **isolated point** of X; this means precisely that there is some open interval around the point which contains no other points of X. We see that the accumulation points of $\mathrm{dom}(g)$ are

$$\mathrm{acc}(\mathrm{dom}(g)) = (-\infty,-1] \cup [1,\infty)\,;$$

the isolated point 0 of $\mathrm{dom}(g)$ is deleted from, and ± 1 are appended to, the domain.

Suppose now that we have x_0 an accumulation point of $\mathrm{dom}(f)$, say $x_0 = \lim x_n$ with $x_n \in \mathrm{dom}(f) \setminus \{x_0\}$, but it is *not* true that $f(x_0) = \lim f(x_n)$. This can be the case for several reasons. If the sequence of values $\{f(x_n)\}$ is divergent (*i.e.*, $\lim f(x_n)$ does not exist) then the situation is pretty messy. But suppose the sequence of values converges; then the failure is either because $f(x_0)$ is defined to be different from $\lim f(x_n)$, or because $f(x_0)$ is undefined (that is, x_0 is an accumulation point, but not an element of the domain). Then we might hope that by (re)defining $f(x_0)$ to be $\lim f(x_n)$, we could regain continuity. However, this naively overlooks the requirement that $f(x_0)$ must equal $\lim f(x_n)$ not just for *one* sequence $x_n \to x_0$ in $\mathrm{dom}(f) \setminus \{x_0\}$, but for *every* such sequence. Thus, we need some consistency between *different* sequences in $\mathrm{dom}(f)$ which converge to the *same* point.

All of these considerations are incorporated into the following formal definition.

Definition 3.4.2. *Suppose f is a function, $y \in \mathbb{R}$ and x_0 is an accumulation point of* $\mathrm{dom}(f)$. *We say that $f(x)$* ***converges to*** *y* ***as*** *x* ***approaches*** *x_0, and write*

$$y = \lim_{x \to x_0} f(x)$$

(pronounced "y is the **limit** *of f(x) as x approaches x_0") if every sequence* $\{x_k\}$ *in* $\operatorname{dom}(f) \setminus \{x_0\}$ *which converges to* x_0

$$x_k \to x_0$$

satisfies

$$f(x_k) \to y.$$

To check this definition for a given function, we replace x in $f(x)$ with a sequence x_k converging to x_0, and investigate the convergence of the resulting sequence $f(x_k)$. For example, to check

$$\lim_{x \to 3} (x^2 - 2x + 2)$$

we take a generic sequence x_k converging to 3 (with each $x_k \neq 3$) and consider the corresponding sequence

$$f(x_k) = x_k^2 - 2x_k + 2.$$

By Theorem 2.4.1, we know that

$$\begin{aligned} \lim(x_k^2 - 2x_k + 2) &= (\lim x_k)^2 - 2(\lim x_k) + 2 \\ &= 3^2 - 2 \cdot 3 + 2 \\ &= 5. \end{aligned}$$

Our use of Theorem 2.4.1 in this case amounted to finding the limit of $f(x)$ as $x \to 3$ by calculating the value of the function at $x = 3$. In fact, we can make the rather obvious

Remark 3.4.3. *If a function f is continuous on an open interval, then*

$$\lim_{x \to x_0} f(x) = f(x_0)$$

for any point x_0 in that interval.

In particular, this is true for any point x_0 in the domain of any polynomial or rational function.
The story is a little different when x_0 is *not* in the domain of f. For the rational function

$$f(x) = \frac{x^2 - 1}{x^3 - 1}$$

the point $x_0 = 1$ is not in the domain, but it is an accumulation point of the domain. To determine its limit at $x_0 = 1$, we substitute a generic sequence x_k converging to 1 into the function, and investigate the behavior of the resulting sequence $f(x_k)$. In this situation, it is convenient to express the sequence as

$$x_k = 1 + t_k, \quad \text{where } t_k \to 0 \text{ (and } t_k \neq 0).$$

The substitution leads to

$$\begin{aligned}\lim f(x_k) = \lim \frac{(1+t_k)^2 - 1}{(1+t_k)^3 - 1} &= \lim \frac{1 + 2t_k + t_k^2 - 1}{1 + 3t_k + 3t_k^2 + t_k^3 - 1} \\ = \lim \frac{2t_k + t_k^2}{3t_k + 3t_k^2 + t_k^3}\end{aligned}$$

We cannot apply Theorem 2.4.1 to this form, since both the numerator and denominator go to 0, but if we divide both by their common factor $t_k (\neq 0)$, we have

$$\lim f(x_k) = \lim \frac{2 + t_k}{3 + 3t_k + t_k^2} = \frac{2 + 0}{3 + 3 \cdot 0 + 0^2} = \frac{2}{3}.$$

A different example is provided by

$$g(x) = \sin\left(\frac{1}{x}\right).$$

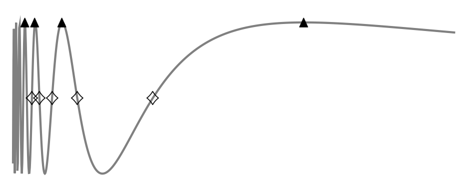


Figure 3.16: $y = \frac{1}{x}$ on $(0, \infty)$

The point $x_0 = 0$ is an accumulation point outside the domain of g, so to find the limit at 0 we pick a sequence $x_k \to 0$ and consider

$$\lim g(x_k) = \lim \sin\left(\frac{1}{x_k}\right).$$

The function $\sin\theta$ is understood to interpret θ as the angle measured in radians, so in particular for every integer k we have

$$\sin k\pi = 0.$$

Thus the sequence

$$x_k = \frac{1}{k\pi}, \quad k = 1, 2, ...$$

(corresonding to the points marked by $\diamondsuit$ in Figure 3.16) has $x_k \to 0$ and leads to

$$\lim g(x_k) = \lim \sin\left(\frac{1}{x_k}\right) = \lim \sin(k\pi) = \lim 0 = 0.$$

But also,

$$\sin\left((2k + \frac{1}{2})\pi\right) = 1$$

so

$$x_k = \frac{1}{(2k+\frac{1}{2})\pi} = \frac{2}{(4k+1)\pi}$$

(corresonding to the points marked by ▲ in Figure 3.16) has $x_k \to 0$ and leads to

$$\lim g(x_k) = \lim \left(\sin(\frac{1}{x_k})\right) \qquad = \lim\left(\sin\left(\frac{4k+1}{2}\pi\right)\right)$$
$$= \lim\left(\sin(2k\pi + \frac{\pi}{2})\right) = \lim(1) = 1.$$

We have found two sequences with $x_k \to 0$, $x_k \neq 0$, for which $\lim g(x_k)$ has distinct values. It follows that

$$\lim_{x\to 0} \sin\left(\frac{1}{x}\right) \text{ does not exist: } \sin\left(\frac{1}{x}\right) \text{ diverges as } x \to 0.$$

Before proceeding with further examples, we note that our basic results on limits of sequences easily translate to results on limits of functions.

Theorem 3.4.4 (Arithmetic for Limits of Functions)**.** *Suppose f and g are functions, x_0 is an accumulation point of* $\text{dom}(f) \cap \text{dom}(g)$*, and*

$$\lim_{x\to a} f(x) = L_f, \quad \lim_{x\to a} g(x) = L_g.$$

Then

1. $\lim_{x\to a}(f(x) \pm g(x)) = L_f \pm L_g$;
2. $\lim_{x\to a} f(x)\, g(x) = L_f L_g$;
3. $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{L_f}{L_g}$, *provided* $L_g \neq 0$;

4. $\lim_{x\to a} \sqrt[n]{f(x)} = \sqrt[n]{L_f}$, *provided* $L_f > 0$.

The proof of these statements is straightforward (Exercise 3): in each case we substitute a sequence $x_k \to a$ $(x_k \neq a)$ into the function and use Theorem 2.4.1 together with the known facts that $\lim_{x\to a} f(x) = L_f$, $\lim_{x\to a} g(x_k) = L_g$.
The Squeeze Theorem also lets us handle some limits involving possibly divergent functions.

Theorem 3.4.5 (Squeeze Theorem for Functions)**.** *Suppose f, g and h are functions, and x_0 is an accumulation point of a set D which contains all of the domain of h near $x = a$, and on which both f and g are defined. Furthermore, suppose*

1. $f(x) \leq h(x) \leq g(x)$ *for all* $x \in D$
2. $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = L$.

Then also

$$L = \lim_{x\to a} h(x).$$

Again, this is proved by substituting a sequence $x_k \in D$ with $x_k \to a$, $x_k \neq a$ and applying Theorem 2.4.7 to the sequences $f(x_k)$, $g(x_k)$, and $h(x_k)$.

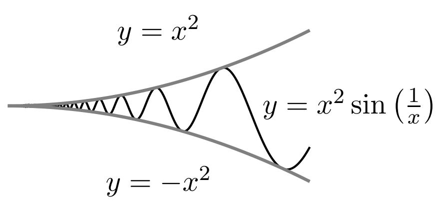


Figure 3.17: Theorem 3.4.5 (Squeeze Theorem)

The Squeeze Theorem is useful, for example, in finding

$$\lim_{x\to 0} x^2 \sin\left(\frac{1}{x}\right).$$

Since

$$-1 \leq \sin\frac{1}{x} \leq 1,$$

we know that

$$-x^2 \leq x^2 \sin \frac{1}{x} \leq x^2$$

and

$$\lim_{x \to 0} x^2 = \lim_{x \to 0} (-x^2) = 0.$$

The Squeeze Theorem then guarantees that

$$\lim_{x \to 0} x^2 \sin \left(\frac{1}{x} \right) = 0.$$

Another useful concept, particularly for functions defined in pieces, is that of *one-sided limits.* Consider, for example, the function defined in pieces (see Figure 3.18) by

$$f(x) = \begin{cases} 1 - x & \text{if } x < 0 \\ x^2 + 1 & \text{if } x > 0. \end{cases}$$

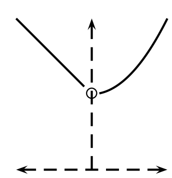


Figure 3.18: $f(x)$ defined in pieces

If we try to find $\lim_{x \to 0} f(x)$ by substituting a sequence x_k which converges to 0, we have no *a priori* control over which side of 0 our terms hit, and so cannot effectively decide whether to deal with $1 - x_k$ or $x_k^2 + 1$. However, if we *know* that our sequence *always* satisfies $x_k < 0$, then we easily see from Theorem 2.4.1 that

$$\lim f(x_k) = \lim(1 - x_k) = 1 - 0 = 1$$

and similarly if $x_k > 0$ for all k, then

$$\lim f(x_k) = \lim(x_k^2 + 1) = 0 + 1 = 1.$$

These special sequences[12] actually determine the one-sided limits of $f(x)$ at $x = 0$.

[12]We will not repeat in these definitions the stipulations that $x_k \in \mathrm{dom}(f)$ and that at least one sequence of the required type must exist for the limit to make sense.

Definition 3.4.6. *The* ***one-sided limits*** *of a function f at x_0 are*

1. ***limit from the right:***

$$L_+ = \lim_{x \to a^+} f(x)$$

(pronounced "L_+ is the limit of $f(x)$ as x **approaches** x_0 **from the right** *(or from above)") if*

$$f(x_k) \to L_+$$

for each sequence $x_k \to a$, $x_k \neq a$ with

$$x_k > a \quad \textit{for all } k.$$

2. ***limit from the left:***

$$L_- = \lim_{x \to a^-} f(x)$$

(pronounced "L_- is the limit of $f(x)$ as x **approaches** x_0 **from the left** *(or from below)") if*

$$f(x_k) \to L_-$$

for each sequence $x_k \to a$, $x_k \neq a$ with

$$x_k < a \quad \textit{for all } k.$$

In our example we saw that

$$\begin{aligned}
\lim_{x \to 0^-} f(x) &= \lim_{x \to 0^-} (1 - x) &= 1 - 0 = 1 \\
\lim_{x \to 0^+} f(x) &= \lim_{x \to 0^+} (x^2 + 1) &= 0 + 1 = 1.
\end{aligned}$$

Since these two one-sided limits both equal 1 we suspect that *any* sequence $x_k \to a$ (with $x_k \neq a$) will satisfy $f(x_k) \to 1$, so that

$$\lim_{x \to 0} f(x) = 1.$$

The following result justifies this.

Lemma 3.4.7. *If the one-sided limits of $f(x)$ at a*

$$\lim_{x\to a^+} f(x) = L_+, \quad \lim_{x\to a^-} f(x) = L_-$$

both exist and are equal,

$$L_+ = L_- = L$$

then

$$\lim_{x\to a} f(x) = L.$$

Proof. We know that any sequence $x_k^{\pm} \to a$ which stays on one side of a has $f(x_k^{\pm}) \to L$. We need to show this is also true for sequences $x_k \to a$ which switch sides.

If the sequence x_k switches sides only a finite number of times, the proof is trivial (Exercise 4). Otherwise, we can consider the two subsequences consisting of the x_k lying to the *right* (*resp.* left) of a; each of these subsequences converges to a from one side, the corresponding subsequences of values $f(x_k)$ both converge to L. This means that given $\varepsilon > 0$ we can find a place in each subsequence such that for every term beyond that place, the corresponding value of f is within distance ε of L: in other words, we have N^+ and N^- so that $f(x_k)$ is within distance ε of L in case either $x_k > a$ and $k > N^+$ or $x_k < a$ and $k > N^-$; but then taking $N = \max(N^+, N^-)$, $k > N$ guarantees that $f(x_k)$ is within distance ε of L regardless of which side of a x_k is on. This shows that $x_k \to L$, and since we started with an arbitrary sequence $x_k \to a$, $x_k \neq a$, we have shown

$$\lim_{x\to a} f(x) = L$$

as required. □

Of course, if the two one-sided limits are *different*, then the (two-sided) limit *diverges*.[13] For example,

$$f(x) = \frac{|x|}{x}(x^2 + 1)$$

has

$$f(x) = \begin{cases} -x^2 - 1 & \text{if } x < 0 \\ x^2 + 1 & \text{if } x > 0 \end{cases}$$

(see Figure 3.19)

[13] Strictly speaking, the *limit* does not exist; the *function* diverges as $x \to a$.

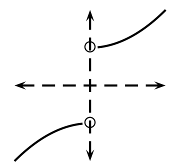


Figure 3.19: Different one-sided limits

so

$$\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-}(-x^2-1) = -0-1 = -1$$
$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+}(x^2+1) = 0+1 = 1$$

and since these one-sided limits are different,

$$\lim_{x\to 0} f(x) \text{ does not exist: } f(x) \text{ diverges as } x \to 0.$$

The lemma can sometimes be useful in situations that are not defined in pieces. We saw in Proposition 3.1.7 that $\lim_{\theta\to 0}\sin\theta = 0$ and $\lim_{\theta\to 0}\cos\theta = 1$; in effect, these proofs consisted of showing that the one-sided limits agree and equal the given value.
A much more subtle (and very important) trig limit is the following

Lemma 3.4.8. $\lim_{\theta\to 0}\frac{\sin\theta}{\theta} = 1$.

IMPORTANT: Up to now, we have used the convention that θ inside a trig function indicates an angle of θ *radians*, but this has not affected the results very much. However, for this limit result, it is *crucial* that this convention be in force. If we understood $\sin\theta$ to mean the sine of an angle of θ *degrees*, the limit would be $\frac{\pi}{180}$! (Exercise 5)

Proof. We elaborate on Figure 3.2 from § 3.1 (see Figure 3.20) by extending the line segment $\overline{OB}$ to a point D directly above A; recall that O, A, and B have respective coordinates $(0,0)$, $(0,1)$, and $(\cos\theta, \sin\theta)$, and the point on the x-axis directly below B is C, with coordinates $(\cos\theta, 0)$. The triangles $\triangle OCB$ and $\triangle OAD$ are similar, so that

$$\frac{\overline{DA}}{\overline{BC}} = \frac{\overline{OA}}{\overline{OC}}.$$

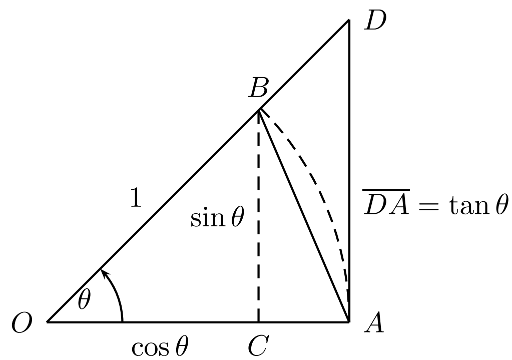


Figure 3.20: Lemma 3.4.8

Substituting the known quantities in this relation, we obtain

$$\frac{\overline{DA}}{\sin\theta} = \frac{1}{\cos\theta}$$

or

$$\overline{DA} = \frac{\sin\theta}{\cos\theta} = \tan\theta.$$

This time, we will compare areas for the following:

- the triangle OAB has base $\overline{OA} = 1$ and height $\overline{CB} = \sin\theta$, so

$$\text{area}(\triangle OAB) = \frac{1}{2}\sin\theta$$

- the sector $\measuredangle OAB$ cut out of the circle by the radii OA and OC has a central angle of θ radians, which constitutes a proportion $\frac{\theta}{2\pi}$ of the whole circle; since the circle has area[14] $\pi r^2 = \pi$, the sector has area

$$\text{area}(\measuredangle OAB) = \frac{\theta}{2\pi} \cdot \pi = \frac{\theta}{2}$$

- the triangle $\triangle OAD$ has base $\overline{OA} = 1$ and height $\overline{AD} = \tan\theta$, so its area is

$$\text{area}(\triangle OAD) = \frac{1}{2}\tan\theta = \frac{1}{2}\frac{\sin\theta}{\cos\theta}.$$

[14]See Exercise 7 in § 5.8, and Exercise 11 in this section.

Since the area of a subregion is no more than that of the whole region, the inclusions

$$\triangle OAB \subset \measuredangle OAB \subset \triangle OAD$$

yield the inequalities

$$\frac{1}{2}\sin\theta \leq \frac{\theta}{2} \leq \frac{1}{2}\frac{\sin\theta}{\cos\theta}.$$

Now, for $0 < \theta < \frac{\pi}{2}$, $\sin\theta > 0$, so we can divide all three quantities above by $\frac{1}{2}\sin\theta$ without reversing any inequalities, to get

$$1 \leq \frac{\theta}{\sin\theta} \leq \frac{1}{\cos\theta}.$$

Now as $\theta \to 0^+$, the right-hand quantity (as well as the left) goes to 1, hence by the Squeeze Theorem

$$\lim_{\theta\to 0^+} \frac{\theta}{\sin\theta} = 1.$$

But then Theorem 3.4.4 together with the observation

$$\frac{\sin\theta}{\theta} = \frac{1}{\left(\frac{\theta}{\sin\theta}\right)}$$

gives

$$\lim_{\theta\to 0^+} \frac{\sin\theta}{\theta} = \frac{1}{\lim_{\theta\to 0^+}\left(\frac{\theta}{\sin\theta}\right)} = \frac{1}{1} = 1.$$

It is relatively easy to see (Exercise 6) that this also gives

$$\lim_{\theta\to 0^-} \frac{\sin\theta}{\theta} = 1$$

and hence the desired (two-sided) limit.[15] □

This limit and its twin

$$\lim_{\theta\to 0} \frac{1-\cos\theta}{\theta} = 0$$

(see Exercise 7) play a very important role in the calculation of derivatives for trig functions. Before leaving this topic, we comment briefly on some variants of this.

First, consider the limit

$$\lim_{x\to 0} \frac{\sin 2x}{2x}.$$

[15] Exercise 11 gives a different proof of this result.

This is easy: we can rewrite the whole business in terms of $\theta = 2x$, which also goes to 0 as $x \to 0$. Thus

$$\lim_{x\to 0} \frac{\sin 2x}{2x} = \lim_{\theta \to 0} \frac{\sin\theta}{\theta} = 1.$$

Although intuitively we seem to have "cancelled" the factor of 2 from in front of both x's, this is *not* an allowed move, because $\sin 2x$ and $2\sin x$ are quite distinct[16]. A second example is

$$\lim_{x\to 0} \frac{\sin 2x}{3x}.$$

This can actually be handled the same way as the preceding, by rewriting in terms of $\theta = 2x$ (or $x = \frac{\theta}{2}$): we have

$$\begin{aligned}\lim_{x\to 0} \frac{\sin 2x}{3x} \underset{\theta=2x}{=} \lim_{\theta\to 0} \frac{\sin\theta}{\frac{3}{2}\theta} &= \lim_{\theta\to 0} \frac{2}{3} \cdot \frac{\sin\theta}{\theta} \\ = \frac{2}{3} \lim_{\theta\to 0} \frac{\sin\theta}{\theta} &= \frac{2}{3}(1) = \frac{2}{3}.\end{aligned}$$

Again, one must go through a procedure like this rather than blindly "factoring out" the 2 and 3.

We turn now to two kinds of "limits" involving "infinity".

First, we sometimes ask about the behavior of a function, defined for all (large) real numbers, as the input goes to one or the other "end" of the number line. This can be formulated by simply replacing "$x \to a$" with "$x \to \infty$" in the old definition of limit:

Definition 3.4.9. *Given a function f (defined for all sufficiently high values of x), we say $L \in \mathbb{R}$ is the **limit of** $f(x)$ **at** ∞(resp. at $-\infty$)*

$$L = \lim_{x\to\infty} f(x) \quad \textit{(resp. } L = \lim_{x\to -\infty} f(x)\textit{)}$$

if $f(x_k) \to L$ for every sequence $x_k \in \mathrm{dom}(f)$ that diverges to ∞ (resp. diverges to $-\infty$).

We note that the tools previously at our disposal, Theorem 3.4.4 and Theorem 3.4.5, can also be used when "a" is one of the the virtual numbers $\pm\infty$. This is also most effective when combined with the following (intuitively obvious) limit result.

Lemma 3.4.10. $\lim_{x\to\infty} \frac{1}{x} = 0 = \lim_{x\to-\infty} \frac{1}{x}$.

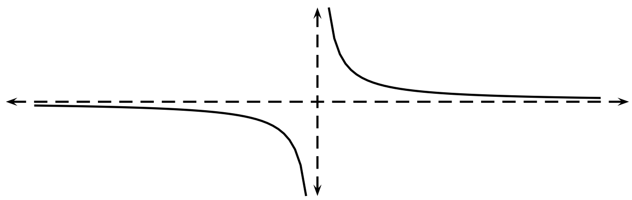


Figure 3.21: $y = \frac{1}{x}$

(See Figure 3.21.)

Proof. We need to show that if a sequence x_k diverges to ∞, then $\frac{1}{x_k} \to 0$. Given $\alpha > 0$, there exists K so that $k \geq K$ guarantees

$$\frac{1}{\alpha} < x_k.$$

But then (since both are positive) we have (for all $k \geq K$)

$$\alpha > x_k > 0,$$

in particular, $k \geq K$ guarantees

$$|x_k - 0| < \alpha.$$

(The other limit is an easy variant on this one; see Exercise 8.) □

The limit at infinity of a rational function is easily calculated if we first divide the numerator and denominator by the highest power of x present in the denominator. For example,

$$\lim_{x\to\infty} \frac{3x^2 - 2x + 1}{x^2 + 3x - 2} = \lim_{x\to\infty} \frac{3 - \frac{2}{x} + \frac{1}{x^2}}{1 + \frac{3}{x} - \frac{2}{x^2}} = \frac{3 - 0 + 0}{1 + 0 - 0} = 3$$

$$\lim_{x\to\infty} \frac{3x^2 + 1}{x^3 - 2x^2 + 1} = \lim_{x\to\infty} \frac{\frac{3}{x} + \frac{1}{x^3}}{1 - \frac{2}{x} + \frac{1}{x^2}} = \frac{0 + 0}{1 - 0 + 0} = 0$$

$$\lim_{x\to\infty} \frac{3x^2 - 2x + 1}{x - 2} = \lim_{x\to\infty} \frac{3x - 2 + \frac{1}{x}}{1 - \frac{2}{x}}$$

This last limit doesn't go quite as automatically, since the first term on top diverges to ∞. However, we can now reason as follows: for x large positive, the denominator is close to $1 - 0 = 1$, and the terms other than the first in

[16] for example, when $x = \frac{\pi}{4}$, $\sin 2x = \sin\frac{\pi}{2} = 1$, while $2\sin x = 2\sin\frac{\pi}{4} = \sqrt{2}$

the numerator are near $-2+0=-2$. So the whole fraction consists (for x large positive) of the large number $3x$, minus (-2)) (so the numerator is large positive) divided by a number near 1. This is clearly large positive, so the limit diverges (to ∞, as we shall see).
As a second example, consider

$$\lim_{x\to-\infty}\frac{\sin x^2}{x^2}$$

(see Figure 3.22).

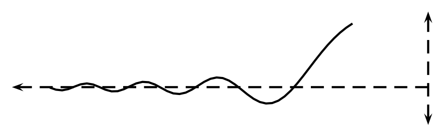


Figure 3.22: $y=\frac{\sin x^2}{x^2}$ on $(-\infty,-1)$

Here, $\sin x^2$ does not converge (as $x\to-\infty$), but we know that it is always between -1 and $+1$. Thus

$$-\frac{1}{x^2}\le\frac{\sin x^2}{x^2}\le\frac{1}{x^2}$$

and since

$$\lim_{x\to-\infty}-\frac{1}{x^2}=-\{\lim_{x\to-\infty}\left(\frac{1}{x}\right)\}^2=0$$
$$\lim_{x\to-\infty}\frac{1}{x^2}=\{\lim_{x\to-\infty}\left(\frac{1}{x}\right)\}^2\quad=0$$

we have

$$\lim_{x\to-\infty}\frac{\sin x^2}{x^2}=0$$

by the Squeeze Theorem.
We also want to deal with divergence to infinity. We do this by replacing "$f(x_k)\to L$" with "$f(x_k)\to\infty$" in the definition of limit.

Definition 3.4.11. *Given $a\in\mathbb{R}$ and f.*

1. *We say $f(x)$ **diverges to ∞ as** $x\to a$ **from the right** and write*

$$\lim_{x\to a^+}f(x)=+\infty$$

if every sequence $x_k \in \mathrm{dom}(f)$ *with* $x_k \to a$ *and* $x_k > a$ *yields* $f(x_k) \to \infty$.

Similarly

$$\lim_{x \to a^+} f(x) = -\infty$$

if $f(x_k) \to -\infty$ *for every sequence* $x_k \to a$ *with* $x_k > a$ *and*

$$\lim_{x \to a^-} f(x) = \pm\infty$$

if every sequence $x_k \in \mathrm{dom}(f)$ *with* $x_k \to a$ *and* $x_k < a$ *yields* $f(x_k) \to +(-)\infty$.

2. $f(x)$ ***diverges to*** $+\infty$ ***at*** $+\infty$

$$\lim_{x \to \infty} f(x) = \infty$$

if every sequence $x_k \in \mathrm{dom}(f)$ *which diverges to* $+\infty$ *yields* $f(x_k) \to +\infty$.

We saw already that

$$\lim_{x \to \infty} \frac{3x^2 - 2x + 1}{x - 2} = \infty.$$

If we try to find

$$\lim_{x \to -1^+} \frac{x^3 - 1}{x^2 - 1}$$

by substituting $x_k = -1 + t_k$, where $t_k > 0$ and $t_k \to 0$, we find

$$\begin{aligned} f(x_k) &= \frac{(-1 + t_k)^3 - 1}{(-1 + t_k)^2 - 1} \qquad = \frac{-1 + 3t_k - 3t_k^2 + t_k^3 - 1}{1 - 2t_k + t_k^2 - 1} \\ &= \frac{-2 + 3t_k - 3t_k^2 + t_k^3}{-2t_k + t_k^2} \end{aligned}$$

and noticing that for t_k small positive, the numerator is near -2, while for small positive t_k (even just $0 < t_k < 2$) the denominator, which can be written as $t_k(t_k - 2)$, is small and *negative*, we can conclude that the fraction is a large positive number, so

$$\lim_{x \to -1^+} \frac{x^3 - 1}{x^2 - 1} = +\infty.$$

Exercises for § 3.4

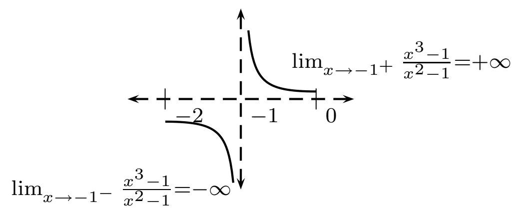


Figure 3.23: $y = \frac{x^3-1}{x^2-1}$ near $x = -1$

Answers to Exercises 1acegikmoqsu are given in Appendix B.

Practice problems:

1. Calculate each limit below, or explain why it does not exist.

(a) $\displaystyle\lim_{x\to1} \frac{x^3-1}{x^2-1}$ (b) $\displaystyle\lim_{x\to-1} \frac{x^3-1}{x^2-1}$

(c) $\displaystyle\lim_{x\to-1} \frac{x^2-1}{x^3-1}$ (d) $\displaystyle\lim_{x\to\infty} \frac{x^3-1}{x^2-1}$

(e) $\displaystyle\lim_{x\to\infty} \frac{x^2-1}{x^3-1}$ (f) $\displaystyle\lim_{x\to\infty} \frac{3x^2-2x-5}{4x^2+3x-1}$

(g) $\displaystyle\lim_{x\to-1} \frac{3x^2-2x-5}{4x^2+3x-1}$ (h) $\displaystyle\lim_{x\to1} \frac{x^3-x^2-x-1}{x^3-x^2+x-1}$

(i) $\displaystyle\lim_{x\to\infty} \arctan x$ (j) $\displaystyle\lim_{x\to-\infty} \arctan x$

(k) $\displaystyle\lim_{x\to0^+} \frac{|x|}{x}$ (l) $\displaystyle\lim_{x\to0^-} \frac{|x|}{x}$

(m) $\displaystyle\lim_{x\to0} \frac{|x|}{x}$ (n) $\displaystyle\lim_{\theta\to0} \frac{\sin 2\theta}{\sin 3\theta}$

(o) $\displaystyle\lim_{\theta\to0} \frac{\tan 5\theta}{3\theta}$ (p) $\displaystyle\lim_{\theta\to0} \frac{\tan 5\theta}{\sec 5\theta}$

(q) $\displaystyle\lim_{\theta\to0} \frac{\tan 5\theta}{\sec 3\theta}$ (r) $\displaystyle\lim_{\theta\to0} \frac{1-\cos 2\theta}{\theta}$

(s) $\displaystyle\lim_{\theta\to0} \frac{1-\cos\theta}{\sin\theta}$ (t) $\displaystyle\lim_{x\to\infty} (\sqrt{x^2+1}-x)$

(u) $\displaystyle\lim_{x\to0} x\sin\frac{1}{2x}$

(v) $\displaystyle\lim_{x\to1} f(x)$, where $f(x) = \begin{cases} x+1 & \text{for } x<1 \\ x^2-2x+3 & \text{for } x>1 \end{cases}$

Theory problems:

2. Use the Squeeze Theorem (Theorem 3.4.5) to show that
$$\lim_{x\to\infty} \frac{\cos 2x}{x^2+1} = 0.$$

3. Prove each part of Theorem 3.4.4, by substituting a sequence $\{x_k\}$ ($x_k \neq a$) into $f(x)$ and $g(x)$ and using Theorem 2.4.1 together with the hypotheses of the theorem.

4. In the proof of Lemma 3.4.7, show that if the sequence x_k *eventually* lies on one side of a, then we can get the result immediately from the hypotheses.

5. Explain the comment following Lemma 3.4.8 by showing that if $f(x)$ is the function assigning to a number x the sine of the angle x *degrees*, then
$$\lim_{x\to 0} \frac{f(x)}{x} = \frac{\pi}{180}.$$

6. Show that the equality
$$\lim_{\theta\to 0^+} \frac{\sin\theta}{\theta} = 1$$
implies the equality
$$\lim_{\theta\to 0^-} \frac{\sin\theta}{\theta} = 1.$$
(*Hint:* $\sin\theta$ is an *odd* function; what about $\frac{\sin\theta}{\theta}$?)

7. Show that
$$\lim_{\theta\to 0} \frac{1-\cos\theta}{\theta} = 0.$$
(*Hint:* Rewrite the fraction by multiplying top and bottom by $1+\cos\theta$, and using a trigonometric identity, as well as some rules about the arithmetic of limits.)

8. How can the proof that
$$\lim_{x\to\infty} \frac{1}{x} = 0$$
in Lemma 3.4.10 be modified to handle the limit as $x \to -\infty$?

Challenge problems:

9. Show that for any rational function f, if either of $\lim_{x\to\infty} f(x)$ or $\lim_{x\to-\infty} f(x)$ is finite, then the two are equal.

10. Suppose $\lim_{x\to a} g(x) = L$ and f is bounded on an open interval containing $x = a$ (but not necessarily continuous, nor not even necessarily having a limit as $x \to a$).

    (a) Show that if $L = 0$, then $\lim_{x\to a} [f(x)\, g(x)] = 0$.

    (b) Give an example to show that if $L \neq 0$ then $\lim_{x\to a} [f(x)\, g(x)]$ need not even exist.

History note:

11. **Newton on $\frac{\sin\theta}{\theta}$:** In Lemma 7, Book 1 of the *Principia* [15, p.436], Newton proves the following (see Figure 3.24): Let RA and RB be radii of a circle, and let the line tangent to the circle at A meet the line RB at D. As the point point B approaches A along the circle, the ratios of lengths of the arc $\overset{\frown}{AB}$, the chord $\overline{AB}$ and the tangent segment $\overline{AD}$ all tend to 1:

$$\overset{\frown}{AB}/\overline{AB},\ \overline{AB}/\overline{AD},\ \overset{\frown}{AB}/\overline{AD} \to 1.$$

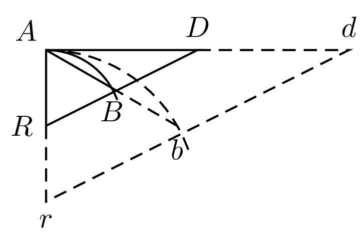


Figure 3.24: Newton's Lemma 1.7

    (a) Newton argues as follows: extend AD to a point d; note that AD does not depend on the location of B, so we will keep d fixed as B approaches A. Now extend AR to r so that the triangle

$\triangle Adr$ is similar to the triangle ADR, and let AB meet rd at b. Then similarity gives equality of the following pairs of ratios:

$$\overset{\frown}{AB}/\overline{AB} = \overset{\frown}{Ab}/\overline{Ab}$$
$$\overline{AB}/\overline{AD} = \overline{Ab}/\overline{Ad}$$
$$\overset{\frown}{AB}/\overline{AD} = \overset{\frown}{Ab}/\overline{Ad}$$

and as B approaches A, the angle $\angle BAD$ approaches zero (this is the content of Newton's preceding lemma); but this is also $\angle bAd$, and so the point b approaches d, and hence the right-hand ratios above all go to 1.

(b) Show that, if the radius of the circle is $\overline{AR} = \rho$ and the angle subtended by the two radii is $\angle ARB = \theta$, then the length of the chord is

$$\overline{AB} = 2\rho \sin \frac{\theta}{2}.$$

(*Hint:* See Exercise 16 in § 4.2.)

(c) Use this to show that Newton's result for the ratio of the chord to the arc is equivalent to

$$\frac{\sin \frac{\theta}{2}}{\frac{\theta}{2}} \to 1$$

as $\theta \to 0$.

(d) As an alternative formulation of Newton's proof, let $K = \overline{AB}$, and let E be the point on AD where it meets the perpendicular at B to RD. Using the analysis of angles in Exercise 16 in § 4.2, show that

$$\overline{AE} = \overline{EB} = \frac{K}{2} \sec \frac{\theta}{2}$$
$$\overline{ED} = \frac{K}{2} \sec \frac{\theta}{2} \sec \theta$$

and use this to show that

$$\overline{AB}/\overline{AD} \to 1.$$

Also, use Archimedes' inequality for convex curves (see Exercise 11 in § 5.1) to show that

$$\overline{AB} \leq \overset{\frown}{AB} \leq \overline{AD}$$

so that the other ratios go to 1 by the Squeeze Theorem.

3.5 Discontinuities

Using the notion of the limit of a function at a point, we can "localize" the notion of continuity.

Definition 3.5.1. *Given a function f and a point $a \in \mathbb{R}$,*

1. *f is* ***continuous at $x = a$*** *if $a \in \text{dom}(f)$ and either a is isolated in $\text{dom}(f)$ or $\lim_{x \to a} f(x) = f(a)$.*

2. *f has a* ***discontinuity at $x = a$*** *if (a is an accumulation point of $\text{dom}(f)$ but) f is not continuous at $x = a$.*

Essentially, f has a discontinuity if the equation

$$f(a) = \lim_{x \to a} f(x)$$

fails to hold. This can happen in several ways:

- if
$$\lim_{x \to a} f(x) = L \in \mathbb{R}$$
exists but $f(a)$ either fails to exist ($a \notin \text{dom}(f)$) or exists but is different from L, we say f has a **removable discontinuity** at $x = a$. In this case, a simple redefinition at $x = a$
$$g(x) = \begin{cases} f(x) & \text{if } x \neq a \\ L & \text{if } x = a \end{cases}$$
yields a function which is continuous at $x = a$ (and agrees with f everywhere else);

- if $\lim_{x \to a} f(x)$ *fails to exist*, then f has an **essential discontinuity** at $x = a$. In particular,
  - if $f(x)$ diverges to $\pm\infty$ as x approaches a from one side (or the other), we say that f **blows up** at $x = a$;
  - if the one-sided limits at $x = a$ both exist but are unequal, we say that f has a **jump discontinuity** at $x = a$.

A common type of removable discontinuity is illustrated by the rational function

$$f(x) = \frac{x^2 - 3x + 2}{x^2 - 1}$$

at $x = 1$. The function is undefined at $x = \pm 1$, since the denominator is zero. However, if we write $x = 1 + t$ and let $t \to 0$, we see that

$$\lim_{x\to 1} f(x) = \lim_{t\to 0} f(1+t) = \lim_{t\to 0} \frac{(t+1)^2 - 3(t+1) + 2}{(t+1)^2 - 1}$$
$$= \lim_{t\to 0} \frac{t^2 + 2t + 1 - 3t - 3 + 2}{t^2 + 2t + 1 - 1} = \lim_{t\to 0} \frac{t^2 - t}{t^2 + 2t}.$$

In this last expression, as long as $t \neq 0$, we can divide top and bottom by t to obtain

$$\lim_{t\to 0} f(1+t) = \lim_{t\to 0} \frac{t-1}{t+2} = \frac{0-1}{0+2} = -\frac{1}{2}.$$

It follows that the function

$$g(x) = \begin{cases} \frac{x^2-3x+2}{x^2-1} & \text{if } x \neq \pm 1 \\ -\frac{1}{2} & \text{if } x = 1 \end{cases}$$

has $\text{dom}(g) = \{x \,|\, x \neq -1\} = (-\infty, -1) \cup (-1, \infty)$, $g(x) = f(x)$ for $x \in \text{dom}(f)$, and is continuous. Of course, when $x \neq 1$, we can factor the numerator and denominator of $f(x)$:

$$\frac{x^2 - 3x + 2}{x^2 - 1} = \frac{(x-1)(x-2)}{(x-1)(x+1)} \underset{x\neq 1}{=} \frac{x-2}{x+1}$$

and in fact this formula also defines $g(x)$.
At $x = -1$ we have a different situation: the numerator approaches

$$\lim_{x\to -1} x^2 - 3x + 2 = 6$$

while the denominator approaches zero: for x slightly *below* -1, this gives $f(x)$ large *negative*, while for x slightly *above* -1, $f(x)$ is large *positive.* Thus

$$\lim_{x\to -1^-} f(x) = -\infty, \quad \lim_{x\to -1^+} f(x) = +\infty$$

and f blows up at $x = -1$. This is typical for rational functions (Exercise 5).
Another example is given by the function

$$f(x) = \frac{\sin x}{x}.$$

This is defined except at $x = 0$. But we know from Lemma 3.4.8 that

$$\lim_{x\to 0} f(x) = 1.$$

Hence we have a removable singularity, and

$$g(x) = \begin{cases} \frac{\sin x}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$$

is continuous on $(-\infty, \infty)$. . An example of a jump discontinuity is the behavior of

$$f(x) = \frac{|x|}{x}$$

at $x = 0$: we know that

$$\lim_{x \to 0^+} f(x) = 1, \quad \lim_{x \to 0^-} f(x) = -1.$$

There is *no* "definition" for $f(0)$ that would yield continuity: the discontinuity is *essential*.
An interesting, and sometimes useful, observation in this connection concerns monotone functions.

Remark 3.5.2. *Suppose f is strictly increasing on the interval (a, b). Then*

$$\lim_{x \to a^+} f(x) = \inf_{x \in (a,b)} f(x)$$
$$\lim_{x \to b^-} f(x) = \sup_{x \in (a,b)} f(x)\,.$$

To see the first statement, simply note that if we take any sequence $\{x_k\}$ decreasing monotonically to a, then $\{f(x_k)\}$ is strictly decreasing, so either $f(x_k) \to -\infty$, which then is the infimum of $f(x)$, or if f is bounded below, then $f(x_k) \to L$, which we claim is also equal to $\inf_{(a,b)} f(x)$. This is because for any $x \in (a, b)$ there must exist some $x_k < x$, and hence $f(x_k) < f(x)$, showing that L is a lower bound for $f(x)$ on (a, b) which is also a limit of values for f, hence $L = \inf_{(a,b)} f(x)$ by Lemma 2.5.5(2.5.5). The second statement follows by analogous reasoning.
Finally, note that the function

$$f(x) = \sin \frac{1}{x}$$

(see Figure 3.16 on p. 131) also has an essential discontinuity at $x = 0$, but it neither blows up nor "jumps" there: the one-sided limits at $x = 0$ do not exist.

Exercises for § 3.5

Answers to Exercises 1acegikmoq, 2, 3ace, 4-5c, 7 are given in Appendix B.

Practice problems:

1. For each function below, (i) determine the points of discontinuity; (ii) classify each as a removable discontinuity, a point where the function blows up, or a jump discontinuity; (iii) for each removable discontinuity $x = a$, find the value of $f(a)$ making f continuous there; (iv) for each essential discontinuity, determine the two one-sided "limits" of $f(x)$ at $x = a$ (including divergence to $\pm\infty$):

   (a) $f(x) = \dfrac{1}{x-1}$ (b) $f(x) = \dfrac{1}{x^2-1}$ (c) $f(x) = \dfrac{1}{x^3-1}$

   (d) $f(x) = \dfrac{x-1}{x^2-1}$ (e) $f(x) = \dfrac{x^2-1}{x^3-1}$ (f) $f(x) = \dfrac{x-1}{x^3-1}$

   (g) $f(x) = \dfrac{(x-1)^2}{x^2-1}$ (h) $f(x) = \dfrac{x^2-1}{(x-1)^2}$ (i) $f(x) = \dfrac{x}{x^2-1}$

   (j) $f(x) = \dfrac{x}{|x|+1}$ (k) $f(x) = \dfrac{x}{|x|-1}$ (l) $f(x) = \dfrac{x}{|x|}$

   (m) $\dfrac{\sin x}{x}$ (n) $\dfrac{\cos x}{x}$ (o) $\dfrac{1-\cos^2 x}{\sin x}$

   (p) $f(x) = \begin{cases} x^2+1 & \text{for } x < 1, \\ 2x & \text{for } x > 1. \end{cases}$

   (q) $f(x) = \begin{cases} x^2-1 & \text{for } x < 1, \\ 2x & \text{for } x > 1. \end{cases}$

2. For each part below, give an example of a rational function $f(x) = \frac{p(x)}{q(x)}$ with a point $x = a$ where $q(a) = 0$ and

   (a) f has a removable discontinuity at $x = a$;

   (b) $\lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x) = \infty$

   (c) $\lim_{x\to a^-} f(x) = -\infty$, $\lim_{x\to a^+} f(x) = +\infty$

3. For each function below, find all values of the constant α for which f has a removable discontinuity at the "interface" point, or explain why no such value exists:

(a) $f(x) = \begin{cases} x^2 & \text{for } x < 0, \\ x + \alpha & \text{for } x > 0. \end{cases}$

(b) $f(x) = \begin{cases} \alpha x^2 & \text{for } x < 3, \\ x & \text{for } x > 3. \end{cases}$

(c) $f(x) = \begin{cases} x - 1 & \text{for } x < 1, \\ x^2 + \alpha x + 3 & \text{for } x > 1. \end{cases}$

(d) $f(x) = \begin{cases} x^2 + 1 & \text{for } x < 0, \\ \alpha x^2 & \text{for } x > 0. \end{cases}$

(e) $f(x) = \begin{cases} \sin \pi x & \text{for } x < 0, \\ \cos \alpha x & \text{for } x > 0. \end{cases}$

(f) $f(x) = \begin{cases} \sin \pi x & \text{for } x < 1, \\ \cos \alpha x & \text{for } x > 1. \end{cases}$

4. Consider the two-parameter family of functions

$$f_{\alpha,\beta}(x) = \begin{cases} x + \alpha & \text{for } x < -1 \text{ and } x > 2, \\ \beta x^2 & \text{for } -1 < x < 2. \end{cases}$$

   (a) Find all values of α and β for which $f_{\alpha,\beta}$ has a removable discontinuity at $x = -1$, or explain why no such values exist.

   (b) Find all values of α and β for which $f_{\alpha,\beta}$ has a removable discontinuity at $x = 2$, or explain why no such values exist.

   (c) Find all values of α and β for which $f_{\alpha,\beta}$ has removable discontinuities at *both* $x = -1$ *and* $x = 2$, or explain why no such values exist.

Theory problems:

5. Suppose $f(x) = \dfrac{p(x)}{q(x)}$ is a rational function.

   (a) Show that f has a removable discontinuity at $x = a$ precisely if $x = a$ is a zero of $p(x)$ with multiplicity m and of $q(x)$ with multiplicity n, where $m \geq n > 0$.

   (b) Show that if $q(a) = 0$ and $p(a) \neq 0$ then f blows up at $x = a$ (*i.e.*, each one-sided limit of $f(x)$ at $x = a$ is infinite).

(c) Determine general conditions on $p(x)$ and $q(x)$ at $x = a$ under which each of the following happens:

i. One of the one-sided limits of $f(x)$ at $x = a$ is $-\infty$ and the other is $+\infty$.

ii. $\lim_{x \to a} f(x) = \infty$.

iii. $\lim_{x \to a} f(x) = -\infty$.

6. Show that if f is defined and strictly increasing on the interval (a, b) then for any point $c \in (a, b)$ f is either continuous or has a jump discontinuity at $x = c$.

Challenge problem:

7. Consider the function f defined on the rational numbers in $(0, 1)$ that assigns to the rational number $x = \frac{p}{q}$ in lowest terms the value $f(x) = \frac{1}{q}$. Identify the essential discontinuities and the removable discontinuities of f in $[0, 1]$.

3.6 Exponentials and Logarithms

A spectacular application of the idea of "removing" discontinuities is provided by the exponential functions. We will concentrate on the function 2^x, but the relevance of the discussion for other "bases" (b^x) will be clear. What is the definition of the function

$$\exp_2(x) = 2^x?$$

For a positive integer n, we know that 2^n means "multiply together n copies of 2", so

$$\exp_2(1) = 2^1 = 2, \ \exp_2(2) = 2^2 = 2 \cdot 2 = 4, \ \exp_2(3) = 2^3 = 2 \cdot 2 \cdot 2 = 8,$$

and so on. For these values, we also know that $\exp_2(x)$ turns sums into products

$$\exp_2(m + n) = 2^{m+n} = 2^m \cdot 2^n = \exp_2(m) \exp_2(n)$$

which easily leads to the only reasonable definition of $\exp_2(x)$ for *negative* integers and zero:

$$\exp_2(-n) = 2^{-n} = \frac{1}{2^n} = \frac{1}{\exp_2(x)}, \ \exp_2(0) = 2^0 = 1.$$

We also know that a power of a power is the same as a product power:

$$\big(\exp_2(m)\big)^n = (2^m)^n = 2^{mn} = \exp_2(mn).$$

This in turn leads to the only reasonable definition of $\exp_2(x)$ for *rational* numbers, namely, since we must have

$$\big(\exp_2(\frac{1}{q})\big)^q = \exp_2(1) = 2$$

it follows that we want

$$\exp_2(\frac{1}{q}) = \sqrt[q]{2},$$

and then

$$\exp_2(\frac{p}{q}) = \big(\exp_2(\frac{1}{q})\big)^p = \Big(\sqrt[q]{2}\Big)^p = \sqrt[q]{2^p} = (\exp_2(p))^{\frac{1}{q}}.$$

So we have defined a function $\exp_2(x)$ whose domain is the set $\mathbb{Q}$ of *rational* numbers. We collect together the useful properties which can be established (for $x, y \in \mathbb{Q}$) from the definition, and our earlier results on arithmetic:

Proposition 3.6.1 (Properties of 2^x for x rational)**.** *The function* $\exp_2$ *defined for rational numbers* $\frac{p}{q}$ *by*

$$\exp_2(\frac{p}{q}) = \sqrt[q]{p}$$

satisfies the following (x, y are understood to be rational):

1. $\exp_2(0) = 1$, $\exp_2(1) = 2$, *and* $\exp_2(x) > 0$
2. $\exp_2(x + y) = \exp_2(x)\exp_2(y)$
3. $\exp_2(x - y) = \frac{\exp_2(x)}{\exp_2(y)}$
4. $\exp_2(xy) = (\exp_2(x))^y = (\exp_2(y))^x$
5. $\exp_2$ *is strictly increasing: if* $x < y$ *then* $\exp_2(x) < \exp_2(y)$
6. $\lim \exp_2(\frac{1}{k}) = 1$.

The proof of this is a fairly straightforward matter of chasing definitions; we leave it as Exercise 3.

Now, we would like to extend the definition of $\exp_2(x)$ to x *irrational*. The idea would be to make $\exp_2$ continuous, so we want to have

$$\exp_2(x) = \lim \exp_2(x_k) \quad \text{whenever } x_k \to x.$$

Since *every* real number can be approximated by *rationals* (for example, its decimal expansion), this gives a scheme for defining $\exp_2$ at *any* number: merely pick a sequence x_k of rationals converging to x, and decree that $\exp_2(x)$ is to be the limit of $\exp_2(x_k)$. However, such a scheme presents three potential problems:

1. does the limit $\lim \exp_2(x_k)$ always exist (for $x_k \to x$)?
2. do *different* sequences $x_k \to x$ yield the *same* value for $\lim \exp_2(x_k)$?
3. in case x itself is rational, does this definition give the old value of $\exp_2(x)$?

We will see that all of these questions have the correct answer: if x_k is a convergent sequence of rationals, then (1) $\lim \exp_2(x_k)$ exists, and (2) depends only on the value of $x = \lim x_k$, and (3) if x happens to be rational, then $\exp_2(x) = \lim \exp_2(x_k)$.

We will start with what seems like a very special case of (1), namely $x_k \to 0$.

Lemma 3.6.2. *If x_k is a sequence of rational numbers converging to 0, then*

$$\exp_2(x_k) \to 1 = \exp_2(0).$$

Proof. We know that $\{\frac{1}{k}\}$ and $\{\exp_2(\frac{1}{k})\}=\{\sqrt[k]{2}\}$ are strictly decreasing sequences, with

$$\frac{1}{k} \to 0, \quad \exp_2(\frac{1}{k}) \to 1,$$

and we easily conclude that $\{-\frac{1}{k}\}$ and $\{\exp_2(-\frac{1}{k})\}=\{\sqrt[k]{\frac{1}{2}}\}$ are strictly increasing, with

$$-\frac{1}{k} \to 0, \quad \exp_2(-\frac{1}{k}) \to 1.$$

It follows that, given $\varepsilon > 0$ there exists N such that

$$\left|\exp_2(-\frac{1}{N})\right| = 1 - \frac{1}{\sqrt[N]{2}} < \varepsilon$$

$$\left|\exp_2(\frac{1}{N})\right| = \frac{1}{\sqrt[N]{2}} - 1 < \varepsilon.$$

Now, suppose x_k is any sequence of rationals with $x_k \to 0$. Eventually, we must have

$$-\frac{1}{N} < x_k < \frac{1}{N}.$$

But since $\exp_2(x)$ is an increasing function, this implies

$$\frac{1}{\sqrt[N]{2}} = \exp_2(-\frac{1}{N}) < \exp_2(x_k) < \exp_2(\frac{1}{N}) = \sqrt[N]{2}.$$

From our choice of N, this gives

$$|\exp_2(x_k) - 1| < \varepsilon.$$

Since this happens eventually for *every* $\varepsilon > 0$, we have

$$\lim \exp_2(x_k) = 1$$

whenever $x_k \to 0$, as required. □

Now, we can use Lemma 3.6.2 together with Proposition 3.6.1 to answer our questions. Suppose first that we have two sequences x_k, y_k of rationals with $\lim x_k = \lim y_k$. This implies that

$$\lim(x_k - y_k) = 0.$$

But then we know from Lemma 3.6.2 that

$$\lim \exp_2(x_k - y_k) = 1.$$

However, we also have

$$\exp_2(x_k - y_k) = \frac{\exp_2(x_k)}{\exp_2(y_k)},$$

so

$$\lim \frac{\exp_2(x_k)}{\exp_2(y_k)} = 1.$$

From this we see that, if either of the sequences $\exp_2(x_k)$ or $\exp_2(y_k)$ converges, so does the other, and their limits agree. In other words, if *some* sequence of rationals with $x_k \to x$ has $\exp_2(x_k)$ convergent, then *every* sequence with $y_k \to x$ has $\exp_2(y_k)$ converging to the same limit. Now, the decimal expression for any x gives a *monotone* sequence x_k converging to x, and if M (*resp.* m) are integers with $m < x < M$, then eventually $\exp_2(m) < \exp_2(x_k) < \exp_2(M)$, so x_k is (eventually) bounded, hence convergent. Thus, the limit

$$\lim \exp_2(x_k)$$

exists, and is the same, for *every* sequence $x_k \to x$. Finally, if x is rational, one sequence x_k converging to x is the constant sequence, $x_k = x$ for all k, and clearly this sequence satisfies $\exp_2(x_k) = \exp_2(x) \to \exp_2(x)$.
These arguments together tell us that the definition of 2^x via

$$2^x = \lim \exp_2(\frac{p_k}{q_k}), \quad \text{where } \frac{p_k}{q_k} \text{ is any rational sequence with } x = \lim \frac{p_k}{q_k} \tag{3.8}$$

gives a well-defined function $f(x) = 2^x$ with domain $(-\infty, \infty)$ and all of the properties mentioned earlier, and which is automatically continuous, as well (Exercise 4).
These arguments apply as well to any functon of the form

$$\exp_b(x) = b^x$$

provided the "base" b is a positive number.

Proposition 3.6.3. *Suppose $b > 0$ is a fixed positive number. Define a function $\exp_b$ with domain the rationals by*

$$\exp_b(\frac{p}{q}) = \sqrt[q]{b^p} = \left(\sqrt[q]{b}\right)^p.$$

Then $\exp_b$ has the following properties (x, y are rational):

1. $\exp_b(0) = 1$, $\exp_b(1) = b$, *and* $\exp_b(x) > 0$ *for all* $x \in \mathbb{Q}$;
2. $\exp_b$ *is monotone:*
   (a) *if* $b > 1$, *then* $x < y \Rightarrow b^x < b^y$;
   (b) *if* $b = 1$, *then* $b^x = 1$ *for all* x;
   (c) *if* $0 < b < 1$, *then* $x < y \Rightarrow b^x > b^y$;
3. $\exp_b(x + y) = \exp_b(x)\exp_b(y)$, $\exp_b(x - y) = \frac{\exp_b(x)}{\exp_b(y)}$;
4. $\exp_b(xy) = (\exp_b(x))^y = (\exp_b(y))^x$.

At each irrational *number x,* $\exp_b$ *has a removable singularity: if we define f by*

$$f(x) = \lim \exp_b(x_k), \quad \textit{where } x_k \textit{ is any } \text{rational } \textit{sequence converging to } x$$

then f is well-defined, continuous, $f(x) = \exp_b(x)$ for every rational x, and the properties above hold for f in place of $\exp_b$ *and x, y any real numbers.*

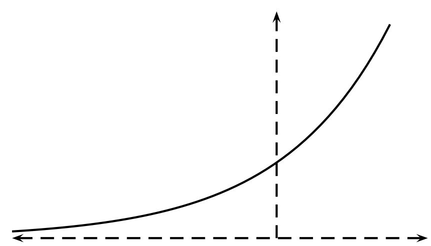


Figure 3.25: $y = 2^x$

We will henceforth use the (usual) notation b^x in place of $\exp_b(x)$. In this notation, we have

- $b^0 = 1$, $b^1 = b$, and $b^x > 0$ for all x
- $b^{x_1+x_2} = b^{x_1}b^{x_2}$, $b^{x_1-x_2} = \frac{b^{x_1}}{b^{x_2}}$, $b^{x_1x_2} = (b^{x_1})^{x_2}$.

Now, when $b = 10$, you are probably familiar with the logarithm function, which assigns to each number $x > 0$ the power of 10 giving x. Such a definition is possible for any "base" $b > 0$, except for $b = 1$ (for which the function b^x has the *constant* value 1).

Definition 3.6.4. *For $b > 0$ ($b \neq 1$) and $x > 0$, define the* ***logarithm with base b*** *of x by*

$$y = \log_b x \Leftrightarrow b^y = x.$$

Does this definition make sense? For $b > 1$ we argue as follows. Given any pair of positive integers $0 < k_1 < k_2$, we know that the function $f(t) = b^t$ is continuous on the interval $[-k_1, k_2]$, and by Proposition 3.6.3(2a) it is strictly increasing there. Applying Proposition 3.2.5 with $a = -k_1$ and $b = k_2$ (and $A = f(-k_1)$, $B = f(k_2)$), we see that the inverse function $g = f^{-1}$ is defined and continuous on $[f(-k_1), f(k_2)]$. But the definition of

$g(x)$ is precisely the same as our definition of $\log_b x$ in Definition 3.6.4. This shows that the function $\log_b$ is defined and continuous on any closed interval with positive integer endpoints. But we also know that $b^k \to \infty$ and $b^{-k} \to 0$, so given $x > 0$ there are positive integers k_1 and k_2 with

$$b^{-k_1} < x < b^{k_2}.$$

Thus, $\log_b x$ is defined for any $x \in (0, \infty)$, and is continuous there.

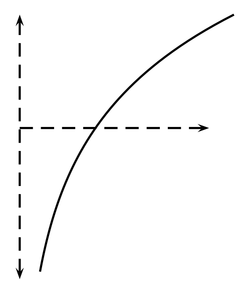


Figure 3.26: $y = \log_2 x$

For $0 < b < 1$, one can modify this argument slightly (Exercise 5) or use an algebraic trick (Exercise 6) to justify the definition.

Remark 3.6.5. *The function* $\log_b$ *defined above on* $(0, \infty)$ *is continuous and has the following properties:*

1. $\log_b 1 = 0$, $\log_b b = 1$
2. $\log_b(x_1 x_2) = \log_b x_1 + \log_b x_2$
3. $\log_b(\frac{x_1}{x_2}) = \log_b x_1 - \log_b x_2$
4. $\log_b(x_1^{x_2}) = x_2 \log_b x_1$.

Proof. These come from chasing definitions. We prove one, leaving the rest to you (Exercise 7). Suppose

$$y_1 = \log_b x_1$$
$$y_2 = \log_b x_2$$

that is,

$$x_1 = b^{y_1}, \ x_2 = b^{y_2}.$$

Then by the properties of b^t, $y = y_1 + y_2$ satisfies

$$b^y = b^{y_1+y_2} = b^{y_1}b^{y_2} = x_1x_2;$$

since $y = y_1 + y_2$ satisfies the definition of $\log_b(x_1x_2)$, it *is* $\log_b(x_1x_2)$. $\square$

We note also that for $b > 1$ (*resp.* $0 < b < 1$), the function $f(x) = \log_b x$ is strictly increasing (*resp.* strictly decreasing) (Exercise 8).

The development we give here of the exponential and logarithm functions is almost exactly the reverse of the historical development of these concepts. The properties we associate with logarithms were discovered by John Napier (1550-1617), a Scottish laird who published a set of tables in 1614 based on comparing a geometric series with an arithmetic one [20, pp. 148-153]. Henry Briggs (1561-1631), after visiting Napier in Scotland in 1615, published tables of logarithms based on 10. In 1647, the Belgian Jesuit Gregory of St. Vincent (1584-1667) noticed that the areas under the hyperbola $xy = 1$ between successive points in geometric progression are equal, and in 1649 his friend and pupil Alfons A. de Sarasa (1618-1667) pointed out that this meant that the area "acts like a logarithm". Newton also noted this connection, using his binomial series. In 1668, Nicolaus Mercator (1620-1687) also gave a series for the logarithm (Exercise 49 in Chapter 6).

The exponential function was first introduced by Euler in his *Introductio* (1748) [22, Chap. 6-7] and he defined the (natural) logarithm as the inverse of the exponential function. He derived power series for both functions, defined the number e as the number whose logarithm is 1, and showed that it is given both by an infinite series and the limit of instantaneously compound interest (Exercise 50 in § 6.5).

Exercises for § 3.6

Answers to Exercises 1ac, 2a are given in Appendix B.

Practice problems:

1. Simplify each expression below, using the definitions and Remark 3.6.5.

   (a) $\log_{\frac{1}{2}} 8$ (b) $\log_2 10 + \log_2 15 - \log_2 75$

   (c) $\log_2 \left(\left(\frac{1}{2} \right)^3 \right)$ (d) $\left(\frac{1}{2} \right)^{\log_2 x}$

2. (a) Is the function $f(x) = b^x$ bounded below? bounded above? (*Hint:* Your answers should depend on the value of b.)
   (b) What does this tell you about the function $\log_b$?

Theory problems:

3. Prove Proposition 3.6.1.
4. Show that the function defined by Equation (3.8) (Page 157)
   (a) is *well-defined* on $(-\infty, \infty)$,
   (b) is *strictly increasing* on $(-\infty, \infty)$,
   (c) is *continuous* on $(-\infty, \infty)$,
   (d) and satisfies all of the arithmetic properties given in Proposition 3.6.1 (without the restriction to rational inputs).
5. How can the *argument* following Definition 3.6.4 be modified to apply to $0 < b < 1$?
6. How can we use an *algebraic trick* to handle b^x for $0 < b < 1$, if we can already handle it for $b > 1$?

Challenge problems:

7. Show that for $b > 1$ the function b^x is strictly increasing, as follows:
   (a) Show that for any rational $x > 0$,
   $$b^x > 1.$$
   (b) Use this to show that if $x_2 - x_1$ is *rational* and positive, then
   $$b^{x_2} > b^{x_1}.$$
   (c) Use limits to show that if $x_1 < x_2$ (not necessarily rational), then
   $$b^{x_2} > b^{x_1}.$$

   How does this argument change to show b^x *decreasing* for $0 < b < 1$?
8. (a) Show that if $b > 1$ then the function $f(x) = \log_b x$ is strictly increasing. (*Hint:* Use Exercise 7)

(b) Show that if $0 < b < 1$ then the function $f(x) = \log_b x$ is strictly *decreasing*. (*Hint:* Use an algebraic trick and the first part of this problem.)

9. Assume that you know the inequality

$$e^x > x$$

holds for $0 \leq x \leq 2$. Use this to prove that it also holds for $2 < x \leq 4$. (You may use the fact that $e > 2$.)

3.7 Epsilons and Deltas (Optional)

In this section we re-examine continuity and limits in terms of approximations.

We should point out (and those readers who have had a standard calculus course before this will already have noticed) that our definitions of continuity and limits for functions, based on sequences, are not the standard ones. We have taken that route because it is easier to comprehend; here, we will examine the standard "epsilon-delta" versions of these definitions, which were given by Bernhard Bolzano (1781-1848) in 1817 and by Augustin-Louis Cauchy (1789-1857) in 1821. These did not explicitly use "epsilon-delta" terminology, but as Judith Grabiner argues [26], it was more than implicit—in fact, this notation (literally: using the letters ε and δ) *was* used by Cauchy in his definition of derivative in the same work.

Recall our motivating example, the calculation of $\sqrt{2}$. We have a function $f(x) = \sqrt{x}$ defined at least for all x near the point $x_0 = 2$; in principle, the value $f(x_0) = \sqrt{2}$ is determined, but we want to nail it on the number line. There are other values of f which we know exactly, like $f(1) = 1$, $f(1.96) = 1.4$, and $f(1.9881) = 1.41$; none of these is *exactly* equal to $\sqrt{2}$, but we hope that, by using one of these "known" values $f(x)$ with x close to x_0, we will be making only a small error in locating $f(x_0) = \sqrt{2}$. This is embodied in the following

Definition 3.7.1. *Suppose f is a function defined at all points x near x_0; we say that f has* ***error controls*** *at x_0 if given any error bound $\varepsilon > 0$, no matter how small, it is possible to find a corresponding deviation allowance $\delta > 0$ so that* any *input x deviating from x_0 by less than δ is guaranteed to produce an output $f(x)$ which approximates $f(x_0)$ with error less than ε:*

$$|x - x_0| < \delta \textit{ implies } |f(x) - f(x_0)| < \varepsilon.$$

Definition 3.7.1 has a geometric interpretation (*cf* Figure 3.27). If we graph the function $y = f(x)$ over the points near x_0, then the set of points (x, y) satisfying the deviation estimate $|x - x_0| < \delta$ (*i.e.*, $x_0 - \delta < x < x_0 + \delta$) is the (open) vertical band between the lines $x = x_0 - \delta$ and $x = x_0 + \delta$; our definition requires that the part of the graph crossing this band lies between the horizontal lines $y = y_0 - \varepsilon$ and $y = y_0 + \varepsilon$, where $y_0 = f(x_0)$. In other words, it must lie inside the "box" centered at (x_0, y_0) with vertical edges δ units to either side and horizontal edges ε units above and below. We require that "squeezing" the box horizontally allows us to squeeze it vertically as well.

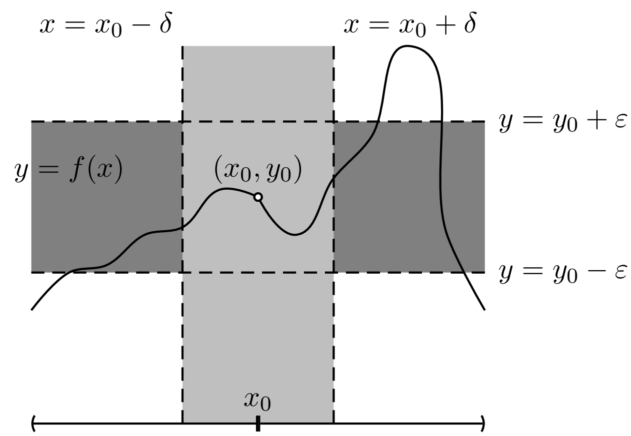


Figure 3.27: Error Controls

How do we establish error controls for our example? We are given an error bound $\varepsilon > 0$ and need to specify a deviation allowance $\delta > 0$ so that

$$|x - x_0| < \delta \text{ implies } |f(x) - f(x_0)| < \varepsilon.$$

An algebraic trick helps us here: since $x = (\sqrt{x})^2$ and $2 = \left(\sqrt{2}\right)^2$, we have

$$x - 2 = (\sqrt{x})^2 - \left(\sqrt{2}\right)^2 = \left(\sqrt{x} - \sqrt{2}\right)\left(\sqrt{x} + \sqrt{2}\right)$$

or

$$\left|\sqrt{x} - \sqrt{2}\right| = \frac{|x - 2|}{\sqrt{x} + \sqrt{2}}.$$

We can control the size of the numerator on the right, and we want to make sure that the quantity on the left is less than ε: if $|x - 2| < \delta$ then

$$\left|\sqrt{x} - \sqrt{2}\right| < \frac{\delta}{\sqrt{x} + \sqrt{2}}$$

and as long as we make sure the right side of this is $< \varepsilon$, we are OK. So, given $\varepsilon > 0$, we want to pick δ such that, when $|x - 2| < \delta$, we can guarantee

$$\frac{|x-2|}{\sqrt{x}+\sqrt{2}} \leq \varepsilon.$$

In other words, we want

$$\delta \leq \varepsilon\left(\sqrt{x}+\sqrt{2}\right).$$

This constitutes a kind of Catch-22, though: we don't *a priori* know x (never mind $\sqrt{x}$!) on the right, so we can't use it to determine δ. However, all we need is *some* choice of δ that works, and if we notice that, regardless of what x is (well, at least $x \geq 0$ to make sure $\sqrt{x}$ is defined), its square root is non-negative, so regardless of x

$$\varepsilon\sqrt{2} \leq \varepsilon\left(\sqrt{x}+\sqrt{2}\right).$$

This means that $\delta \leq \varepsilon\sqrt{2}$ is enough. *Hmm....*we don't actually *know* $\sqrt{2}$, but we want to use it to specify δ? *I don't think so.* But there is another trick we can use: true, we don't know $\sqrt{2}$ *exactly*, but we *do* know that it is bigger than 1, so we can be sure that

$$\varepsilon \leq \varepsilon\sqrt{2}.$$

Thus,

$$\delta \leq \varepsilon$$

should be good enough.

Let's turn this into a FORMAL PROOF: Given $\varepsilon > 0$, pick any δ satisfying

$$0 < \delta \leq \varepsilon.$$

Then if we know that

$$|x-2| < \delta,$$

we also know that

$$\left|\sqrt{x}-\sqrt{2}\right| = \frac{|x-2|}{\sqrt{x}+\sqrt{2}} < \frac{\delta}{\sqrt{x}+\sqrt{2}} \leq \frac{\varepsilon}{\sqrt{x}+\sqrt{2}}$$

but

$$\sqrt{x}+\sqrt{2} \geq \sqrt{2} > 1$$

so the last fraction on the right is $< \varepsilon$, as required. □

Notice that as with limits of sequences, we are not required to produce the *best* deviation allowance, just *some* deviation allowance for which we can *prove* that the desired bound on the error follows. Always take advantage of any tricks which, by making your deviation allowance stricter (*i.e.*, making δ smaller), allow you to prove the bound on the errror more easily. Also, don't be concerned with *large* values of ε—these are already taken care of whenever you can handle smaller ones. However, you need a scheme that lets you produce an effective deviation allowance $\delta > 0$ *no matter how small* an $\varepsilon > 0$ is thrown at you.

Continuity of f at x_0 in terms of convergent sequences (Definition 3.1.2) and error controls (Definition 3.7.1) both encode the idea that we can get a good approximation of $f(x_0)$ by calculating $f(x)$ for x close enough to x_0. In fact the two properties are equivalent.[17]

Proposition 3.7.2 (Error controls versus sequences)**.** *Suppose f is defined at all points in some open interval I containing x_0. Then f is continuous at x_0 precisely if it has error controls at x_0.*

Proof. We need to prove two assertions:

1. If f has error controls at x_0 then it is continuous at x_0.
2. If f is continuous at x_0, then it has error controls at x_0.

To prove (1), assume f has error controls at x_0; we need to show that whenever a sequence x_k in I converges to x_0 we must have

$$f(x_0) = \lim f(x_k)\,.$$

So let $x_k \to x_0$; given $\varepsilon > 0$, we need to find a place in the sequence, say K, beyond which (*i.e.*, for all $k > K$) we can guarantee $|f(x_k) - f(x_0)| < \varepsilon$. Using the error controls, pick $\delta > 0$ such that

$$|f(x) - f(x_0)| < \varepsilon \text{ whenever } |x - x_0| < \delta.$$

Since $x_k \to x_0$, there is a place K beyond which we have $|x_k - x_0| < \delta$; but then the error controls with $x = x_k$ guarantee that for all $k > K$ we also have

$$|f(x_k) - f(x_0)| < \varepsilon \text{ for all } k > K$$

[17] What we call "error controls" is the standard "epsilon-delta" definition of continuity at x_0.

as required.
The proof of (2) is less direct. From knowledge that $x_k \to x_0$ implies $f(x_k) \to f(x_0)$ we need to produce error controls at x_0; we need to produce a deviation allowance $\delta > 0$ independent of any sequence. It is hard to see how to do this directly, so instead we argue by contradiction. Suppose for some error bound $\varepsilon > 0$ there is no effective deviation allowance $\delta > 0$: this means that no matter what $\delta > 0$ we try, there is always *some* point with $|x - x_0| < \delta$ but $|f(x) - f(x_0)| \geq \varepsilon$. In particular, taking $\delta = \frac{1}{k}$ for $k = 1, 2, ...$, we can find a sequence of points x_k with $|x_k - x_0| < \frac{1}{k}$ but $f(x_k) - f(x_0) \geq \varepsilon$. Of course, the first of these conditions says that $x_k \to x_0$, and so, since f is assumed to be continuous at x_0, we must have $f(x_k) \to f(x_0)$. But this means that eventually $|f(x_k) - f(x_0)| < \varepsilon$, contrary to our second condition on the x_k's, and so it is impossible that there are no error controls—in other words, f *has* error controls at x_0.

□

There are also "error control" characterizations of limits:

$$L = \lim_{x \to x_0} f(x)$$

encodes the idea that for $x \in \text{dom}(f)$ near x_0, $f(x)$ is near L. Remember that here, we are not interested in $x = x_0$.

Lemma 3.7.3. *Suppose x_0 is an accumulation point of* $\text{dom}(f)$ *and $L \in \mathbb{R}$. Then $L = \lim_{x \to x_0} f(x)$ precisely if given any error bound $\varepsilon > 0$ there exists a deviation allowance $\delta > 0$ so that, whenever $x \in \text{dom}(f)$ satisfies*

$$0 < |x - x_0| < \delta$$

we can guarantee the bound on the error

$$|f(x) - L| < \varepsilon.$$

The proof of this is very similar to that of Proposition 3.7.2; we leave it to you (Exercise 5).
The variants of this when x_0 or L is infinite can also be formulated in terms of error control, provided that whenever we would be required to write a nonsense statement like $|x - \infty| < \delta$, we replace it with the statement $x > \Delta$ (where, instead of imagining $\delta > 0$ to be *small*, we imagine $\Delta > 0$ to be *large*). For example

Lemma 3.7.4. *Suppose f is defined for all $x \in (a, \infty)$. Then*

$$\lim_{x \to +\infty} f(x) = L \in \mathbb{R}$$

precisely if, given an error bound $\varepsilon > 0$, there exists a deviation allowance $\Delta \in \mathbb{R}$ such that

$$x > \Delta \Rightarrow |f(x) - L| < \varepsilon.$$

Again we leave the proof of this, and the "error control" characterization of divergence to infinity as $x \to x_0$, to you (Exercises 6 and 7).
The alert reader may wonder why we have concentrated in our discussion on limits and continuity at one point: what about continuity on an interval? Well, here things get more involved. We saw (Remark 3.4.3) that a function f is continuous on and interval I if it is continuous at every point of the interval; thus it is continuous on I if we have error controls at every point of I. But this involves a subtle complication: *a priori*, for a given error bound $\varepsilon > 0$, we can only expect to pick a derivation allowance $\delta > 0$ once we know the basepoint x_0 from which we are measuring the deviation—so in general, we expect to produce $\delta > 0$ which depends on *both* ε and x_0.
For example, we found error controls for $f(x) = \sqrt{x}$ at $x_0 = 2$ by picking $0 < \delta \leq \varepsilon$, but if we tried this at $x_0 = 0.01$, where $f(x_0) = 0.1$, we would find that, for example, with $\varepsilon = 0.005$ and $x = 0.0121$, we have $|x - x_0| = 0.0021 < \varepsilon$ but $|f(x) - f(x_0)| = 0.01 > \varepsilon$.
Even worse is the situation of $f(x) = \frac{1}{x}$ on the interval $(0, 1)$; it is continuous at every point, but for any fixed deviation allowance $\delta > 0$, if we try to use it at $x_0 = \frac{1}{N}$, where N is an integer with $\frac{1}{2N} < \delta$, then $x = \frac{1}{2N}$ satisfies $|x - x_0| < \delta$ but $|f(x) - f(x_0)| = N$, which can be arbitrarily *large*.
There are times, however, when this difficulty does not arise. Let us formulate what this means.

Definition 3.7.5. *Suppose f is defined at all points in some interval I. We say that f is* ***uniformly continuous*** *on I if we can use the same error controls at every point of I: that is, given any error bound $\varepsilon > 0$, no matter how small, it is possible to find a deviation allowance $\delta > 0$ so that whenever $x, x_0 \in I$ deviate by less than δ*

$$|x - x_0| < \delta$$

we can guarantee that the bound

$$|f(x) - f(x_0)| < \varepsilon$$

holds.

The point here is that δ can be picked *independent* of x_0. We have seen that $f(x) = \frac{1}{x}$ is *not* uniformly continuous on the interval $I = (0, 1)$, even though it is continuous at every point of I. However, the following positive result holds.

Theorem 3.7.6. *Suppose f is continuous at every point of a* ***bounded, closed*** *interval $I = [\alpha, \beta]$. Then f is uniformly continuous on I.*

The notion of uniform continuity was first formulated by Bernhard Georg Friedrich Riemann (1826-1866) in 1851, in connection with his theory of integration (see p. 343) and Johann Peter Gustav Lejeune Dirichlet (1805-1859), who proved Theorem 3.7.6 in lectures in Berlin in 1854. The formulation in terms of ε and δ was given by Karl Theodor Wilhelm Weierstrass (1815-1897) [8, p. 216]; the term "uniform continuity" was formulated by Eduard Heine (1821-1881), a student of Weierstrass, in a pair of papers written in 1871 and 1874 [5, pp. 14, 23-6], where a generalization of Theorem 3.7.6 to higher dimensions was proved. While the concept of uniform continuity is for the most part not needed in this book (with the notable exception of Proposition 5.2.3), it plays an important role in the more advanced study of functions. Augustin-Louis Cauchy (1789-1857) in some of his proofs implicitly (and unknowingly) used subtle assumptions of uniform continuity; however, it appears that in a paper published in 1854 (a year before he died) he finally came to grips with this concept.

Proof. We proceed by contradiction: suppose f is continuous but not uniformly continuous on $I = [a, b]$. This means that for some $\varepsilon > 0$, we cannot find a single deviation allowance $\delta > 0$ which works at *every* point of I; in other words, for any $\delta > 0$ we can find two points $x, x' \in I$ with $|x - x'| < \delta$ but $|f(x) - f(x')| \geq \varepsilon$. Taking $\delta = \frac{1}{k}$ for $k = 1, 2, ...$, this gives us two sequences $x_k, x'_k \in I$ such that

$$\left|x_k - x'_k\right| < \frac{1}{k}$$

but

$$\left|f(x_k) - f\left(x'_k\right)\right| \geq \varepsilon.$$

By the Bolzano-Weierstrass Theorem (Proposition 2.3.8), we can find a convergent subsequence of $\{x_k\}$, say $y_i = x_{k_i}$; set $y'_i = x'_{k_i}$. Let

$$y = \lim y_i.$$

Since I is closed, $y \in I$; we claim that also

$$y = \lim y_i'.$$

To see this, given a new error bound[18] $\alpha > 0$, let K be a place such that for $i > K$

$$|y_i - y| < \frac{\alpha}{2};$$

furthermore, by upping K we can assume that for all $i > K$ also

$$\frac{1}{k_i} < \frac{\alpha}{2}.$$

But then the triangle inequality says that for all $i > K$

$$\left|y_i' - y\right| \leq \left|y_i' - y_i\right| + |y_i - y| < \frac{1}{k_i} + \frac{\alpha}{2} \leq \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha.$$

This shows that $\lim y_i = y = \lim y_i'$; since f is continuous at y, we have

$$\lim f(y_i) = f(y) = \lim f\left(y_i'\right)$$

and in particular, for i sufficiently large,

$$\left|f(y_i) - f\left(y_i'\right)\right| \leq |f(y_i) - f(y)| + \left|f(y) - f\left(y_i'\right)\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

contradicting the original choice of x_k, x_k', which requires

$$\left|f(y_i) - f\left(y_i'\right)\right| = \left|f(x_{k_i}) - f\left(x_{k_i}'\right)\right| \geq \varepsilon.$$

□

As we noted above, the uniform continuity of continuous functions on closed bounded intervals will play a useful role in Chapter 5, especially in Proposition 5.2.3.

Exercises for § 3.7

Answers to Exercises 1c, 7ab are given in Appendix B.

Practice problems:

1. Consider the function $f(x) = x^2$.

[18] We have already used ε for a different error bound

(a) Suppose we want to get error controls for $f(x)$ at $x_0 = 1$ with $\varepsilon = 0.1$. Show that $\delta = 0.05$ is not sufficient, but $\delta = 0.04$ is.

(b) Now try to do the same thing at $x_0 = 2$: show that $\delta = 0.04$ is not sufficient, but $\delta = 0.02$ is.

(c) Theorem 3.7.6 says that we should be able to find one value of δ that works for $\varepsilon = 0.1$ at *every* point $x_0 \in [1, 2]$. Give such a value. (*Hint:* Where is it harder to control the error?)

2. Show directly that any function of the form $f(x) = mx + b$, $m > 0$, is uniformly continuous on $(-\infty, \infty)$, by showing that for any $\varepsilon > 0$ and any $x_0 \in \mathbb{R}$, the value $\delta = \dfrac{\varepsilon}{m}$ gives error controls at x_0.

3. Show that the function $f(x) = \frac{1}{x}$ is *not* uniformly continuous on $(0, 1)$. (*Hint:* For any $\delta > 0$, find a pair of points $x, x' \in (0, 1)$ for which $|x - x'| < \delta$ but $|f(x) - f(x')| > 1$.)

4. Show that the function $f(x) = \frac{1}{x}$ *is* uniformly continuous on $(1, \infty)$, by showing that the choice
$$\delta = \frac{\varepsilon}{2}$$
always works at any $x_0 > 1$, at least provided $0 < \varepsilon < 1$.

Theory problems:

5. Mimic the proof of Proposition 3.7.2 to establish Lemma 3.7.3.

6. Modify the proof of Lemma 3.7.3 to prove Lemma 3.7.4.

7. Formulate an "error control" version of the definition of

   (a) $\lim_{x \to a} f(x) = \infty$

   (b) $\lim_{x \to \infty} f(x) = \infty$

And what are these fluxions? The velocities of evanescent increments? And what are these same evanescent increments? They are neither finite quantities, nor quantities infinitely small, nor yet nothing. May we not call them the ghosts of departed quantities?

George Berkeley, Bishop of Cloyne
The Analyst (1734)[51, p. 338]

...these increments must be conceived to become continuously smaller, and in this way, their ratio is represented as continuously approaching a certain limit, which is finally attained when the increment becomes absolutely nothing. This limit, which is, as it were, the final ratio of those increments, is the true object of differential calculus.

Leonhard Euler
Institutiones Calculi Differentialis (1755) [23]

4 Differentiation

There are two classes of geometric problems treated by calculus: the determination of *tangents* is associated with derivatives, and the determination of *areas* with integrals. The standard practice in calculus texts (which we follow) is to look at derivatives first, but historically area problems were studied far more extensively and far earlier.

In Euclid and Archimedes, the notion of a line tangent to a curve was limited to circles and other *convex* curves; a line was tangent to such a curve if it touched it at exactly one point. (It is a tribute to the care that Archimedes took that he explicitly defined what it means for a curve to be convex, and set forth which properties he assumed followed from this definition [31, pp. 2-4].) A more general idea of tangent lines was given in Fermat's method for finding maxima and minima (Exercise 15 in § 4.7), which was primarily an algebraic algorithm, but supported at least in part by a concept similar to the modern one via secant lines. Descartes had a method for finding the tangent line by first finding the *normal* or perpendicular line (the *circle method* [20, pp. 125-7]) in connection with finding double roots of equations. Isaac Barrow (1630-1677), Newton's teacher and predecessor in the Lucasian Chair at Cambridge, found tangent lines via a method similar to Fermat's, using the "characteristic triangle" which we discuss below.

The approaches of Leibniz and Newton to tangents were quite different. Newton was coming at curves from a kinematic point of view; using the synthetic geometric language of Euclid, many of his tangent/derivative constructions consisted of taking a point moving toward another along a curve, and considering the chord between the two; he would then consider the positions of these chords and conclude that the "ultimate ratio" between these chords and the tangent line at the latter point was "that of equality". In general, his curve was *parametrized*: that is, it was generated by a combination of motions for two quantities, which we would label x (the abcissa) and y (the ordinate); he would consider their "fluxions" $\dot{x}$ and $\dot{y}$ (in our terms, the rates of change of the two quantities) and conclude that the tangent line has slope given by $\frac{\dot{y}}{\dot{x}}$.

Leibniz, on the other hand, thought in terms of a "characteristic triangle" whose sides were "infinitesimal" (ie, of vanishing length): this was a right triangle whose horizontal leg was the "differenital" dx, the vertical leg was dy, and the hypotenuse was ds. These were to represent the infinitesimal increments of the abcissa, ordinate, and of the position along the curve. He would quite happily use the similarity between this triangle and various more conventional ones to deduce results about the tangent and related items. Our modern view combines some of Newton's kinematic vision with Leibniz's differential notation, rejecting the idea of actually taking a ratio of the form $\frac{dy}{dx}$ as Leibniz might do, but considering the limiting value of the ratio. However, Leibniz's notation has survived in all aspects of calculus.

The first quotation at the head of this chapter is a famous passage from a pamphlet published in 1734 by George Berkeley (1685-1753), the Bishop of Cloyne. Berkeley attacked the reasoning behind the calculus of Newton, Leibniz and their colleagues to counter attacks on religious thinking as inexact; his point was that (even though the results *work*), the logical basis for the calculus was no more rigorous (no, less so) than the theological thinking of the time. His criticisms found their target; Newton among others tried to firm up the foundations of their thinking; the quotation from Euler is an indication of the ways this was attempted. But rigorous formulations of the basic ideas of limits, continuity and derivatives really only came in the nineteenth century.

The first careful formulations of the idea of derivative were given by Bernhard Bolzano (1781-1848) in 1817 and Augustin-Louis Cauchy (1789-1857) in 1821; we have already mentioned several times that Bolzano's work, while somewhat earlier, was not generally known. Cauchy's definition of the derivative is much the same as ours.

There was a presumption in the seventeenth through mid-nineteenth centuries that all continuous functions can be differentiated, with perhaps a few exceptional points (such as interface points for functions defined in pieces). In fact, Joseph Louis Lagrange (1736-1813) built a whole theory of functions around this assumption (using power series—see Chapter 6) [37]. At the end of this chapter, we will construct an example which shows that this assumption can fail spectacularly. The first such example was probably given by Bolzano in 1834, but went unnoticed; however, when Karl Theodor Wilhelm Weierstrass (1815-1897) constructed such an example in the 1860's, it had an effect analogous to the discovery of the irrationality of $\sqrt{2}$ in Hellenic times, and was a major push toward the development of a far more careful formulation of the ideas of continuity, limits and derivatives.

4.1 Slope, Speed and Tangents

The **graph** (or **locus**) of an equation in x and y is the collection of points in the plane whose coordinates (x, y) satisfy the equation. The simplest equations involve sums of terms which are constant multiples of x or y, or just constants

$$Ax + By = C; \tag{4.1}$$

these are called **linear equations**, because their graphs are straight lines. When A and B are both nonzero, then the line crosses each coordinate axis at a unique point: it crosses the x-axis (which is the set of points where $y = 0$) at the point $(\frac{C}{A}, 0)$ and the y-axis (where $x = 0$) at $(0, \frac{C}{B})$; $\frac{C}{A}$ (*resp.* $\frac{C}{B}$) is called the ***x*-intercept** (*resp.* ***y*-intercept**) of the line.

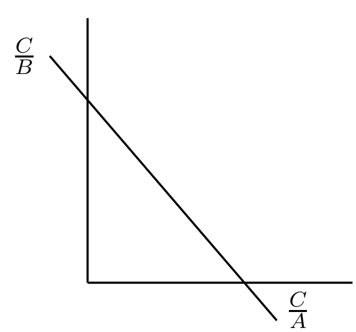


Figure 4.1: $Ax + By = C$

Of course, an equation can be manipulated algebraically, either by multiplying both sides by the same non-zero number or adding the same term to both sides, without changing the graph: in this way (provided $B \neq 0$) we can rewrite Equation (4.1) as the definition of a polynomial function of degree one

$$y = f(x) := mx + b \tag{4.2}$$

where $m = -\frac{A}{B}$ and $b = \frac{C}{B}$. Technically, such a function is not called "linear" (even though its graph is a straight line) unless $b = 0$; in general it is called an **affine function**. The coefficients of this polynomial have geometric interpretations. You can see immediately that b is the y-intercept of the graph. The various lines going through the point $(0, b)$ (or any other given point, for that matter) are distinguished by their "tilt", which is determined by the first-order coefficient m. Geometrically, if we compare the point $(0, b)$ (corresponding to $f(x)$ at $x = 0$) and the point corresponding to another, higher value of x, we see that the value $f(a) = ma + b$ at $x = a$ is precisely ma units higher than that at $x = 0$ (assuming $m > 0$; otherwise we adopt the convention that going "up" (*resp.* "right") a *negative* number of units is the same as going "down" (*resp.* "left") by its absolute value). Thus, m represents the ratio between the change, or **increment** $\triangle y$ in the output $y = f(x)$ and the increment $\triangle x$ in the input x as we go from $x = 0$ to any other value of x. In fact, if we start at *any* input value $x = x_1$ (with corresponding output value $y_1 = f(x_1)$) and go to any *other* input $x = x_2$ (with $y = y_2 = f(x_2)$), the increment in x is

$$\triangle x := x_2 - x_1$$

and the corresponding increment in $y = f(x)$ is

$$\triangle y := y_2 - y_1 = f(x_2) - f(x_1)$$

(sometimes also denoted "$\triangle f$") so (when f is affine, given by Equation (4.2))

$$\triangle y = (mx_1 + b) - (mx_2 + b) = m(x_2 - x_1) = m\triangle x$$

and thus we have

$$m = \frac{\triangle y}{\triangle x}$$

for *any* pair of *distinct* inputs $x_1 \neq x_2$. The fraction on the right is pronounced "*delta y over delta x*", but another colloquial expression for it is "*rise over run*", reflecting the fact that $\triangle y$ corresponds to vertical displacement and $\triangle x$ to horizontal displacement along the line.

The value $m = \frac{\triangle y}{\triangle x}$ is called the **slope** of the line $y = mx + b$; geometrically,

$$m = \frac{\triangle y}{\triangle x} = \tan\theta \tag{4.3}$$

where θ is the angle between the line and any horizontal line (Figure 4.2): *positive* slope means the line tilts *up* to the right while *negative* slope

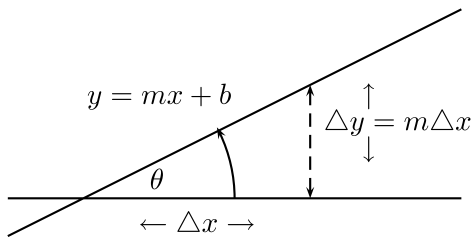


Figure 4.2: $m = \tan\theta$

means it tilts *down* to the right; a horizontal line is the graph of a constant function, with slope zero (Figure 4.3). It also follows that two lines are

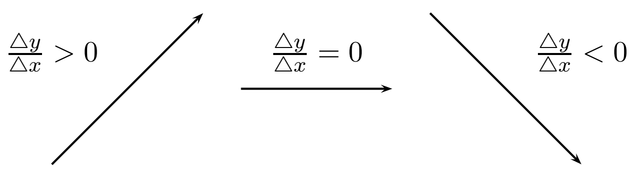


Figure 4.3: Sign of slope

parallel precisely if they have the same slope (Exercise 7).
The fact that $\frac{\triangle y}{\triangle x}$ is the same for any pair of points when y is an affine function of x can be used to write down the equation of the line through any pair of distinct points. For example, for $(x_1, y_1) = (-1, 3)$ and $(x_2, y_2) = (2, -3)$, the increments are

$$\begin{aligned} \triangle y &= y_2 - y_1 = (-3) - (3) = -6 \\ \triangle x &= x_2 - x_1 = 2 - (-1) \quad = 3 \end{aligned}$$

so the affine function $f(x) := mx + b$ with $f(-1) = 3$ and $f(2) = -3$ has

$$m = \frac{\triangle y}{\triangle x} = \frac{-6}{3} = -2;$$

then the fact that $f(-1) = 3$ means

$$2(-1) + b = 3$$

or

$$b = 5$$

so the equation of the given line is

$$y = -2x + 5.$$

(You should check that $(x_2, y_2) = (2, -3)$ also lies on this line.)
The slope $\frac{\triangle y}{\triangle x}$ can also be interpreted as the **rate of change** of $y = f(x)$ between $x = x_1$ and $x = x_2$: in our example, a (horizontal) change in x of 3 units leads to a (vertical) change in $y = -2x + 5$ of -6 units, so y changes -2 units for every unit of change in x.
A little thought, fortified by some calculation (Exercise 8), shows that affine functions (including constants) are the *only* functions with domain $(-\infty, \infty)$ for which $\frac{\triangle y}{\triangle x}$ is the same for any pair of points. For example if we carried out the corresponding calculations for

$$f(x) = 3x^2 - 2$$

with $x_1 = -1$ and $x_2 = 2$, we would get $y_1 = 3(-1)^2 = 1$, $y_2 = 3(2)^2 - 2 = 10$, so

$$\frac{\triangle y}{\triangle x} = \frac{9}{3} = 3$$

while if we changed x_2 to $x_2 = 1$, we would have $y_2 = 3(1)^2 - 2 = 1$, so

$$\begin{aligned} \triangle y &= y_2 - y_1 &= 1 - 1 &= 0 \\ \triangle x &= x_2 - x_1 &= 1 - (-1) &= 2 \end{aligned}$$

and

$$\frac{\triangle y}{\triangle x} = 0;$$

similarly, if $x_1 = 1$ and $x_2 = 2$, then

$$\begin{aligned} \triangle y &= 10 - 1 &= 9 \\ \triangle x &= 2 - 1 &= 1 \end{aligned}$$

so

$$\frac{\triangle y}{\triangle x} = 9.$$

How, then, do we formulate the rate of change of a non-affine function? First, we have to be more specific: clearly $y = 3x^2 - 2$ is *increasing* when x is close to $x = 2$ but *decreasing* when x is near $x = -1$, so we need to specify *where* we want to find the rate of change. Second, the quantity $\frac{\triangle y}{\triangle x}$ represents a kind of *average* rate of change: we can only say that, for example between $x = -1$ and $x = 2$ the function $y = 3x^2 - 2$ has a net increase of 9 units for a 3-unit increase in x, giving an average rate of change of 3 units increment in y per unit increment in x, while between $x = -1$ and $x = 1$ the function has no *net* change (of course, it does not stay constant in betweeen—it just happens to return to its initial value), so the *average* rate of change between $x = -1$ and $x = 1$ is zero.
If we try to formulate "how fast" a function is changing *at a particular point*, we need to resort to limits. The idea is that if, say, we want to know how fast $y = 3x^2 - 2$ is changing *at* $x = 1$, we work from the assumption that the *average* rate of change over a *small* interval containing $x = 1$ should approximate the "true" or **instantaneous rate of change** at $x = 1$, and the smaller the interval, the better the approximation. This leads to the notion of a derivative.

Definition 4.1.1. *Suppose the function f is defined for all input values near $x = a$. The* ***derivative*** *of $y = f(x)$ at $x = a$ is the limit defined as follows:*

- *for each input x near $x = a$, set*

$$\begin{aligned} \triangle x &= x - a \\ \triangle y &= f(x) - f(a) \end{aligned}$$

- *the derivative is then the limit*

$$\lim_{x \to a} \frac{\triangle y}{\triangle x} = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

While this idea was there in Newton's work, in the form of "ultimate ratios", it was not stated so clearly, particularly not in terms of functions (this notion, in fact, came later). The first statement of the derivative of a *function* in the terms above appears to have been by Jean le Rond D'Alembert (1717-1783) in an article in an encyclopedia (1754). [20, p. 295] has a translated quotation of d'Alembert's statement. Augustin-Louis Cauchy (1789-1857) gave a *rigorous* definition of derivative, more or less as above, in 1823 and used it along with his rigorous formulations of the notions of limit and integral to develop a careful theory of basic calculus.

Note that in this process, $\frac{\triangle y}{\triangle x}$ *does not make sense* when $x = a$. Also, notice that we could equally well have formulated this in terms of $\triangle x$ instead of x, using $x = a + \triangle x$:

$$\frac{\triangle y}{\triangle x} = \frac{f(a + \triangle x) - f(a)}{\triangle x}$$

and taken the limit as $\triangle x \to 0$.[1] Of course, as $\triangle x \to 0$, so[2] does $\triangle y$. A standard notational convention in calculus, which we shall see again in Chapter 5, says that when a quantity is defined in terms of increments of certain variables (like $\frac{\triangle y}{\triangle x}$), then the limiting value as all of these go to zero is denoted by replacing the Greek letter $\boldsymbol{\Delta}$ (denoting a concrete small increment, calculated from two distinct but close points) with the corresponding English[3] letter d. Thus, a standard notation for the derivative is

$$\frac{dy}{dx} := \lim_{\triangle x \to 0} \frac{\triangle y}{\triangle x}.$$

Note that the expression $\frac{dy}{dx}$ is *not* the ratio between two quantities called "dx" and "dy"; rather we look at the *whole expression* as the limit of the ratio $\frac{\triangle y}{\triangle x}$. If we want to explicitly indicate *where* the derivative is being taken, we write

$$\left.\frac{dy}{dx}\right|_{x=a} = \lim_{x \to a} \frac{\triangle y}{\triangle x};$$

also, when $y = f(x)$, we sometimes use "f" in place of "y"

$$\left.\frac{dy}{dx}\right|_{x=a} = \left.\frac{df}{dx}\right|_{x=a} = \frac{df}{dx}.$$

An entirely different notation (due to Lagrange), which indicates in a more graceful way the function being differentiated and the place where the differentiation is being performed, is

$$\left.\frac{df}{dx}\right|_{x=a} = f'(a)\,,$$

and finally, we sometimes write y' in place of $f'(a)$ when $y = f(x)$.

---

[1]It is common to replace the slightly awkward symbol $\triangle x$ with h in the definition of derivative, so one frequently sees the definition of the derivative written in the (clearly equivalent) form $f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$.

[2]at least if f is continuous

[3]or French, in the case of Cauchy

Let us examine this definition in case $f(x) = 3x^2 - 2$ and $a = 1$. We will think in terms of $\triangle x$. We have

$$\begin{aligned}\triangle y = \triangle f = f(1+\triangle x) - f(1) &= [3(1+\triangle x)^2 - 2] - [1] \\ &= 6\triangle x + \triangle x^2\end{aligned}$$

so

$$\frac{\triangle y}{\triangle x} = \frac{\triangle f}{\triangle x} = \frac{6\triangle x + \triangle x^2}{\triangle x} = \frac{(6+\triangle x)\triangle x}{\triangle x} = 6 + \triangle x$$

and clearly

$$y' = f'(1) = \frac{df}{dx} = \frac{dy}{dx} = \lim_{\triangle x \to 0} \frac{\triangle y}{\triangle x} = \lim_{\triangle x \to 0} (6 + \triangle x) = 6.$$

There is a geometric way of thinking about the process of differentiation. When we have two distinct points (x_1, y_1) and (x_2, y_2) on the graph of $y = f(x)$, the line joining them is sometimes called a **secant line**; the slope of this line is the ratio $\frac{\triangle y}{\triangle x}$ of increments calculated using this pair of points (Figure 4.4). To calculate the derivative $f'(a)$, we nail down the

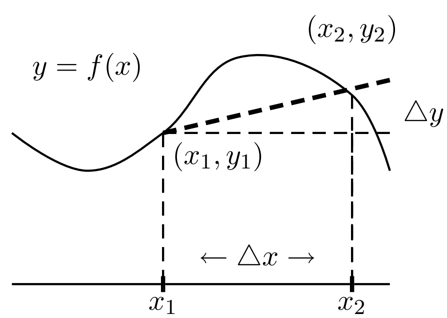


Figure 4.4: Secant Line

point (x_1, y_1) at $x_1 = a$ (and $y_1 = f(a)$) and consider the secant line joining this to the variable point $x_2 = a + \triangle x$ (and $y_2 = f(a + \triangle x) = y_1 + \triangle y$); now consider the limit as $\triangle x \to 0$ (*i.e.*, $x_2 \to x_1$). All the secant lines we consider in this process go through the point $(x_1, y_1) = (a, f(a))$, but as x_2 changes they rotate around that point: it is geometrically clear that as $\triangle x \to 0$ this rotation brings them ever closer to a limiting position, with slope $\frac{dy}{dx}$: this defines the "tangent line" (to the graph) at the point (Figure 4.5).

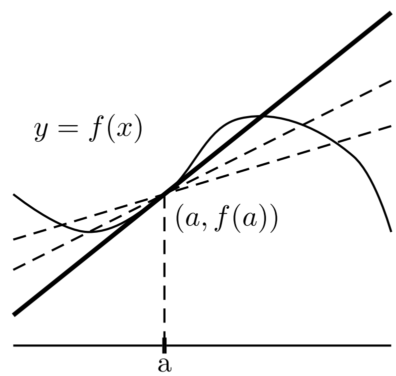


Figure 4.5: Tangent Line

Definition 4.1.2. *Suppose the function f is defined for all x near $x = a$. The line* ***tangent*** *to the graph $y = f(x)$ at $x = a$ is the limit of the secant lines through $(x_1, y_1) = (a, f(a))$ and nearby points (x_2, y_2) on the graph as $\triangle x = x_2 - x_1 \to 0$; the tangent line is the line through $(a, f(a))$ with slope $m = f'(a)$.*

From this description, we can deduce that the equation of the tangent line at $(a, f(a))$ is

$$y - f(a) = f'(a)\,(x - a)$$

or

$$y = f(a) + f'(a)\,(x - a). \tag{4.4}$$

The function defined by Equation (4.4) (whose graph is the tangent line) plays a very important role in the application of derivatives to the analysis of f. We will clarify later (Lemma 6.1.1 and Proposition 6.1.2) that this is the affine function which best approximates the behavior of f near $x = a$. We will call this function by several names: it is the **tangent map** or **affine approximation** to $f(x)$ at $x = a$; we denote it by

$$T_a f(x) := f(a) + f'(a)\,(x - a).$$

In all of this discussion, we have ignored a crucial point: does every function *have* a derivative, or equivalently, does every curve have a tangent line? After all, these are defined as limits, and we have already seen that whenever we work with limits we should first make sure they exist. The following shows that not every function can be differentiated.

Remark 4.1.3. *In order for (the limit defining) the derivative $f'(a)$ to exist, the function f must be continuous at $x = a$.*

This is simply a consequence of the definition of the derivative: we are interested in the limit of the fraction $\frac{\triangle y}{\triangle x}$ as $\triangle x \to 0$, and if $\triangle y$ does *not* go to zero, this *must* diverge (why?). But saying $\triangle y \to 0$ is the same as saying $f(a + \triangle x) - f(a) \to 0$, and this is precisely the same as

$$f(a) = \lim_{x \to a} f(x)$$

(Exercise 9).
We say f is **differentiable** at $x = a$ if the limit defining $f'(a)$ exists; the remark tells us that *every differentiable function is continuous.* Thus for example the function

$$f(x) := \begin{cases} \frac{|x|}{x} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0 \end{cases}$$

is *not* differentiable at $x = 0$. We have seen that $f(x) = 3x^2 - 2$ *is* differentiable at $x = 1$, and in the next section we shall establish that most of our "everyday" functions are differentiable wherever they are defined. But there is one glaring exception: the absolute value

$$f(x) = |x|.$$

First, look at the graph of the absolute value (Figure 4.6): for $x \geq 0$, we

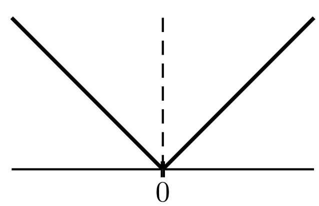


Figure 4.6: $y = |x|$

have $f(x) = x$, a line through the origin with slope $+1$, but for $x \leq 0$, $f(x) = -x$ is a line of slope -1. Certainly near any point $x = a$ *different from* $x = a$ the function looks just like the function $y = x$ (*resp.* $y = -x$) if $a > 0$ (*resp.* $a < 0$), so it is differentiable with derivative $y' = \pm 1$. But what is the tangent line at the origin? *All* secants through $(x_1, y_1) = (0, 0)$

and (x_2, y_2) with $x_2 > 0$ (and hence $y_2 = x_2$) have slope $+1$ while *all* secants with $x_2 < 0$ (so $y_2 = -x_2$) have slope -1. Thus, in the definition of $\frac{dy}{dx}$ as the limit of $\frac{\triangle y}{\triangle x}$, we have distinct one-sided limits

$$\lim_{\triangle x \to 0^+} \frac{\triangle y}{\triangle x} = \lim_{\triangle x \to 0^+} +1 = +1$$
$$\lim_{\triangle x \to 0^-} \frac{\triangle y}{\triangle x} = \lim_{\triangle x \to 0^-} -1 = -1.$$

Thus, the (two-sided) *limit* does *not* exist: *the absolute value function $y = |x|$ is continuous, but NOT differentiable at $x = 0$.*
We shall deal with other examples of functions defined in pieces in the next section.

Exercises for § 4.1

Answers to Exercises 1-2ace, 3a, 4aceg, 5acf, 6acd, 10ade, 11c, 12a are given in Appendix B.

Practice problems:

1. Find the slope of each line below:

   (a) $y = 3x + 5$
   (b) $x = 3y + 5$
   (c) $x + y = 0$
   (d) $3x + 2y = 5$
   (e) $3x = 2y + 5$
   (f) $2y = 3x + 5$

2. Below are several lines described in geometric terms. Write an equation for each line.

   (a) A line through the point $(1, 2)$ with slope 3
   (b) A line through the point $(2, -1)$ with slope -3
   (c) A line through the point $(3, -2)$ with slope 0
   (d) A line through the points $(1, 2)$ and $(2, 4)$
   (e) A line through the points $(1, 2)$ and $(2, 1)$
   (f) A line through the points $(1, 2)$ and $(-1, 3)$

3. (a) Use Definition 4.1.1 and Equation (4.4) to find the equation of the line tangent to the curve $y = x^3$ at the point $(-1, -1)$.

(b) Show that this line also crosses the graph of $y = x^3$ at the point $(2, 8)$.

4. Use Definition 4.1.1 to find the derivative of the given function at the given point.

(a) $f(x) = 3x^2$, $x = 1$

(b) $f(x) = 3x^2$, $x = 2$

(c) $f(x) = 5x^2 - 2x + 1$, $x = 2$

(d) $f(x) = \frac{2}{x}$, $x = 3$

(e) $f(x) = \frac{x^2-2x}{x+1}$, $x = 0$

(f) $f(x) = \sqrt{x}$, $x = 4$

(g)
$$f(x) = \begin{cases} x^2 + 1 & \text{for } x < 1, \\ 2x & \text{for } x \geq 1 \end{cases}, \quad x = 1.$$

5. Consider the function
$$f(x) = 5x^2 - 3x.$$

(a) Find the average rate of change of f as x goes from $x = 1$ to $x = 2$.

(b) Find the average rate of change of f as x goes from $x = 1$ to $x = 1.5$.

(c) Find the average rate of change of f as x goes from $x = 1$ to $x = 1.1$.

(d) Find the average rate of change of f as x goes from $x = 1$ to $x = 0.9$.

(e) Use Definition 4.1.1 to find the instantaneous rate of change of $f(x)$ at $x = 1$.

(f) Find the affine approximation $T_1 f(x)$ to $f(x)$ at $x = 1$.

6. Consider the curve
$$y = x^2 - 3x + 1.$$

(a) Find the equations of the secant lines through the point $(1, -1)$ and each of the points corresponding to $x = 1.1, 1.5, 2$.

(b) Find the equation of the secant line through the point $(1, -1)$ and the point corresponding to $x = 1 + \triangle x$.

(c) Find the equation of the line tangent to the curve $y = x^2 - 3x + 1$ at the point $(1, -1)$.

(d) A line is **normal** to a curve at a point if it is perpendicular to the tangent line at that point. Find the equation of the line normal to the curve $y = x^2 - 3x + 1$ at the point $(1, -1)$. (*Hint:* Use Exercise 11.)

Theory problems:

7. Show that two lines are parallel precisely if they have the same slope. (*Hint:* Show that the two equations $y = mx + b$ and $y = \tilde{m}x + \tilde{b}$ always have a common solution if $m \neq \tilde{m}$, and never have a common solution if $m = \tilde{m}$ and $b \neq \tilde{b}$.)

8. Given three distinct points (x_i, y_i), $i = 1, 2, 3$,

   (a) Show that if these points all lie on the curve with equation

   $$Ax + By = C$$

   then

   $$\frac{y_3 - y_1}{x_3 - x_1} = \frac{y_3 - y_2}{x_3 - x_2} = \frac{y_2 - y_1}{x_2 - x_1}.$$

   (That is, the three secant lines determined by these points all have the same slope.) Show also that in case one of these fractions has zero denominator, they all do.

   (b) Show that if any two of the three fractions above are equal, then there is a curve with equation of the form $Ax + By = C$ on which they all lie.

9. Explain why the statements "$f(a + \triangle x) - f(a) \to 0$" and "$f(a) = \lim_{x \to a} f(x)$" mean precisely the same thing.

10. Give an example of each phenomenon, or explain why none exists.

   (a) A function defined on $[0, 1]$ which is not constant, but whose average rate of change over this interval is zero.

   (b) A function such that the secant line joining any pair of distinct points on its graph has slope 1.

   (c) A function defined on $[0, 1]$ and two sequences of points x_k, x'_k converging to $x = \frac{1}{2}$ such that the two sequences of difference quotients converge to different limits:

   $$\lim \frac{f(x_k) - f\left(\frac{1}{2}\right)}{x_k - \frac{1}{2}} = L \neq L' = \lim \frac{f(x'_k) - f\left(\frac{1}{2}\right)}{x'_k - \frac{1}{2}}.$$

(d) A function defined on $[0,1]$ which has a point of discontinuity at $x = \frac{1}{2}$ with $f'\left(\frac{1}{2}\right) = 0$.

(e) A function defined and continuous on $[0,1]$ for which $f'\left(\frac{1}{2}\right)$ does not exist.

Challenge problems:

11. (a) Show that for any angle θ,

$$\tan(\theta \pm \frac{\pi}{2}) = -\frac{1}{\tan\theta}.$$

(*Hint:* Use angle-summation formulas for the sine and cosine.)

(b) Use this to show that two lines with nonzero slopes m_1 and m_2 are *perpendicular* to each other precisely if

$$m_1 m_2 = -1.$$

(c) Find an equation for the line through $(-1, 2)$ perpendicular to the line $3x + y = 1$.

12. In this problem, we investigate a different possible idea for defining the derivative.

(a) Show that, if $f(x)$ is differentiable at $x = a$, then

$$\lim_{h \to 0} \frac{f(a+h) - f(a-h)}{2h} = f'(a)\,.$$

Note that this is the limit of slopes of secant lines through points symmetrically located on either side of $x = a$.

(b) Give an example of a function which is *not* differentiable at $x = a$ for which this limit exists.

13. Show that if $f(x)$ has a jump discontinuity at $x = a$, then at least one of the one-sided limits involved in defining the derivative must diverge to $\pm\infty$.

4.2 Formal Differentiation I: The Algebra of Derivatives

In this section, we establish some fundamental differentiation formulas. If a function f is defined and differentiable at all points on an interval I, then we can consider the function f' which assigns to each point in I the derivative of f at that point.

General Algebraic Formulas

The easiest differentiation formula concerns constant functions.

Remark 4.2.1. *A constant function has derivative zero.*

To see this, it suffices to note that when y is constant, $\triangle y = 0$ for any pair of points. $\Diamond$

Slightly more involved are monomials.

Proposition 4.2.2. *If $f(x) = cx^n$ (n a positive integer), then*

$$f'(x) = ncx^{n-1}.$$

Proof. Pick two distinct values of x, say $x = x_1$ and $x = x_2$. Then the increment in $y = f(x)$ is

$$\triangle y := f(x_2) - f(x_1) = cx_1^n - cx_2^n = c(x_1^n - x_2^n)$$

so that

$$\frac{\triangle y}{\triangle x} = \frac{c(x_2^n - x_1^n)}{x_2 - x_1};$$

now we invoke the factorization

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + ... + ab^{n-2} + b^{n-1}$$

to write

$$\frac{\triangle y}{\triangle x} = c(x_2^{n-1} + x_2^{n-2}x_1 + x_2^{n-3}x_1^2 + ... + x_2x_1^{n-2} + x_1^{n-1}).$$

Note that this consists of n terms, in each of which the powers of x_1 and x_2 add up to $n-1$. Holding x_1 fixed, and letting $x_2 \to x_1$, we get n copies of x_1^{n-1}, times c.
This shows that the derivative of f at $x = x_1$ is ncx_1^{n-1}, as required. $\square$

A way to remember this formula is: *bring the power down to the front, and reduce the exponent by one.*
We form polynomials by adding together monomials, so we can find their derivatives by combining Proposition 4.2.2 with knowledge of the behavior of sums under differentiation. For stating some of our formulas, it will be easier to use the "$\frac{d}{dx}$" notation in place of primes to indicate differentiation:

$$\frac{d}{dx}[f] := \frac{df}{dx}.$$

Proposition 4.2.3. *If f and g are both differentiable functions, then*

1. *For any constant $c \in \mathbb{R}$,*

$$\frac{d}{dx}\left[cf(x)\right] = c\frac{d}{dx}\left[f(x)\right];$$

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

Proof. This is an easy consequence of the fact that $\triangle[cf(x)] = c\triangle[f(x)]$ (*resp.* $\triangle[f(x) + g(x)] = \triangle[f(x)] + \triangle[g(x)]$) and Theorem 2.4.1. We leave the details to you (Exercise 9). □

The two formulas in Proposition 4.2.3 can be combined in one statement:

> *For any two differentiable functions f and g and any two constants $\alpha, \beta \in \mathbb{R}$, the function $\alpha f + \beta g$ is also differentiable, and*

$$\frac{d}{dx}\left[\alpha f(x) + \beta g(x)\right] = \alpha\frac{d}{dx}\left[f(x)\right] + \beta\frac{d}{dx}\left[g(x)\right]. \tag{4.5}$$

A **linear combination** of two functions is any function which can be obtained by multiplying each of the given functions by a constant and adding; Equation (4.5) is sometimes stated as *differentiation respects linear combinations*. One easy application of this is that any polynomial can be differentiated term-by-term: for example, if

$$f(x) = 3x^2 - 5x + 4$$

then we have

$$f'(x) = 3\frac{d}{dx}\left[x^2\right] - 5\frac{d}{dx}\left[x\right] + \frac{d}{dx}\left[4\right] = 3(2x) - 5(1) + 0 = 6x - 5.$$

What about products? Since the limit of a product is the product of the limits, it might be tempting to think that the same holds for derivatives. However, things are not quite so simple here (if it were, then for example the derivative of cx would be the product of the derivative of c–which is zero–with the derivative of x–which is one; but their product is zero, and that *can't* be right). The correct formula is

Proposition 4.2.4 (Product Rule)**.** [4] *If $u = f(x)$ and $v = g(x)$ are both differentiable, then so is their product, and*

$$\frac{d[uv]}{dx} = \left(\frac{du}{dx}\right) \cdot (v) + (u) \cdot \left(\frac{dv}{dx}\right); \tag{4.6}$$

equivalently,

$$\left.\frac{d[fg]}{dx}\right|_{x=a} = f'(a)\, g(a) + f(a)\, g'(a)\,.$$

Proof. The heart of the argument here is to express the increment $\triangle(uv) := \triangle(f(x)\, g(x))$ in terms of $\triangle u := \triangle(f(x))$ and $\triangle v := \triangle(g(x))$, via a trick—given the two values $x = a$ and $x = x_2$, subtract and add $f(a)\, g(x_2)$ in the middle of the expression for $\triangle(uv)$:

$$\begin{aligned}
\triangle(uv) &= f(x_2)\, g(x_2) - f(a)\, g(a) \\
&= f(x_2)\, g(x_2) - f(a)\, g(x_2) + f(a)\, g(x_2) - f(a)\, g(a) \\
&= (f(x_2) - f(a))\, g(x_2) + f(a)\, (g(x_2) - g(a)) \\
&= (\triangle u)\, g(x_2) + f(a)\, (\triangle v)
\end{aligned}$$

Dividing the last expression by $\triangle x$, we obtain the formula

$$\frac{\triangle[uv]}{\triangle x} = \left(\frac{\triangle u}{\triangle x}\right)(g(x_2)) + (f(a))\left(\frac{\triangle v}{\triangle x}\right)$$

and taking limits as $x_2 \to a$, we have

$$\frac{d[uv]}{dx} = \left(\frac{du}{dx}\right)\left(\lim_{x_2\to a} g(x_2)\right) + \left(\lim_{x_2\to a} f(x_2)\right)\left(\frac{dv}{dx}\right).$$

Finally, since f and g are differentiable at a, they are also continuous there (by Remark 4.1.3), so we see that the two limits equal the values of f (*resp.* g) at $x = a$, and gives us the appropriate formula. □

Exercise 20 discusses the more geometric version of this proof given by Newton.

This formula can be summarized as: *the derivative of the product is a sum of products in which each term has one factor differentiated.* It is useful for differentiating polynomials in factored form. For example, to differentiate

[4] The product rule in the differential notation of Equation (4.6) was first given by Leibniz; analogous formulas occur in other areas of mathematics, and in such contexts the product rule is sometimes called the *Leibniz formula*.

$f(x) = (x^2 + 3x - 2)(x^3 + x^2 + 5)$, we can take $u = x^2 + 3x - 2$ and $v = x^3 + x^2 + 5$; then

$$f'(x) =$$
$$(x^2 + 3x - 2)\left(\frac{d}{dx}(x^3 + x^2 + 5)\right) + \left(\frac{d}{dx}(x^2 + 3x - 2)\right)(x^3 + x^2 + 5)$$

$$= (x^2 + 3x - 2)(3x^2 + 2x) + (2x + 3)(x^3 + x^2 + 5).$$

You can check that first multiplying out the two factors and then differentiating term-by-term yields the same thing as multiplying out the expression above.
What about quotients? We will give the formula here without proof: there are two different ways to prove it: one is a variant of the proof above (Exercise 10) and the other uses the Chain Rule (see § 4.5, Exercise 5).

Proposition 4.2.5 (Quotient Rule)**.** *If $u = f(x)$ and $v = g(x)$ are both differentiable, then so is their quotient (provided the denominator is not zero), and*

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}. \tag{4.7}$$

or equivalently

$$\frac{d}{dx}\Big|_{x=a}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(a)\,g(a) - f(a)\,g'(a)}{(g(a))^2}.$$

This looks just like the product formula, with two changes: the term in which the *denominator* is differentiated has a minus in front, and the whole business is divided by the denominator squared. This allows us to differentiate all rational functions. For example,

$$\frac{d}{dx}\left[\frac{x}{x^2+1}\right] = \frac{(1)(x^2+1) - (x)(2x)}{(x^2+1)^2}$$
$$= \frac{1 - x^2}{(x^2+1)^2}.$$

A particular case of the quotient rule is

Remark 4.2.6. *For any positive integer n,*

$$\frac{d}{dx}\left[x^{-n}\right] = -nx^{-(n+1)}. \tag{4.8}$$

(You should be able to derive this from the Quotient Rule yourself.) Notice that this is consistent with the pattern for positive powers of x: *bring down the power ($-n$) and lower the exponent by one (to $-n-1=-(n+1)$).* We can also take care of roots, and note the pattern continuing.

Proposition 4.2.7. *For any positive integer n,*

$$\frac{d}{dx}\left[\sqrt[n]{x}\right]=\frac{1}{n\,(\sqrt[n]{x})^{n-1}}$$

or

$$\frac{d}{dx}[x^{\frac{1}{n}}]=\frac{1}{n}x^{\frac{1}{n}-1}.$$

Proof. If $y=\sqrt[n]{x}$, then we can use the factorization

$$a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+a^{n-3}b^2+...+b^{n-1})$$

with $a=\sqrt[n]{x+h}$ and $b=\sqrt[n]{x}$ to write

$$\begin{aligned}\frac{\triangle y}{\triangle x}&=\frac{\sqrt[n]{x+h}-\sqrt[n]{x}}{h}\\&=\frac{(\sqrt[n]{x+h}-\sqrt[n]{x})\left((\sqrt[n]{x})^{n-1}+(\sqrt[n]{x})^{n-2}\sqrt[n]{x}+...+(\sqrt[n]{x})^{n-1}\right)}{h\,((\sqrt[n]{x})^{n-1}+(\sqrt[n]{x})^{n-2}\sqrt[n]{x}+...+(\sqrt[n]{x})^{n-1})}\\&=\frac{\left(\sqrt[n]{x+h}\right)^n-(\sqrt[n]{x})^n}{h\,((\sqrt[n]{x})^{n-1}+(\sqrt[n]{x})^{n-2}\sqrt[n]{x}+...+(\sqrt[n]{x})^{n-1})}\\&=\frac{h}{h\,((\sqrt[n]{x})^{n-1}+(\sqrt[n]{x})^{n-2}\sqrt[n]{x}+...+(\sqrt[n]{x})^{n-1})}\\&=\frac{1}{((\sqrt[n]{x})^{n-1}+(\sqrt[n]{x})^{n-2}\sqrt[n]{x}+...+(\sqrt[n]{x})^{n-1})}.\end{aligned}$$

Now note that the denominator consists of n terms, each a product of powers of $\sqrt[n]{x+h}$ and of $\sqrt[n]{x}$ with the powers adding up to $n-1$; as $h\to 0$, each of these products goes to $(\sqrt[n]{x})^{n-1}$, and so we get the formula required. □

Trigonometric Functions

What about the trig functions? For this, we will need the two limits

$$\lim_{\theta\to 0}\frac{\sin\theta}{\theta}=1 \tag{4.9}$$

$$\lim_{\theta\to 0}\frac{1-\cos\theta}{\theta}=0 \tag{4.10}$$

established in § 3.4, as well as the angle summation formulas for the sine and cosine (see Exercise 16 for a proof of these): for any angles α and β,

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta \tag{4.11}$$

$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta \tag{4.12}$$

Proposition 4.2.8.

$$\frac{d}{dx}[\sin x] = \cos x \tag{4.13}$$

$$\frac{d}{dx}[\cos x] = -\sin x \tag{4.14}$$

Proof. We will prove the first formula, Equation (4.13), and leave the second to you (Exercise 13). We will use "h" in place of $\triangle x$ in writing down our computations:

$$\begin{aligned}\frac{\triangle \sin x}{\triangle x} &= \frac{\sin(x+h) - \sin x}{h} \\ &= \frac{\sin x\cos h + \cos x\sin h - \sin x}{h} \\ &= (\sin x)\frac{\cos h - 1}{h} + (\cos x)\frac{\sin h}{h}\end{aligned}$$

so (remember that x is fixed in the limit below; only h is moving)

$$\begin{aligned}\frac{d}{dx}[\sin x] &:= \lim_{h\to 0}\frac{\triangle \sin x}{\triangle x} \\ &= (\sin x)\lim_{h\to 0}\frac{\cos h - 1}{h} + (\cos x)\lim_{h\to 0}\frac{\sin h}{h} \\ &= (\sin x)(0) + (\cos x)(1) \\ &= \cos x.\end{aligned}$$

□

Notice that this formula depends *essentially* on the use of radians to measure angles: if we had used degrees instead, then the limit in Equation (4.9) would be $\frac{\pi}{180}$, and then every differentiation formula involving trig functions would require a similar factor in front! This is why the *natural* unit for angles in calculus is the radian.

Combining Proposition 4.2.8 with the Quotient Rule, we can also derive the differentiation formulas for the other trig functions; we leave the details to you (Exercises 3-4).

Corollary 4.2.9.

$$\frac{d}{dx}[\tan x] = \sec^2 x \qquad \frac{d}{dx}[\sec x] = \tan x \sec x \tag{4.15}$$

$$\frac{d}{dx}[\cot x] = -\csc^2 x \qquad \frac{d}{dx}[\csc x] = -\cot x \csc x. \tag{4.16}$$

The basic differentiation formulas, summarized in Table 4.1, should be memorized; they form the underpinning of all calculations with derivatives. Table 4.1 summarizes the basic differentiation formulas, which should be memorized.

Table 4.1: Differentiation Formulas

Function	Derivative
$af(x) \pm bg(x)$	$af'(x) \pm bg'(x)$
$f(x)\, g(x)$	$f'(x)\, g(x) + f(x)\, g'(x)$
$\frac{f(x)}{g(x)}$	$\frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$
constant	0
x^p	px^{p-1}
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\sec^2 x$
$\sec x$	$\tan x \sec x$

Functions Defined in Pieces

Finally we consider differentiation for functions defined in pieces. We should clarify that not every function defined by cases is of this sort: for example the function

$$f(x) = \begin{cases} 0 & \text{for } x \text{ rational,} \\ 1 & \text{for } x \text{ irrational} \end{cases}$$

is *not* defined "in pieces", because the two cases are intertwined; we reserve the term **function defined in pieces** for a function defined by different formulas in different intervals; a common endpoint of two such intervals is an **interface point** for the function. At an interface point, nearby inputs to the *right* of the point are governed by *one* formula, and those on the *left* are defined by *another*. We saw in § 3.4 that to find the limit of a function defined in pieces at an interface point, it is enough to compare the two one-sided limits there, which in general means that (when the formula is defined by algebraic functions) we need only compare the values of the competing formulas at the point: if they agree, the limit exists and equals their common value, and if they disagree, then the two one-sided limits are different and hence the limit does not exist. This reasoning applies as well to the limits defining derivatives. For example, consider the function

$$f(x) = \begin{cases} x^2 & \text{for } x \leq 0, \\ x^3 & \text{for } x > 0 \end{cases}$$

(see Figure 4.7).

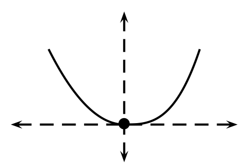


Figure 4.7: Derivative at an interface point

At any point $x = a$ with $a \neq 0$, the function is easy to analyze: if $a < 0$, then all points x sufficiently close to a also satisfy $x < 0$, and so we need never consider the formula x^3–as far as we are concerned, the function looks like $f(x) = x^2$ near $x = a$, it is clearly continuous, and we can just formally differentiate x^2 to get

$$f'(a) = 2a \quad \text{for } a < 0;$$

similarly, we get

$$f'(a) = 3a^2 \quad \text{for } a > 0.$$

But what about *at* $a = 0$? The reasoning is this: we need to find

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \to 0} \frac{f(h) - (f(0) = 0)}{h}$$

From the definition of $f(x)$, we see that

$$\frac{f(h) - f(0)}{h} = \begin{cases} \frac{h^2}{h} = h & \text{for } h < 0, \\ \frac{h^3}{h} = h^2 & \text{for } h > 0 \end{cases}$$

so that

$$\lim_{h\to 0^-} \frac{f(h)-f(0)}{h} = \lim_{h\to 0^-} h = 0$$
$$\lim_{h\to 0^+} \frac{f(h)-f(0)}{h} = \lim_{h\to 0^+} h^2 = 0$$

and since the two one-sided limits agree, the limit exists and equals their common value:

$$f'(0) = 0.$$

In effect, what we did was to evaluate the *derivative* of the formulas in force on each side of $x = 0$ and compare their values at $x = 0$; since they agreed, we concluded that the derivative exists and equals this common value. However, there is one important caveat. Suppose instead of f as above we considered g defined by

$$g(x) = \begin{cases} x^2 + 1 & \text{for } x \leq 0, \\ x^3 & \text{for } x > 0 \end{cases}$$

(see Figure 4.8).

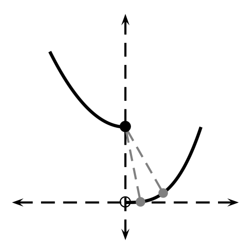


Figure 4.8: Jump discontinuity at an interface point (secants → vertical)

The derivative formulas both give 0 at $x = 0$, but *the function is not continuous* at $x = 0$, so it can't possibly be differentiable there! The problem is a subtle one: the derivative of $x^2 + 1$ at $x = 0$ agrees perfectly with the limit from the left of $\frac{\triangle y}{\triangle x}$, but in the limit defining the derivative of x^3 at $x = 0$, we are assuming $g(0) = 0$, when in fact $g(0) = 1$, so *that* derivative does *not* give the correct limit from the right: in fact, that limit does not exist (*check it out!*). So let us state this with some care:

Remark 4.2.10. *Suppose f is defined in pieces, and $x = a_0$ is an interface point at which f is continuous: that is, at least for x near $x = a_0$, suppose we have*

$$f(x) = \begin{cases} F_1(x) & \textit{for } x < a_0, \\ y_0 & \textit{at } x = a_0 \\ F_2(x) & \textit{for } x > a_0 \end{cases}$$

where F_1 and F_2 are algebraic functions with $F_1(a_0) = F_2(a_0) = y_0$. Then

$$f'(x) = \begin{cases} F_1'(x) & \textit{for } x < a_0, \\ F_2'(x) & \textit{for } x > a_0 \end{cases}$$

and at $x = x_0$ the derivative is specified by:

- *if the derivatives of the two neighboring formulas agree at a_0, say*

  $$F_1'(a_0) = F_2'(a_0) = m$$

  then this common value equals the derivative

  $$f'(a_0) = m;$$

- *if*

  $$F_1'(a_0) \neq F_2'(a_0)$$

  then f is not differentiable at $x = a_0$ (see Figure 4.9).

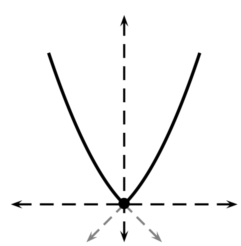


Figure 4.9: $F_1'(a_0) \neq F_2'(a_0)$

Pay attention to the requirement that $F_1(a_0) = F_2(a_0)$ in the hypotheses: if the two original formulas fail to agree at $x = a_0$ then the function is not continuous, and hence also not differentiable, at the interface, as in Figure 4.8.

We note also that Remark 4.2.10 does not really require that the two formulas $F_i(x)$, $i = 1, 2$ be *algebraic*: all we need is that the functions given by these formulas are both differentiable at and near $x = a_0$.

Higher Derivatives

Note that if a function f is differentiable on an interval I, then its derivative f' is a new function, which we can try to differentiate again: the derivative of the derivative is called the **second derivative** of f and is denoted in one of several ways, for example:

$$\begin{aligned}\frac{d}{dx}\left[\frac{dy}{dx}\right] &= \frac{d^2y}{dx^2} = \\ \left(f'(x)\right)' &= f''(x) = \\ (y')' &= y''.\end{aligned}$$

Thus, if

$$y = x^3$$

then

$$\begin{aligned}y' &= \frac{dy}{dx} = 3x^2, \\ y'' &= \frac{d^2y}{dx^2} \\ &= \frac{d}{dx}\left[\frac{dy}{dx}\right] \\ &= \frac{d}{dx}\left[3x^2\right] &= 6x.\end{aligned}$$

This process can be repeated: the third derivative is the derivative of the second, and so on: as long as we have a differentable function at the n^{th} stage, we can define the $\boldsymbol{(n+1)^{st}}$ **derivative** as

$$\begin{aligned}\frac{d^{n+1}y}{dx^{n+1}} &= \frac{d}{dx}\left[\frac{d^ny}{dx^n}\right] \\ = f^{(n+1)}(x) &= \left(f^{(n)}(x)\right)'.\end{aligned}$$

The second derivative plays an important role in § 4.8, and the higher derivatives are central to Chapter 6.

Exercises for § 4.2

Answers to Exercises 1-2acegikmoqsuw, 3a, 5abe, 6ace, 7ac, 8ac are given in Appendix B.

Practice problems:

1. Use the formal rules of this section to calculate the derivative of each function f at the given point $x = a$:

   (a) $f(x) = 3x^2 - 2x + 5$, $a = 1$

   (b) $f(x) = 3x^{-2} + 2x^{-1} + 5x$, $a = 1$

   (c) $f(x) = x^3 - \sqrt{x} + 2$, $a = 4$

   (d) $f(x) = \sqrt[3]{x} - 2\sqrt{x}$, $a = 1$

   (e) $f(x) = 2x^{1/3} - x^{1/2} + 5x^{-1}$, $a = 64$

   (f) $f(x) = \sqrt{x} - 2\sqrt[4]{x}$, $a = 16$

   (g) $f(x) = x^2(x^3 + 1)$, $a = 1$

   (h) $f(x) = (x^2 - 1)(x + 1)$, $a = 0$

   (i) $f(x) = (x^3 - 2x + 5)(3x^2 + 1)$, $a = 1$

   (j) $f(x) = x \sin x$, $a = \frac{\pi}{2}$

   (k) $f(x) = \dfrac{\sin x}{x}$, $a = \frac{\pi}{3}$

   (l) $f(x) = \tan x + \sec x$, $a = \frac{2\pi}{3}$

   (m) $f(x) = 3 \sec x(\sin x + \cos x)$, $a = 0$

   (n) $f(x) = \dfrac{x}{x^2 + 1}$, $a = 1$

   (o) $f(x) = \dfrac{x - 1}{x + 2}$, $a = 1$

   (p) $f(x) = \dfrac{x - 1}{x^2 + 2x - 3}$, $a = -1$

   (q) $f(x) = \dfrac{x^2 - 1}{x + 1}$, $a = 1$

   (r) $f(x) = \dfrac{x^{1/2} - 1}{x - 1}$, $a = 4$

   (s) $f(x) = \dfrac{\sin x}{1 + \cos x}$, $a = 0$

   (t) $f(x) = \dfrac{\sin x - \cos x}{\sin x + \cos x}$, $a = 0$

   (u) $f(x) = x^{-1} \cot x$, $a = \frac{\pi}{4}$

   (v) $f(x) = (x + 1) \csc x$, $a = \frac{4\pi}{3}$

   (w) $f(x) = \sin 2x$, $a = \frac{\pi}{3}$ (*Hint:* Use trig identities.)

   (x) $f(x) = \cos 2x$, $a = \frac{3\pi}{2}$ (*Hint:* Use trig identities.)

2. Find the equation of the line tangent to the graph of the given function at the given point, for each item in the preceding problem.

3. (a) Use Proposition 4.2.8 and Proposition 4.2.5 to calculate the derivative of $f(x) = \tan x$ and $g(x) = \sec x$ at $x = \frac{\pi}{3}$.

(b) Use Proposition 4.2.5, together with Proposition 4.2.8, to prove Equation (4.15).

4. Use Proposition 4.2.5, together with Proposition 4.2.8, to prove Equation (4.16).

5. For each function f defined in pieces below, find the function $f'(x)$. In particular, determine whether the derivative exists at the interface point, and if so, find its value.

(a) $f(x) = \begin{cases} x^2+1 & x \le 1, \\ 2x & x > 1. \end{cases}$ (b) $f(x) = \begin{cases} x^2 & x \le 1, \\ 2x & x > 1. \end{cases}$

(c) $f(x) = \begin{cases} x^4 - 2x^2 + 1 & x \le 1, \\ 2x^3 - 4x^2 + 2x & x > 1. \end{cases}$

(d) $f(x) = \begin{cases} -(x-1)^2 & x \le 1, \\ (x-1)^3 & x > 1. \end{cases}$

(e) $f(x) = \begin{cases} \sin x & x < 0, \\ x & x \ge 0. \end{cases}$ (f) $f(x) = \begin{cases} \sin x & x \le 0, \\ \sin x \cos x & x > 0. \end{cases}$

6. For each function f defined in pieces below, find all values of the constant α (or the constants α and β) for which the function is differentiable at the interface point, or explain why no such values exist.

(a) $f(x) = \begin{cases} x^{3/2} & x < 0 \\ \alpha x^2 & x \ge 0 \end{cases}$ (b) $f(x) = \begin{cases} \frac{x^2+\alpha x}{\alpha+1} & x < 1 \\ x^3 & x \ge 1 \end{cases}$

(c) $f(x) = \begin{cases} \alpha \sin x & x \le \frac{\pi}{4} \\ \cos(-x) & x > \frac{\pi}{4} \end{cases}$ (d) $f(x) = \begin{cases} \sin x & x \le \frac{\pi}{4} \\ \alpha \cos x & x > \frac{\pi}{4} \end{cases}$

(e) $f(x) = \begin{cases} \sin x & x < -\frac{\pi}{4} \\ \alpha \cos x & x > -\frac{\pi}{4} \end{cases}$ (f) $f(x) = \begin{cases} \alpha - \beta x^2 & x < 1 \\ x^2 & x \ge 1. \end{cases}$

7. For each function f below, find the first three derivatives: f', f'' and f''':

(a) $f(x) = x^2 + 2x + 3$ (b) $f(x) = \dfrac{1}{2x}$

(c) $f(x) = x^{1/2}$ (d) $f(x) = \tan x$

8. Let $y = \sin x$.

   (a) Find the first four derivatives $\frac{d^n y}{dx^n}$, $n = 1, 2, 3, 4$.

   (b) Find the eighth derivative $\frac{d^8 y}{dx^8}$.

   (c) Find a formula for the even and odd derivatives $\frac{d^{2k} y}{dx^{2k}}$ and $\frac{d^{2k+1} y}{dx^{2k+1}}$.

Theory problems:

9. Suppose f and g are differentiable functions.

   (a) Show that for any constant c, $\triangle [cf(x)] = c\triangle [f(x)]$

   (b) Use this to show that

$$\frac{d}{dx}[cf(x)] = c\frac{d}{dx}[f(x)].$$

   (c) Show that $\triangle [f(x) + g(x)] = \triangle [f(x)] + \triangle [g(x)]$.

   (d) Use this to show that

$$\frac{d}{dx}[f(x) + g(x)] = \frac{d}{dx}[f(x)] + \frac{d}{dx}[g(x)].$$

   (e) Combine the above to establish Equation (4.5).

10. Prove Proposition 4.2.5 as follows:

   (a) Let $u = f(x)$ and $v = g(x)$. Show that

$$\triangle \left(\frac{u}{v}\right) = \frac{f(x_2)\, g(x_1) - f(x_1)\, g(x_2)}{g(x_1)\, g(x_2)}$$

   by putting everything over a common denominator.

   (b) Use a trick similar to that in the proof of Proposition 4.2.4 to show that this equals

$$\frac{(\triangle f) g(x_1) + f(x_1)\, (-\triangle g)}{g(x_1)\, g(x_2)}.$$

   (c) Show that as $x_2 \to x_1$ this converges to

$$\frac{f'(x_1)\, g(x_1) - f(x_1)\, g'(x_1)}{g(x_1)^2}.$$

11. Use the Quotient Rule to prove Remark 4.2.6.

12. Use the Product Rule and mathematical induction to prove that if $f(x)$ is differentiable at $x = a$, then for any natural number $n > 1$, $\frac{d}{dx}\big|_{x=x=a} [(f(x))^n] = n\,(f(a))^{n-1}\, f'(a)$.

13. (a) Calculate the derivative of $f(x) = \cos x$ at $a = \pi$ directly from the definition of the derivative.
    (b) Prove Equation (4.14) as follows:
        i. Show that
        $$\frac{\triangle \cos x}{\triangle x} = (\cos x)\frac{\cos h - 1}{h} - (\sin x)\frac{\sin h}{h}$$
        ii. Take the limit as $h \to 0$.

14. Suppose the function f is strictly *increasing* on the set $S \subset \mathbb{R}$.
    (a) Show that the function $-f(x)$ is strictly *decreasing* on S;
    (b) Show that the function g defined on $-S := \{x \mid -x \in S\}$ by $g(x) = f(-x)$ is strictly *decreasing* on $-S$;
    (c) Combine the above to show that the function $-f(-x)$ is strictly *increasing* on $-S$.

15. Show that for every point on the curve $y = \dfrac{1}{x}$, $x > 0$, the tangent line to the curve, together with the coordinate axes from the origin to their intersection with the tangent line, form a triangle of area 2.

Challenge problems:

16. Prove the **angle-summation formulas for sine and cosine** (Equations (4.11) and (4.12), also Equations (3.5) and (3.6) in § 3.1) as follows:
    (a) Show that the chord AB subtended by an angle of θ radians in a circle of radius 1 has length $2\sin\dfrac{\theta}{2}$ (Figure 4.10):
    (b) Show that the angles $\angle OAB$ and $\angle OBA$ in Figure 4.10 are each equal to $\dfrac{\pi}{2} - \dfrac{\theta}{2}$.
    (c) Now consider the setup in Figure 4.11.
    Show that $\angle CAB = \alpha$ and $\angle ACB = \beta$.

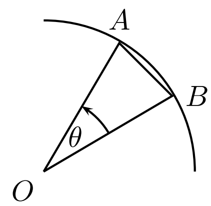


Figure 4.10: Chord subtended by an angle

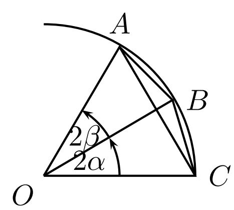


Figure 4.11: $\sin(\alpha + \beta)$

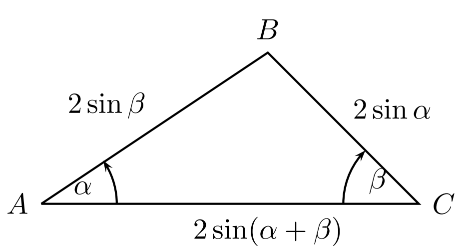


Figure 4.12: The triangle $\triangle ABC$

(d) Show that the triangle $\triangle ABC$ has sides and angles as shown in Figure 4.12.

(e) Use this to prove Equation (3.5) for $\alpha \geq 0$, $\beta \geq 0$ and $\alpha + \beta \leq \frac{\pi}{2}$.

(f) Sketch the relevant diagram if $0 < \alpha \leq \frac{\pi}{2}$ and $0 < \beta \leq \frac{\pi}{2}$ but $\frac{\pi}{2} < \alpha + \beta \leq \pi$, and use the fact that $\sin(\pi - \alpha - \beta) = \sin(\alpha + \beta)$ to show that Figure 4.12 still applies.

(g) Use the identities $\sin(\frac{\pi}{2} - \theta) = \cos\theta$ and $\cos(\frac{\pi}{2} - \theta) = \sin\theta$ to show that Equation (3.5) implies Equation (3.6).

(h) How would you handle angles with other values?

17. Show that if f is defined and bounded near $x = 0$ (but not necessarily differentiable, or even continuous, at $x = 0$) then $g(x) = x^2 f(x)$ is differentiable at $x = 0$, and

$$g'(0) = 0.$$

(*Hint:* Use the Squeeze Theorem on the definition of the derivative.)

18. Here is a "proof" of the Quotient Rule from the Product Rule: Suppose

$$h(x) = \frac{f(x)}{g(x)}.$$

Then

$$h(x) \cdot g(x) = f(x)$$

so differentiating both sides, and using the Product Rule on the left side, we have

$$h'(x) \cdot g(x) + h(x) \cdot g'(x) = f'(x)$$

which we can solve for $h'(x)$:

$$h'(x) = \frac{f'(x) - h(x) \cdot g'(x)}{g(x)}.$$

But then substituting the definition of $h(x)$ gives the desired formula.

Examine this proof. What does it actually show, and what is assumed to be true?

History note:

19. In this problem, we will prove the **Law of Cosines**, which says that if a triangle has sides with lengths a, b, and c, and the angle opposite side c is θ, then

$$c^2 = a^2 + b^2 - 2ab\cos\theta. \tag{4.17}$$

(a) Show that **Pythagoras' Theorem** (Euclid I.47, [32, vol. I, p. 349]), that in a right triangle the square on the hypotenuse is the sum of the squares on the legs, is a special case of the Law of Cosines.

(b) Prove this special case as follows: suppose the lengths of the legs are a and b, and draw a square of side $a + b$. Going clockwise around the edge of the square, mark points a units into each side, and join successive points of this type with line segments (see Figure 4.13).

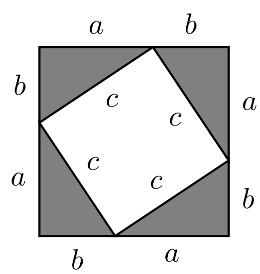


Figure 4.13: Pythagoras' Theorem

Show that the total (dark) area of the four triangles in Figure 4.13 is $2ab$, and that the inside figure (white) is a square (you need to check the angles). Then subtract the dark area from the area of the outside square to show that

$$c^2 = a^2 + b^2. \tag{4.18}$$

Note: Heath has an extensive commentary on this theorem and its proofs [32, vol. I, pp. 350-368]. The figure and proof we give, which is famous and elegant, is not the one in Euclid; Heath gives a specific reference to this proof in the work of the twelfth-century Indian writer Bhāskara (whose proof consisted of a drawing like Figure 4.13 together with the exclamation "Behold!").

(c) Now use Pythagoras' theorem, together with the angle-summation formulas, to prove the general law of cosines as follows. Suppose the triangle $\triangle ABC$ has sides $AC = a$, $BC = b$ and $AB = c$, and denote the angle at C by θ. Drop a perpendicular from C to AB, meeting it at D (Figure 4.14): denote the angles formed by this with the sides by $\angle ACD = \alpha$ and $\angle BCD = \beta$, so

$$\theta = \alpha + \beta.$$

Also, denote the lengths of the two intervals into which D divides AB by $AD = x$, $DB = y$, and the length of the perpendicular $CD = z$; in particular,

$$x + y = c.$$

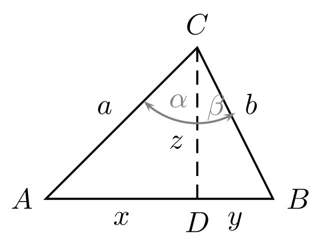


Figure 4.14: Law of Cosines

Show that

$$x = a \sin \alpha$$
$$y = b \sin \beta$$

and

$$z = a \cos \alpha = b \cos \beta.$$

Also, show that

$$a^2 + b^2 = x^2 + y^2 + 2z^2$$
$$c^2 = x^2 + y^2 + 2xy.$$

Conclude from these equations that

$$c^2 = a^2 + b^2 - 2ab \cos(\alpha + \beta).$$

(d) *Caution:* Figure 4.14 assumes that the two base angles of $\triangle ABC$ are both between 0 and $\frac{\pi}{2}$. Sketch the corresponding figure if one of them is between $\frac{\pi}{2}$ and π (why is the other necessarily less than $\frac{\pi}{2}$?), and explain how the argument above has to be modified in this case.

20. **Newton's Proof of the Product Rule:** Newton's *magnum opus*, PHILOSOPHIAE NATURALIS PRINCIPIA MATHEMATICA [41] (generally referred to as the *Principia*–the title translates as "Mathematical Principles of Natural Philosophy"), whose first edition was published in 1687, presents his theory of universal gravitation and is the seminal text of what is now called "Newtonian mechanics". Curiously, although he had formulated the calculus in 1665-6, by the time of the *Principia* twenty years later, he had adopted an approach to mathematics based on the synthetic geometry of the ancient Greeks, as summarized in Euclid.[5] The one place that something like our calculus appears in the *Principia* is Lemma 2 of Book 2 [41, pp. 646-649]. Here in modern language is his proof of the product formula, as given there:

    Given two lengths A and B, with A changing at the rate a and B changing at the rate b (say per unit time), we construct the rectangle with sides A and B and ask how fast its area is changing. If we consider the rectangle at time a half-unit earlier, the rectangle has area

$$(A-\frac{1}{2}a)(B-\frac{1}{2}b) = AB - \frac{1}{2}aB - \frac{1}{2}bA + \frac{ab}{4}$$

    while a half-unit later it has area

$$(A+\frac{1}{2}a)(B+\frac{1}{2}b) = AB + \frac{1}{2}aB + \frac{1}{2}bA + \frac{ab}{4}.$$

    The difference of these two represents the change in the area over a time of one unit, and equals $aB + bA$.

[5] There is a traditional assumption that in fact Newton first obtained the results in *Principia* by calculus, but then translated them into the language of synthetic geometry when writing the book. This was suggested by Newton himself, during his dispute with Leibniz over priority for the discovery of the calculus, but there is not a scrap of historical evidence that this is what he actually did. However, as Niccolò Guicciardini argues [28], Newton had by the time of the *Principia* adopted a "synthetic" treatment of the basic ideas of calculus, and this approach *is* present in Books 1 and 2 of the *Principia*. As examples, see Exercise 11 in § 3.4, as well as Exercise 21 in this section.

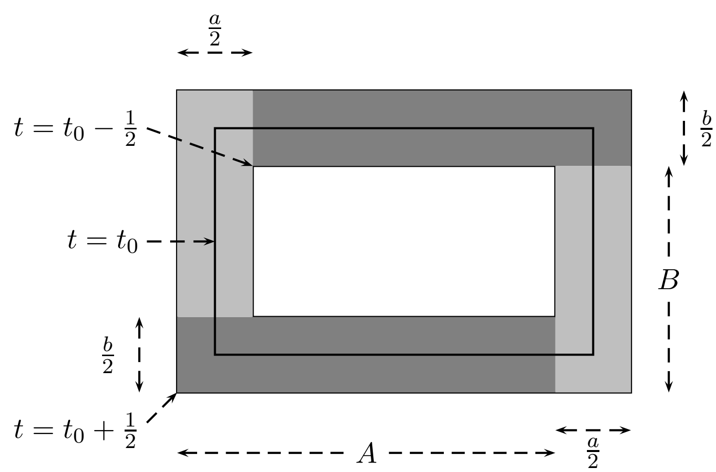


Figure 4.15: Newton's proof of the product rule

The algebraic calculations above are actually there in Newton, but a geometric version is indicated in Figure 4.15

How legitimate is this argument, in modern terms? What does it prove, and what does it simply assume? (In this connection, look at Exercise 12 in § 4.1.)

For comparison, we mention that Leibniz's proof of the product rule is formally like the one we have in the text, except that instead of looking at increments and taking limits, he simply manipulates differentials, throwing away all terms involving more than one differential factor ([51, pp. 272-5], [49, pp. 620-2], [20, p. 255-6]).

21. **Newton differentiates $\sin x$:** As noted in Exercise 20, Newton abandoned the "analytic" version of his calculus in favor of a more "synthetic" geometric approach in the manner of Euclid. This was set forth in a work entitled *Geometria Curvilinea* (The Geometry of Curved Lines) around 1680. He states in this work, without giving a proof, that the derivative of the sine is the cosine[6], but Niccolò Guicciardini [28, pp. 33-34] gives the following hypothetical "synthetic" proof (see Figure 4.16).

    (a) As noted in Exercise 11 in § 3.4, as b tends toward B, the ratio

[6] I am butchering his fluxion-based formulation for the sake of efficiency and clarity.

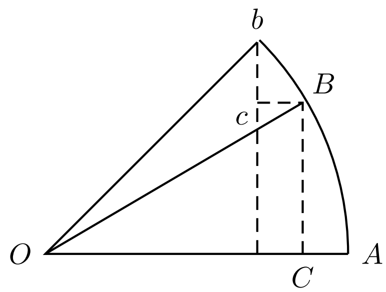


Figure 4.16: Newtonian differentiation of $\sin x$

between the chord $\overline{bB}$ and the arc $\widehat{bB}$ goes to 1:

$$\frac{\overline{bB}}{\widehat{bB}} \to 1.$$

(b) Since the pairs of lines Bc, BC and bc, OC are perpendicular, and the chord bB tends to perpendicularity with OB as b tends toward B, the triangles $\triangle bBc$ and $\triangle OBC$ tend to similarity, and in particular the ratio of the angles $\angle bBc$ and $\angle OBC$ tends to 1.

(c) It follows that the "ultimate ratios" bc/bB and OC/OB are equal. But the first is the ratio between the "fluxion" of the vertical at b (ie, of the sine of the angle $\angle bOA$ times the radius $\overline{OA}$) and the "fluxion" of the arc $\widehat{bA}$, while the latter is precisely $\overline{OA}$ times the cosine of $\angle BOA$.

4.3 Formal Differentiation II: Exponentials and Logarithms

In this section, we will derive a formula for the derivatives of the exponential functions defined in § 3.6; while the final formula is a simple one, it will take considerable work, involving many of the theoretical results we have obtained so far, to derive it. Along the way, we will see

that there is a "natural" exponential function (as well as a related "natural" logarithm) which is best suited for our purposes.
We begin by reducing the differentiability of an exponential function on the whole real line to its differentiability at $x = 0$.

Lemma 4.3.1. *For any base $b > 0$ and any $a \in (-\infty, \infty)$,*

$$\left.\frac{db^x}{dx}\right|_{x=a} = (b^a)\left(\left.\frac{db^x}{dx}\right|_{x=0}\right),$$

provided the derivative on the right exists.

Proof. We will use the laws of exponentiation. The increment in b^x from $x = a$ to $x = a + h$ is

$$\begin{aligned}\triangle b^x &:= b^{a+h} - b^a \\ &= b^a b^h - b^a \\ &= (b^a)\left(b^h - 1\right)\end{aligned}$$

and since $1 = b^0$ for any base b, the quantity inside the parentheses at the end of the last expression is clearly the increment in b^x from $x = 0$ to $x = 0 + h$. But then, by definition,

$$\begin{aligned}\left.\frac{db^x}{dx}\right|_{x=a} &:= \lim_{h\to 0}\frac{\triangle b^x}{\triangle x} \\ &= \lim_{h\to 0}(b^a)\left(\frac{b^h - b^0}{h}\right) \\ &= (b^a)\lim_{h\to 0}\left(\frac{b^{0+h} - b^0}{h}\right) \\ &= (b^a)\left(\left.\frac{db^x}{dx}\right|_{x=0}\right).\end{aligned}$$

□

So to establish the differentiability of exponentials, we can focus without loss of generality on showing that they are differentiable at $x = 0$, in other words, that the limit

$$\lim_{x\to 0}\frac{b^x - 1}{x}$$

exists. This will take a little work.

The function φ_b

First, let us investigate the function given by the fraction above: for any $b > 0$, define the function φ_b on $(-\infty, 0) \cup (0, \infty)$ by

$$\varphi_b(x) := \frac{b^x - 1}{x}.$$

We start with a useful observation about this function.

Remark 4.3.2. *The function φ_b is positive if $b > 1$ and negative if $0 < b < 1$.*

If $b > 1$, this is an immediate consequence of the fact that $b^x - 1$ is positive for x positive and negative for x negative, so the ratio is always positive; the argument for $0 < b < 1$ is analogous.
The following numerical observation will help us in establishing the monotonicity of φ_b.

Lemma 4.3.3. *For $b > 1$ and n any nonzero integer,*

$$\frac{b^{n+1} - 1}{b^n - 1} \geq \frac{n+1}{n}. \tag{4.19}$$

The inequality is strict for $n \neq -1$.

Proof. We establish Equation (4.19) first for $n > 0$. Using the fact that the numerator and denominator are both divisible by $b - 1$, we write

$$\begin{aligned} \frac{b^{n+1} - 1}{b^n - 1} &= \frac{b^n + b^{n-1} + ... + b + 1}{b^{n-1} + ... + b + 1} \\ &= \frac{b^n}{b^{n-1} + ... + b + 1} + 1 \\ &= \frac{1}{b^{-1} + b^{-2} ... + b^{-n}} + 1 \end{aligned}$$

Since $b^{-k} < 1$ for $b > 1$ and $k > 0$, each of the n terms in the last denominator is less than one, so

$$\frac{1}{b^{-1} + b^{-2} ... + b^{-n}} > \frac{1}{n}$$

and

$$\frac{b^{n+1} - 1}{b^n - 1} > \frac{1}{n} + 1 = \frac{n+1}{n},$$

as required.
For the case $n < 0$, we write $n = -m$ where m is a positive integer, and using reasoning parallel to the above, obtain

$$\begin{aligned}
\frac{b^{n+1}-1}{b^n-1} &= \frac{b^{-m+1}-1}{b^{-m}-1} \\
&= \frac{b^{-m}(b-b^m)}{b^{-m}(1-b^m)} = \frac{b-b^m}{1-b^m} \\
&= \frac{b-1}{1-b^m} + 1 = 1 - \frac{1-b}{1-b^m} \\
&= 1 - \frac{1}{1+...+b^{m-1}} \\
&> 1 - \frac{1}{m} = \frac{m-1}{m} = \frac{1-m}{m} \\
&= \frac{1+n}{n}.
\end{aligned}$$

□

From this, we will establish the monotonicity of φ_b, which underpins our whole argument for the differentiability of exponentials.

Lemma 4.3.4. *For every positive $b \neq 1$, φ_b is strictly increasing on its natural domain $(-\infty, 0) \cup (0, \infty)$.*

Proof. $\boldsymbol{b > 1}$: We concentrate first on $b > 1$, proceeding in steps similar to those in the definition of b^x.

Positive (Negative) Integers: First, we prove φ_b is strictly increasing on the *positive integers*. For this, it suffices to prove that for any integer $n > 0$, $\varphi_b(n+1) > \varphi_b(n)$ (use induction!). Since both quantities are positive, we can test their ratio. Using Equation (4.19), we have

$$\frac{\varphi_b(n+1)}{\varphi_b(n)} = \frac{b^{n+1}-1}{n+1} \cdot \frac{n}{b^n-1} = \frac{b^{n+1}-1}{b^n-1} \cdot \frac{n}{n+1} > \frac{n+1}{n} \cdot \frac{n}{n+1} = 1.$$

Note that the same argument gives φ_b strictly increasing on the *negative integers* (where we have to assume $n \neq -1$ to insure that $n+1$ is a negative integer).

Positive (Negative) Rationals: Next, we show that φ_b is increasing on the *positive rationals* (and on the *negative rationals*). If $x < x'$ are both rationals, we can express them as fractions with the same positive denominator

$$x = \frac{p}{q},\ x' = \frac{p'}{q}, \quad q > 0$$

with $p < p'$, and of the same sign if x and x' have the same sign. A simple calculation shows that

$$\varphi_b\left(\frac{p}{q}\right) = \frac{b^{p/q} - 1}{p/q} = q\frac{(b^{1/q})^p - 1}{p} = q\varphi_{b^{1/q}}(p)$$

from which it follows, using the case of integers, that if $p < p'$ are integers of the same sign, then

$$\varphi_b\left(\frac{p}{q}\right) = q\varphi_{b^{1/q}}(p) < q\varphi_{b^{1/q}}(p') = \varphi_b\left(\frac{p'}{q}\right).$$

Positive (Negative) Reals: To extend this to each of the intervals $(-\infty, 0)$ and $(0, \infty)$, we will of course use limits. Suppose $x < x'$ are reals of the same sign. Take two sequences of rationals, $\{x_k\}$ *strictly increasing* with limit x, and $\{x'_k\}$ *strictly decreasing* with limit x'. Then we automatically have $x_k < x < x' < x'_k$, hence $\varphi_b(x_k) < \varphi_b(x'_k)$, and so

$$\varphi_b(x) = \varphi_b(\lim x_k) = \lim \varphi_b(x_k) \le \lim \varphi_b(x'_k) = \varphi_b(\lim x'_k) = \varphi_b(x').$$

However, we will really need *strict* inequality in our conclusion. To insure this, we note that, since $x < x'$, there exist two rational numbers separating them: $x < \frac{p}{q} < \frac{p'}{q'} < x'$, and as above we can conclude that

$$\varphi_b(x) \le \varphi_b\left(\frac{p}{q}\right) < \varphi_b\left(\frac{p'}{q'}\right) \le \varphi_b(x')$$

as required.

Monotonicity on Nonzero Reals: So far, we have established that (for $b > 1$) φ_b is strictly increasing on each of the intervals $(-\infty, 0)$ and $(0, \infty)$; to show that it is strictly increasing on their union, we need to compare values in the two intervals. Another easy calculation shows that for any $b > 0$ and $x \neq 0$

$$\varphi_b(-x) = \frac{b^{-x} - 1}{-x} = b^{-x}\frac{1 - b^x}{-x} = b^{-x}\frac{b^x - 1}{x} = b^{-x}\varphi_b(x). \quad (4.20)$$

In particular, if $b > 1$ and $x > 0$ then $0 < b^{-x} < 1$ and so, since both $\varphi_b(-x)$ and $\varphi_b(x)$ are positive, we can conclude that

$$\varphi_b(-x) = b^{-x}\varphi_b(x) < \varphi_b(x) \quad (b > 1,\, x > 0). \tag{4.21}$$

Thus, given $-x < 0 < x'$, we can use Equation (4.21) to compare $\varphi_b(-x)$ with $\varphi_b(x)$ or $\varphi_b(-x')$ with $\varphi_b(x')$ (whichever of $x = |-x|$ and $x' = |x'|$ is smaller), and then comparison of $\varphi_b(-x)$ with this intermediate value and then with $\varphi_b(x')$ completes the required inequality.

$0 < b < 1$: This shows that φ_b is strictly increasing on its natural domain when $b > 1$. To prove the assertion for $0 < b < 1$, we take advantage of the fact that in this case $\frac{1}{b} > 1$, together with another convenient relation between $\varphi_b(x)$ and $\varphi_{\frac{1}{b}}(x)$:

$$\varphi_{\frac{1}{b}}(x) = \frac{(\frac{1}{b})^x - 1}{x} = \frac{b^{-x} - 1}{x} = -\frac{b^{-x} - 1}{-x} = -\varphi_b(-x)\,. \tag{4.22}$$

It is straightforward to show that if a function f is strictly *increasing*, then so is $-f(-x)$ (Exercise 14), and Equation (4.22) shows that we have this situation here.

□

Corollary 4.3.5. *For each $b > 0$, the limit*

$$\psi(b) := \left.\frac{db^x}{dx}\right|_{x=0} = \lim_{x\to 0} \varphi_b(x)$$

exists, and hence $f(x) = b^x$ is differentiable everywhere, with derivative at $x = a$ given by

$$\frac{d}{dx}\left[b^x\right] = \psi(b) \cdot b^x. \tag{4.23}$$

Proof. It is immediate that $\psi(1) = 0$; for all other positive values of b, the monotonicity of φ_b on $(-\infty, 0) \cup (0, \infty)$ implies that the two one-sided limits of φ_b at $x = 0$ exist (and are finite). Furthermore, taking limits as $x \to 0^+$ in Equation (4.20), we have equality of the two one-sided limits:

$$\lim_{x\to 0^-} \varphi_b(x) = \lim_{x\to 0^+} \varphi_b(-x) = \left(\lim_{x\to 0^+} b^{-x}\right) \lim_{x\to 0^+} \varphi_b(x) = \lim_{x\to 0^+} \varphi_b(x)\,.$$

This establishes the existence of the limit

$$\psi(b) = \lim_{x\to 0} \varphi_b(x)$$

and by Lemma 4.3.1 proves the differentiation formula (4.23) . □

This formula tells us that the differentiation formula for any of the exponential functions b^x is completely determined by the value of $\psi(b)$—in other words, by the slope of the tangent line to the graph of b^x at the y-intercept $(0,1)$.

The function $\psi(b)$ and the number e

Next, we wish to show that by a judicious choice of the base b, we can get $\psi(b)=1$, yielding a particularly simple differentiation formula.

Lemma 4.3.6. *1. For $b>1$, $\psi(b)>0$ while for $0<b<1$ $\psi(b)<0$.*

2. For any $b>0$ and $r\in\mathbb{R}$, $\psi(b^r)=r\psi(b)$.
In particular, there exists a value of b for which

$$\psi(b)=1.$$

Note that Remark 4.3.2 automatically gives us a weak inequality in the first statement above; however, we will need the strict inequality in what follows.

Proof. To prove the first assertion, consider the values of φ_b along the sequence $\{\frac{1}{n}\}$:

$$\varphi_b\left(\frac{1}{n}\right)=\frac{b^{\frac{1}{n}}-1}{\frac{1}{n}}=n(b^{\frac{1}{n}}-1)$$

and using the factoring of $b-1$ as

$$b-1=(b^{\frac{1}{n}})^n-1=\left(b^{\frac{1}{n}}-1\right)\left((b^{\frac{1}{n}})^{n-1}+(b^{\frac{1}{n}})^{n-2}+...+b^{\frac{1}{n}}+1\right),$$

we have

$$\varphi_b\left(\frac{1}{n}\right)=\frac{n(b-1)}{b^{\frac{n-1}{n}}+b^{\frac{n-2}{n}}+...+b^{\frac{1}{n}}+1}.$$

Notice that each term in the denominator can be written as $b^{1-\frac{k}{n}}$ for some $i=0,\dots,n-1$; thus, if $b>1$ each term is less than b and (since there are n terms in the denominator)

$$\varphi_b\left(\frac{1}{n}\right)>\frac{b-1}{b}>0.$$

Since this bound is independent of n, it applies to the limit as well, giving the first assertion when $b>1$. The argument when $0<b<1$ simply reverses the inequalities.

The second assertion follows from taking $x \to 0$ in the identity

$$\varphi_{b^r}(x) = \frac{b^{rx} - 1}{x} = r\varphi_b(rx)\,.$$

Finally, we pick any base $b > 1$, say $b = 2$: then we have for all $r \in \mathbb{R}$,

$$\psi(2^r) = r\psi(2)$$

but by the first assertion

$$\psi(2) > 0$$

so taking $r = \frac{1}{\psi(2)}$, we obtain the desired value $b = 2^r$. □

Lemma 4.3.6 establishes the existence of a (unique) value of b for which $\psi(b) = 1$. We name this special value[7] $b = \boldsymbol{e}$—that is:

Definition 4.3.7. $\boldsymbol{e}$ *is the unique number satisfying*

$$\lim_{h\to 0} \frac{e^h - 1}{h} = 1. \tag{4.24}$$

With this definition, we immediately get the differentiation formula for e^x:

$$\frac{de^x}{dx} = e^x. \tag{4.25}$$

Natural Logarithms

The simplicity of this differentiation formula makes e^x the **natural exponential** for doing calculus. The inverse function to e^x, $\log_e x$, is called the **natural logarithm** and often denoted

$$\ln x := \log_e x.$$

How are logarithms with other bases related to $\ln x$? Applying $\ln x$ to both sides of the identity $x = b^{\log_b x}$, we see that

$$\ln x = \ln(b^{\log_b x}) = (\log_b x)(\ln b)$$

or

$$\log_b x = \frac{\ln x}{\ln b}. \tag{4.26}$$

[7]This is the notation used by **E**uler to define the **e**xponential function—you decide what $\boldsymbol{e}$ stands for .

(In fact, a similar calculation relates the logarithms of x with respect to any pair of bases.) This translates into the formula

$$b^x = e^{x \ln b}. \tag{4.27}$$

From this it is easy to deduce that for any base $b > 0$

$$\begin{aligned} \lim_{h\to 0} \frac{b^h - b^0}{h} &= \lim_{h\to 0} \frac{e^{h\ln b} - 1}{h} \\ &= \lim_{y=h\ln b\to 0} (\ln b) \frac{e^{y=h\ln b} - 1}{y = h\ln b} \\ &= \ln b \end{aligned}$$

and so in view of Lemma 4.3.1 we have the differentiation formula for any exponential map

$$\frac{d}{dx}\left[b^x\right] = (\ln b)(b^x). \tag{4.28}$$

Exercises for § 4.3

Answers to Exercises 1-2ace, 3ac are given in Appendix B.

Practice problems:

1. Use the formal rules of this section to calculate the derivative of each function f at the given point $x = a$:

   (a) $f(x) = xe^x$, $a = \ln 2$
   (b) $f(x) = x^2 e^x$, $a = 1$
   (c) $f(x) = e^x \cos x$, $a = 0$
   (d) $f(x) = (x^2 + 1)e^x \sin x$, $a = 0$
   (e) $f(x) = \dfrac{e^x - 1}{e^x + 1}$, $a = 0$
   (f) $f(x) = \dfrac{e^{2x} - 1}{e^x + 1}$, $a = 0$ (*Hint:* Use some identities.)

2. Find the equation of the line tangent to the graph of the given function at the given point, for each item in the preceding problem.

3. For each function f defined in pieces below, find the function $f'(x)$. In particular, determine whether the derivative exists at the interface point, and if so, find its value.

(a) $f(x) = \begin{cases} e^x & \text{for } x < 0, \\ x+1 & \text{for } x \geq 0. \end{cases}$ (b) $f(x) = \begin{cases} x+1 & \text{for } x < 0, \\ e^x & \text{for } x \geq 0. \end{cases}$

(c) $f(x) = \begin{cases} \frac{e^{2x}-1}{e^x-1} & \text{for } x < 0, \\ x & \text{for } x \geq 0. \end{cases}$ (d) $f(x) = \begin{cases} \frac{x}{e^x} & \text{for } x < 0, \\ x & \text{for } x \geq 0. \end{cases}$

Theory problems:

4. (a) Find the derivative at $x = 0$ of the function e^{2x}, from the definition of derivative.
   (b) Use the Product Rule to calculate the derivative of the function e^{2x}.
   (c) Use either of these approaches to find the derivative of the function e^{nx} when n is a natural number.
   (d) Use identities for exponentiation to show that
   $$\frac{d}{dx}\left[e^{ax}\right] = ae^{ax}$$
   for any real number a.
5. Use Lemma 4.3.6(2) and the definition of the natural logarithm to show that
   $$\psi(b) = \ln b$$
   for any $b > 0$.

4.4 Formal Differentiation III: Inverse Functions

Recall from § 3.2 that when f is continuous and strictly monotone on an interval $[a, b]$, then its *inverse* is the function f^{-1} defined on $[A, B]$ (where $\{A, B\} = \{f(a), f(b)\}$) by

$$y = f^{-1}(x) \iff x = f(y).$$

We saw in Proposition 3.2.5 that f^{-1} is continuous. Suppose f is differentiable; what about f^{-1}?
Let us think for a moment about the relation between the graph of an invertible function and the graph of its inverse. The graph of f^{-1} is the set of points (x, y) whose coordinates satisfy $y = f^{-1}(x)$; by definition, this equation is the same as $f(y) = x$, which is the equation defining the graph

of f, *except that the variables x and y are interchanged.* Geometrically, this interchange corresponds to reflecting the plane across the **diagonal line** $y = x$. So the whole graph $y = f^{-1}(x)$ is just *the reflection across the diagonal line* of the graph $y = f(x)$ (Figure 4.17). This reflection process should also carry any secant or line tangent to the graph of f at a point (x_0, y_0) into a secant or a line tangent to the graph of f^{-1} at the corresponding point $y_0, x_0)$ (Figure 4.17). Note that this reflection takes

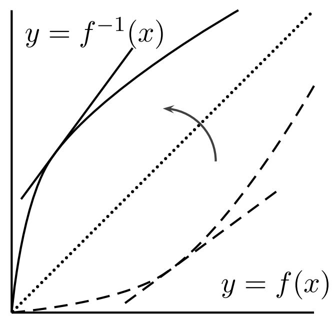


Figure 4.17: Graph of f^{-1} and f, and Tangent Lines

horizontal lines to *vertical* lines: thus if $f'(x_0) = 0$, then the graph of f^{-1} at $x = y_0$ will have a *vertical* tangent, which means f^{-1} will not be differentiable there. However, with this one cautionary proviso, we see that differentiablility of f at a point x_0 should imply differentiability of f^{-1} at its image $y_0 = f(x_0)$.
The statement of this theorem is a little more natural if we deviate from our practice of consistently using x for the input into any function and, guided by the reflection picture, think of $y = f(x)$ as a function of x, but think of its inverse as a function $f^{-1}(y)$ of y.

Theorem 4.4.1 (Derivative of Inverse)**.** *Suppose f is invertible on an interval around $x = x_0$, and $f'(x_0) \neq 0$.*
Then $g(y) := f^{-1}(y)$ is differentiable at $y = y_0 := f(x_0)$, and

$$g'(y_0) = \frac{1}{f'(x_0)}. \tag{4.29}$$

Stated differently, if $y = f(x)$ and $x = f^{-1}(y)$, then

$$\left.\frac{dy}{dx}\right|_{x=x_0} = \left(\left.\frac{dx}{dy}\right|_{y=y_0=f(x_0)}\right)^{-1}.$$

Proof. We essentially translate the reflection picture into formulas. Using the notation of the statement, we can write the definition of $g'(y_0)$ as

$$g'(y_0) := \lim_{y\to y_0} \frac{\triangle g}{\triangle y} = \lim_{y\to y_0} \frac{g(y)-g(y_0)}{y-y_0}.$$

We are interested in the limit of the sequence obtained by picking any sequence $y_k \to y_0$ (of points distinct from y_0) and substituting $y = y_k$ in the difference quotient above. For any such sequence, let $x_k := g(y_k)$; by definition of $g(y) = f^{-1}(y)$, we have the relation $y_k = f(x_k)$ for all k. But then we can write

$$\begin{aligned}\lim \frac{g(y_k)-g(y_0)}{y_k-y_0} &= \lim \frac{x_k-x_0}{f(y_k)-f(x_0)} \\ &= \lim \left(\frac{f(y_k)-f(x_0)}{x_k-x_0}\right)^{-1} \\ &= \left(\lim \frac{f(y_k)-f(x_0)}{x_k-x_0}\right)^{-1}\end{aligned}$$

and by assumption, the limit of the sequence of fractions inside the large parentheses is $f'(x_0)$ (for *any* sequence $x_k \to x_0$). But this means we have shown that for any sequence $y_k \to y_0$, the sequence of difference quotients $\frac{\triangle g}{\triangle y}$ converges to $\frac{1}{f'(x_0)}$, and hence this is the derivative of $g(y)$ at $y = y_0$. □

Theorem 4.4.1 immediately gives us a number of new differentiation formulas.
Perhaps the most important one is for the natural logarithm. We saw that

$$\frac{d}{dx}[e^x] = e^x$$

and since the inverse of $y = e^x$ is $x = \ln y$, we have

Corollary 4.4.2.

$$\frac{d}{dy}[\ln y] = \frac{1}{y}. \tag{4.30}$$

Proof. Setting $y = f(x) = e^x$, we have $f'(x) = e^x = y$, so if $x = g(y) = \ln y = f^{-1}(y)$, Theorem 4.4.1 gives us

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{y}.$$

□

From this and Equation (4.26), we can differentiate the logarithm function with any positive base $b \neq 1$ (Exercise 4):

$$\frac{d}{dx}[\log_b x] = \frac{1}{x \ln b} \quad (b \in (0,1) \cup (1,\infty)). \tag{4.31}$$

Another important class of differentiable functions is the inverse trigonometric functions. Recall that the arcsine of a number $-1 \leq x \leq 1$ is defined to be the angle $\theta = \arcsin x$ lying in the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ with $x = \sin\theta$, while the arccosine is defined similarly, except that the angle is taken in the interval $[0, \pi]$. Since we cannot define these functions for $|x| > 1$, we cannot try to differentiate either function at the endpoints of its domain.

Corollary 4.4.3. *For* $|x| < 1$

$$\frac{d}{dx}[\arcsin x] = \frac{1}{\sqrt{1-x^2}} \tag{4.32}$$

$$\frac{d}{dx}[\arccos x] = -\frac{1}{\sqrt{1-x^2}}. \tag{4.33}$$

Proof. We prove the formula for $g(x) = \arcsin x$ and leave the other case to you (Exercise 6).
Since $\theta = \arcsin x$ is the inverse of $f(\theta) = \sin\theta$, Theorem 4.4.1 gives us

$$\frac{d}{dx}[\arcsin x] = \frac{1}{f'(\theta)} = \frac{1}{\cos\theta}.$$

To proceed, we need to express $\cos\theta$ in terms of $x = \sin\theta$. The identity $\cos^2\theta + \sin^2\theta$ tells us that

$$\cos\theta = \pm\sqrt{1-\sin^2\theta} = \pm\sqrt{1-x^2},$$

but we need to determine which of the two signs applies. However, we know that $\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, and this insures that $\cos\theta > 0$, so the positive sign applies, and the required formula follows.
A similar argument (Exercise 6) gives the formula for $\arccos x$. □

The derivative of the arctangent can be established in a similar way. Recall that $\arctan x$ is defined for *all* x as the angle $\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ satisfying $x = \tan\theta$.

Corollary 4.4.4. *For all x,*

$$\frac{d}{dx}[\arctan x] = \frac{1}{x^2+1}. \tag{4.34}$$

Proof. Since $\theta = \arctan x$ is the inverse of $x = f(\theta) = \tan\theta$, Theorem 4.4.1 gives us

$$\frac{d}{dx}[\arctan x] = \frac{1}{f'(\theta)} = \frac{1}{\sec^2\theta}$$

but the trigonometric identity $\tan^2\theta + 1 = \sec^2\theta$ gives us that in this case $\sec^2\theta = \tan^2\theta + 1 = x^2 + 1$, and Equation (4.34) follows. □

These last few differentiation formulas are remarkable: they relate (the derivative of) a function defined geometrically to an algebraic function. One final example is the arcsecant function; this is defined for $|x| > 1$; when $x > 1$, it is natural to take $\theta = \operatorname{arcsec} x$ (the angle satisfying $x = \sec\theta$ in the range $\theta \in \left(0, \frac{\pi}{2}\right)$ (where both $\cos\theta$ and$\sin\theta$ are positive), but the choice is less clear when $x < -1$, and different conventions about this choice lead to different differentiation formulas [8] for the arcsecant. We will therefore restrict ourselves to considering the derivative of this function only for $x > 1$ (see Exercise 7).

Corollary 4.4.5. *For $x > 1$,*

$$\frac{d}{dx}[\operatorname{arcsec} x] = \frac{1}{x\sqrt{x^2-1}}. \tag{4.35}$$

Proof. Recall that the derivative of $f(\theta) = \sec\theta$ is $f'(\theta) = \sec\theta\tan\theta$. If $x = \sec\theta$, then the identity $\tan^2\theta + 1 = \sec^2\theta$ leads to $\tan\theta = \pm\sqrt{\sec^2 - 1} = \pm\sqrt{x^2-1}$, but if we agree that for $x > 1$ we pick $\theta \in \left(0, \frac{\pi}{2}\right)$, then $\tan\theta > 0$ so the plus sign applies; then as before substituting into Equation (4.29) we get

$$\frac{d}{dx}[\arcsin x] = \frac{1}{\sec\theta\tan\theta} = \frac{1}{x\sqrt{x^2-1}}$$

which is the desired formula. □

Exercises for § 4.4

[8] but differing only by sign

Answers to Exercises 1-2acegi, 3a are given in Appendix B.

Practice problems:

1. For each given function f, find its inverse f^{-1}, specifying the domain of f^{-1}:

   (a) $f(x) = x + 1$
   (b) $f(x) = 3x + 2$
   (c) $f(x) = x^3 - 1$
   (d) $f(x) = \ln x + \ln 2$
   (e) $f(x) = \dfrac{x+1}{x+2}$
   (f) $f(x) = \dfrac{x-1}{x+1}$
   (g) $f(x) = \sqrt{x}$
   (h) $f(x) = \sqrt[3]{8x}$
   (i) $f(x) = e^x + 1$
   (j) $f(x) = e^x - 1$

2. Find the derivative of the given function f at the given point $x = a$:

   (a) $f(x) = x \ln x$, $a = e$
   (b) $f(x) = e^x \ln x$, $a = 1$
   (c) $f(x) = (x^2 + 1) \ln x$, $a = 1$
   (d) $f(x) = \dfrac{\ln x}{x}$, $a = 1$
   (e) $f(x) = x \arctan x$, $a = 1$
   (f) $f(x) = e^x \arccos x$, $a = \frac{1}{2}$
   (g) $f(x) = (\ln x)(\cos x)$, $a = \frac{\pi}{2}$
   (h) $f(x) = e^x \arctan x$, $a = 0$
   (i) $f(x) = \log_2 x$, $a = 8$ (*Hint:* Use some identities) .
   (j) $f(x) = \ln(xe^x)$, $a = 1$ (*Hint:* Use some identities) .

3. Use Theorem 4.4.1 to find the equation of the tangent line to the curve $y = f^{-1}(x)$ at the point $(f(a), a)$. You may assume without proof that the function f has an inverse $g(x)$ near $x = f(a)$ with $g(f(a)) = a$.

   (a) $f(x) = x^3 + 2x + 1$, $a = 1$
   (b) $f(x) = \dfrac{x^3 + 1}{x^5 - 8}$, $a = -1$

Theory problems:

4. Prove Equation (4.31).

5. Show that the formula
$$\frac{d}{dx}[\ln|x|] = \frac{1}{x}$$
holds for all $x \neq 0$.

6. Mimic the proof of Corollary 4.4.3 to prove Equation (4.33).

Challenge problems:

7. As mentioned in the text, there are two "natural" choices for the value of $\operatorname{arcsec} x$ when $x < -1$: either (i) $\frac{\pi}{2} < \operatorname{arcsec} x < \pi$, or (ii) $\pi < \operatorname{arcsec} x < \frac{3\pi}{2}$. Note that the first one is "natural" in the sense that it corresponds to defining $\operatorname{arcsec} x$ to be $\arccos \frac{1}{x}$.
   (a) Show that the first choice leads to the formula
   $$\frac{d}{dx}[\operatorname{arcsec} x] = \frac{1}{|x|\sqrt{x^2-1}}.$$
   (*Hint:* For $x < -1$ this choice gives $\tan\theta = -\sqrt{x^2-1}$.)
   (b) Show that the second choice is "natural" in that it leads to the formula
   $$\frac{d}{dx}[\operatorname{arcsec} x] = \frac{1}{x\sqrt{x^2-1}},$$
   as in Corollary 4.4.5.

8. Use Theorem 4.4.1 to give a different proof of Proposition 4.2.7.

4.5 Formal Differentiation IV: The Chain Rule and Implicit Differentiation

The basic differentiation formulas in the previous section are the building blocks for differentiating a broad variety of functions. The Chain Rule extends our repertoire considerably, allowing us to differentiate a composition using the derivatives of the functions being composed. Suppose we have two functions, f and g, and form the composition

$$h := g \circ f.$$

How is $h'(x)$ related to $f'(x)$ and $g'(x)$? To see what we should expect, let us recall the affine approximation to a function at $x = a$; for f this is

$$T_a f(x) := f(a) + f'(a)(x-a);$$

that is, $y = f(x)$, for x near $x = a$, is approximately equal to $T_a f(x)$. Now, by definition $h(x) = g(y)$, and if x is near $x = a$ then by continuity

$y = f(x)$ is near $y = b := f(a)$. To approximate the value of g at such a y, we expect to use the analogous affine approximation to $g(y)$ at $y = b := f(a)$, namely

$$T_g b(y) = g(b) + g'(b)\,(y - b).$$

Now, it seems reasonable that the *composition of the approximations* to the two functions should be a good *approximation to the composition* of the functions, so we would expect the affine approximation to $h(x)$ to have the form

$$\begin{aligned}((T_b g)\circ(T_a f))(x) &= T_b g\,(T_a f(x)) \\ &= g(b) + g'(b)\,(T_a f(x) - b) \\ &= g(b) + g'(b)\,\big(f(a) + f'(a)\,(x-a) - b\big) \\ &= g(b) + \big(g'(b)\,f'(a)\big)\,(x-a)\end{aligned}$$

where we have used the definition $b = f(a)$ to cancel two terms in the last line. Note also that because of this definition, $g(b) = g(f(a)) = h(a)$, so the first term above matches the first term in the formal expression for the affine approximation to $h(x)$

$$T_h a(x) = h(a) + h'(a)\,(x - a).$$

Hence, we expect that the second terms will also match:

$$h'(a) = g'(b)\,f'(a) = g'(f(a)) \cdot f'(a)\,.$$

We should be careful, of course, to check that the composition actually is differentiable, and then that our formal matching (and assumption that the affine approximation of a composition is the composition of the corresponding affine approximations) is correct.

Theorem 4.5.1 (Chain Rule)**.** *Suppose f is differentiable at $x = a$ and g is differentiable at the image $y = b := f(a)$.*
Then the composition

$$h(x) = (g\circ f)(x) := g(f(x))$$

is differentiable at $x = a$, and its affine approximation at $x = a$ is the composition of the affine approximation to f at $x = a$ with the affine approximation to g at $y = b = f(a)$:

$$T_{(g\circ f)} a(x) = T_h a(x) = T_g b(T_a f(x)) = ((T_b g)\circ(T_a f))(x).$$

In particular, the derivative at $x = a$ of $h = g \circ f$ is the product of the derivatives of f and g at their associated points, $x = a$ (resp. $y = b := f(a)$):

$$h'(a) = (g \circ f)'(a) = g'(b) \cdot f'(a) = g'(f(a)) \cdot f'(a) . \tag{4.36}$$

The conclusion of Theorem 4.5.1, Equation (4.36), can be summarized in the following easy-to-remember form: if

$$y = f(x)$$

and

$$z = g(y)$$

then

$$\frac{dz}{dx} = \left(\frac{dz}{dy}\right) \cdot \left(\frac{dy}{dx}\right) .$$

Proof. Let us use the notation as above

$$\begin{aligned} y &= f(x) \\ z &= h(x) = g(y) = g(f(x)) \end{aligned}$$

and set $b = f(a)$, $c = h(a) = g(b)$. We need to calculate

$$h'(a) := \lim_{\triangle x \to 0} \frac{\triangle z}{\triangle x}.$$

For reasons that will become clear soon, let us calculate this in terms of sequences. Suppose $\{x_k\}$ is a sequence of points in the domain of f distinct from but converging to a; define the sequences

$$\begin{array}{ccc} & y_k := f(x_k) & z_k := h(x_k) = g(y_k) \\ \triangle x_k := x_k - a & \triangle y_k := y_k - b & \triangle z_k := z_k - c. \end{array}$$

Note that since f, g and hence h are continuous (because the first two are differentiable) we automatically have $\triangle x_k, \triangle y_k, \triangle z_k \to 0$. We need to show that

$$\frac{\triangle z_k}{\triangle x_k} \to g'(b) \cdot f'(a) . \tag{4.37}$$

Now, if $\triangle y_k \neq 0$ then we can write

$$\frac{\triangle z_k}{\triangle x_k} = \frac{\triangle z_k}{\triangle y_k} \cdot \frac{\triangle y_k}{\triangle x_k};$$

and then we can write

$$\lim \frac{\triangle z_k}{\triangle x_k} = \left(\lim \frac{\triangle z_k}{\triangle y_k}\right) \cdot \left(\lim \frac{\triangle y_k}{\triangle x_k}\right) = g'(b) \cdot f'(a)\,.$$

But just because $x_k \neq a$ we can't be sure $\triangle y_k \neq 0$ (or equivalently $y_k \neq b$). However, if there *is* some sequence $x_k \to a$ with $x_k \neq a$, but $y_k = b$, for all k, then since $\{y_k\}$ is a constant sequence, so is $\{z_k\}$, which means in particular that both $\{\triangle y_k\}$ and $\{\triangle z_k\}$ are constant *zero* sequences and hence both have limit 0. In particular this means that $f'(a) = 0$ and so (4.37) holds in this case, as well. Note that in fact $f(a) = 0$ if any *sub*sequence of $\{y_k\}$ is constant (argue as above just with the subsequence). Of course, if there is no such (sub)sequence then eventually $\triangle y_k \neq 0$, and our argument works.

We have now shown that any sequence $x_k \to a$ with $x_k \neq a$ for all k satisfies (4.37), and so we have Equation (4.36). □

The Chain Rule greatly enlarges the repertoire of functions we can formally differentiate. We consider two examples.

First consider the function

$$h(x) = \sqrt{x^2 + 3x + 26}.$$

This can be recast as the composition $h(x) = g(f(x)) = (g \circ f)(x)$, where

$$f(x) := x^2 + 3x + 26$$
$$g(y) := \sqrt{y}.$$

To calculate the derivative of h at, say, $x = 2$, we first calculate

$$f'(x) = 2x + 3$$
$$g'(y) = \frac{1}{2\sqrt{y}}$$

and substituting $x = 2$, $y = f(2) = (2)^2 + 3(2) + 26 = 36$, we have

$$f'(2) = 2(2) + 3 = 7$$
$$g'(36) = \frac{1}{2\sqrt{36}} = \frac{1}{12}$$

so the Chain Rule gives

$$h'(2) = \frac{1}{12} \cdot (7) = \frac{7}{12}.$$

As a second example, let

$$h(x) = \sin\left(\sqrt{x^2 - 3x + 1}\right).$$

This can be cast as the composition $h(x) = (g \circ f)(x)$ where

$$f(x) := \sqrt{x^2 - 3x + 1}$$
$$g(y) := \sin y.$$

Now it is easy to see that

$$g'(y) = \cos y$$

but to differentiate f we need to cast *it* as a compsition, say

$$f(x) = (s \circ t)(x)$$

where

$$t(x) := x^2 - 3x + 1$$
$$s(t) := \sqrt{t}.$$

This lets us differentiate f using the Chain Rule:

$$f'(x) = s'(t) \cdot t'(x) = \frac{1}{2\sqrt{t}} \cdot (2x - 3) = \frac{2x - 3}{2\sqrt{x^2 - 3x + 1}}.$$

Substituting this into the Chain Rule for $g \circ f$, we have

$$h'(x) = g'(f(x)) \cdot f'(x) = \left(\cos\sqrt{x^2 - 3x + 1}\right) \cdot \frac{2x - 3}{2\sqrt{x^2 - 3x + 1}}$$
$$= \frac{(2x - 3)\cos\sqrt{x^2 - 3x + 1}}{2\sqrt{x^2 - 3x + 1}}.$$

A moment's reflection should convince you that what we have done is to differentiate a threefold composition (sine *of* the square root *of* $x^2 - 3x_{+}1$) and obtained a threefold product: the derivative of sin, evaluated at $\sqrt{x^2 - 3x + 1}$, times the derivative of the square root function, evaluated at $x^2 - 3x + 1$, times the derivative of $x^2 - 3x + 1$ (evaluated at x). This can be visualized as a process of removing parentheses: to differentiate a composition, we first differentiate the *outermost* function, evaluating it at the point to which it is being applied, then we differentiate the *next* outermost function, evaluating *it* at the point to which *it* is being applied, and so on...

Implicit Differentiation

There is a second application of the chain rule, called **implicit differentiation**. Let us illustrate with a concrete example. Consider the curve given by the equation

$$3x^2 - 2xy + y^2 = 48.$$

You can check that the point $(-4, 0)$ lies on this curve; what is the equation of its tangent line there?

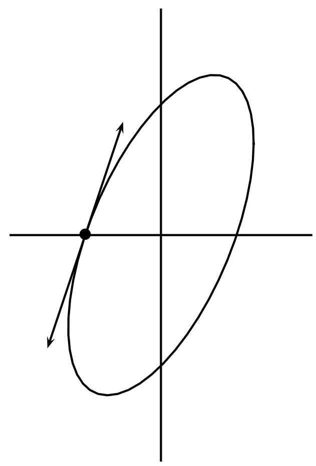


Figure 4.18: $3x^2 - 2xy + y^2$: Tangent at $(-4, 0)$

If we could express this curve as the graph of a function, say $y = f(x)$, then we would simply differentiate the function at $x = -4$; this would give us the slope of the tangent line, and we could use the point-slope formula (the line goes through $(-4, 0)$) to write down the equation. But we *don't* have the function[9] f, so how can we differentiate it? Well, *if* we had f, then the equation of the curve would be a statement about the relation between the value $y = f(x)$ and x itself, at any point on the curve. Ignoring the fact that we don't *know* $f(x)$ itself, we can still perform a formal differentiation on both sides of the equation, simply writing $f'(x)$ (or better yet, $y' = \frac{dy}{dx}$) wherever the (unknown) derivative of our

[9]In fact, the whole curve is an ellipse, and hence is *not* the graph of a function. But the *piece* of it *near the point* is. Nonetheless, finding an expression for this function is not entirely trivial and, as we shall see, unnecessary.

(unknown) function would have to go. That is, we formally apply $\frac{d}{dx}$ to both sides. On the left, we have, by the basic differentiation rules

$$\frac{d}{dx}[3x^2 - 2xy + y^2] = 3\frac{d}{dx}[x^2] - 2\frac{d}{dx}[xy] + \frac{d}{dx}[y^2].$$

The first differentiation is completely straightforward:

$$\frac{d}{dx}[x^2] = 2x.$$

The second is a little more complicated: we are differentiating a product of two functions of x: the "identity" function $u = x$ and the unknown function $v = y = f(x)$. Never mind—we can still apply the Product Rule to conclude that

$$\begin{aligned}\frac{d}{dx}[xy] = \frac{d}{dx}[uv] &= \left(\frac{du}{dx}\right) \cdot (v) + (u) \cdot \left(\frac{dv}{dx}\right) \\ &= \left(\frac{d}{dx}[x]\right) \cdot (y) + (x) \cdot \left(\frac{d}{dx}[y]\right) \\ &= (1)(y) + (x)(y') \\ &= y + xy'.\end{aligned}$$

The third term on the left is even more complicated: it is the composition of the squaring function with the (unknown) function y. But the Chain Rule doesn't care whether we *know* the function or even its derivative—all it says is that the derivative of the square of $y = f(x)$ (as a function of x) is the derivative of the squaring function, evaluated at y, times the derivative of $y = f(x)$ (as a function of x), or

$$\frac{d}{dx}[y^2] = (2y)\left(\frac{d}{dx}[y]\right) = 2yy'.$$

Now, combining all of this, we see that the derivative $\frac{d}{dx}$ of the left-hand side of the equation of the curve is

$$3(2x) - 2(y + xy') + 2yy' = 6x - 2y - 2xy' + 2yy'.$$

The equation of the curve says that the left side (as a function of x) is the same as the right side, which is a *constant* function of x, and hence has derivative zero. Well, this means our expression above for the derivative of the *left* side must match the derivative of the *right* side, in other words, we have the equation involving the three variables x, y, and y'

$$6x - 2y - 2xy' + 2yy' = 0.$$

But at our point, we know the values of two of these variables: $x = -4$ and $y = 0$. Substituting, we have that at our point, the relation which must hold is

$$6(-4) - 2(0) - 2(-4)y' + 2(0)y' = 0.$$

This, of course, can be solved for y':

$$-24 - 0 + 8y' + 0 = 0 \Rightarrow y' = \frac{24}{8} = 3$$

A miracle! We don't know the *function*, but we know its *derivative*, at least at this point! Without spending too much time on Hosannas, we can now proceed to write the equation of the tangent line, which must be a line through the point $(-4, 0)$ with slope 3, so the point-slope formula gives us

$$y - 0 = 3(x - (-4)), \text{ or } y = 3x + 12.$$

In setting forth this technique, we have glossed over some technical difficulties.[10] See Exercise 4.

Logarithmic Differentiation

A particularly powerful version of implicit differentiation is the so-called method of **logarithmic differentiation**. This takes advantage of two facts: that the natural logarithm turns products into sums and powers into multiples, and that its derivative is very simple (Equation (4.30)): combining the derivative formula with the Chain Rule, we obtain for any function $y = f(x) > 0$ the formula

$$\frac{d}{dx}[\ln y] = \frac{d}{dx}[\ln f(x)] = \frac{y'}{y} = \frac{f'(x)}{f(x)}.$$

We illustrate with a simple example first. Let

$$y = \frac{(x^2 - 1)^3}{(x^3 - 2x^2 + x + 3)^2}$$

and let us try to find $y' = \frac{dy}{dx}$ at $x = 2$. Of course we could do this by brute force, using the the power rule, the chain rule, and the quotient rule, but

[10]General conditions under which this technique is justified are given by the *Implicit Function Theorem*, which is dealt with in multivariate calculus. In effect, when application of the technique does not involve division by zero, it can be justified.

the calculations are grotesque. Instead, let us look at a new function,

$$\begin{aligned}\ln y &= \ln\left(\frac{(x^2-1)^3}{(x^3-2x^2+x+3)^2}\right)\\ &= 3\ln(x^2-1) - 2\ln(x^3-2x^2+x+3).\end{aligned}$$

Using our formula on each of the three natural logarithms, we have

$$\begin{aligned}\frac{y'}{y} &= 3\frac{\frac{d}{dx}\left[x^2-1\right]}{x^2-1} - 2\frac{\frac{d}{dx}\left[x^3-2x^2+x+3\right]}{x^3-2x^2+x+3}\\ &= 3\frac{2x}{x^2-1} - 2\frac{3x^2-4x+1}{x^3-2x^2+x+3}.\end{aligned}$$

We want to find y', but let us refrain from doing the algebra to solve for it quite yet; instead, let us evaluate each of these expressions when $x = 2$. At this point we have

$$\begin{aligned}x^2-1 &= (2)^2-1\\ &= 4-1\\ &= 3\\ x^3-2x^2+x+1 &= (2)^3-2(2)^2+2+3\\ &= 8-8+2+3\\ &= 5\\ y &= \frac{3^3}{5^2}\\ &= \frac{27}{25}\\ 2x &= 2(2)\\ &= 4\\ 3x^2-4x+1 &= 3(2)^2-4(2)+1\\ &= 12-8+1\\ &= 5\end{aligned}$$

and substituting all of these values into the preceding equation, we have

$$\frac{y'}{\frac{27}{25}} = 3\frac{4}{3} - 2\frac{5}{5} = 2$$

or

$$y' = 2\frac{27}{25} = \frac{54}{25}.$$

This technique is also very useful in certain more theoretical situations. For example, how would we differentiate the function

$$y = f(x) = x^x, \quad x > 0?$$

We know how to differentiate an exponential expression when either the base or the exponent is constant, but here *both* are changing. However, mimicking our procedure above, we can write

$$\ln y = x \ln x$$

and differentiating both sides (using the Product Rule on the right), we have

$$\frac{y'}{y} = \ln x + x \frac{d}{dx}[\ln x] = \ln x + \frac{x}{x} = \ln x + 1$$

and then multiplying both sides by $y = x^x$ we obtain the formula

$$\frac{d}{dx}[x^x] = x^x(1 + \ln x).$$

Exercises for § 4.5

Answers to Exercises 1acegikmoqsuw, 2ace, 3ab are given in Appendix B.

Practice problems:

1. Use the chain rule to differentiate each function below:

(a) $f(x) = (2x+1)^3$	(b) $f(x) = (x^2+1)^{-3}$
(c) $f(x) = \sqrt{x^2+1}$	(d) $f(x) = \dfrac{(x^2+1)^2}{(x^2-1)^3}$
(e) $f(x) = \sin(2\pi x)$	(f) $f(x) = \cos(\pi x^2)$
(g) $f(x) = e^{x^2/2}$	(h) $f(x) = \ln(\ln x)$
(i) $f(x) = \sin\dfrac{1}{x}$	(j) $f(x) = \ln(e^x + 1)$
(k) $f(x) = e^x \cos 2x$	(l) $f(x) = e^{x \sin 2x}$
(m) $f(x) = \sqrt{\ln(x^2+1)}$	(n) $f(x) = \arctan\sqrt{x^2+1}$
(o) $f(x) = \sqrt{x^2 - x}$	(p) $f(x) = x \sec(x^2 - 1)$
(q) $f(x) = e^{e^x}$	(r) $f(x) = \cos(\arcsin x)$

(s) $f(x) = (3x^2 + 2x + 5)^3(x^3 - 2x)^2$

(t) $f(x) = \ln(\tan x + \sec x)$

(u) $f(x) = (x^{\frac{1}{2}} + x^{\frac{3}{2}} + x^{\frac{1}{3}})^{\frac{2}{3}}$

(v) $f(x) = (4x^{-2} + 2x + 1)^{-\frac{1}{2}}$

(w) $f(x) = \tan\sqrt{x^2 + 1}$

(x) $f(x) = 6(x^3 - x)^{1/3}$

2. For each problem below, you are given the equation of a curve and a point on that curve. Use implicit differentiation to find the equation of the line tangent to the given curve at the given point.

   (a) $3x^2 + 2y^2 = 5$ at $(1, -1)$
   (b) $x^2y^3 = 4$ at $(-2, 1)$
   (c) $x^3 + y^3 + 7 = 0$ at $(-2, 1)$
   (d) $x^4 + 3x^2y + y^3 = 5$ at $(-1, 1)$
   (e) $xy = y^2 + 4\cos \pi x$ at $(3, 4)$
   (f) $x^2e^{y+2} + y^3e^{x-3} = 1$ at $(3, -2)$

3. Use logarithmic differentiation to find $f'(a)$:

   (a) $f(x) = \dfrac{(x^3 + 3x^2 + 2x + 1)^3(x^2 - 5x + 5)^2}{\sqrt{x^2 + 1}}$, $a = 1$
   (b) $f(x) = (1 + x)^{1/x}$, $a = 2$

Theory problems:

4. (a) If you try to use implicit differentiation to find the line tangent to the curve

$$x + y^2 = 1$$

at the point $(1, 0)$, you come up with the nonsense equation

$$0 = 1.$$

Sketch the curve near this point and explain why the method can't work.

   (b) Do the same with the equation

$$x + y^3 = 0$$

at the origin.

(c) What do we need to know to be sure that the method will work?

(d) Suppose we have worked through implicit differentiation for a curve at a point and come up with a reasonable-looking answer. How do we know that this answer is, indeed, correct? (*Hint:* Some of the justification will require Proposition 4.8.3.)

5. Prove the Quotient Rule by means of the Chain Rule and Product Rule. (*Hint:* Write $\frac{f(x)}{g(x)}$ as $f(x) \cdot (g(x))^{-1}$.)

Challenge problem:

6. In this problem, we will see that even if the derivative of a function is defined everywhere on an interval, the derivative might not be a continuous function.

   (a) First, to warm up, show that the function

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0 \end{cases}$$

   is continuous on $[-1, 1]$ but not differentiable at $x = 0$.

   (b) Now consider the function

$$g(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

   Find a formula for the derivative $g'(x)$ for $x \neq 0$.

   (c) Show that $g'(0) = 0$. (*Hint:* Use Exercise 17 in § 4.2) .

   (d) Show that $g'(x)$ diverges as $x \to 0$.

4.6 Related Rates

An important practical application of the Chain Rule, (which could be regarded as an extension of implicit differentiation) is the calculation of **related rates**. We begin with a concrete example.
A rectangle has two sides along the positive coordinate axes, and its upper right corner lies on the curve

$$x^3 - 2xy^2 + y^3 + 1 = 0.$$

How fast is the area of this rectangle changing as the point passes the position $(2, 3)$*, if it is moving so that* $\frac{dx}{dt} = 1$ *unit per second?*
Here, we have both coordinates as (unkown) functions of time

$$x = x(t), \quad y = y(t)$$

related by the equation of the curve:

$$(x(t))^3 - 2\,(x(t))\,(y(t))^2 + (y(t))^3 + 1 = 0.$$

Using the Chain Rule, we can differentiate this relation to obtain a relation between the values of $x = x(t)$, $y = y(t)$ and their rates of change $x' = x'(t)$, $y' = y'(t)$ at any particular moment:

$$3\,(x(t))^2\,x'(t) - 2\{\left(x'(t)\right)\,(y(t))^2 + (x(t))\cdot 2\,(y(t))\,y'(t)\} + 3\,(y(t))^2\,y'(t) = 0.$$

Now, at the moment in question, we have $x(t) = 2$, $y(t) = 3$ and $x'(t) = 1$; substituting this into our relation, we have

$$3(2)^2(1) - 2\{(1)(3)^2 + (2)(2)(3)y'\} + 3(3)^2y' = 0$$

or, solving for y',

$$y'(t) = 2.$$

Now, to find the rate of change of the area of the rectangle (which has sides of length x and y respectively, we differentiate

$$A(t) = x(t)\cdot y(t)$$

using the Product and Chain Rules, then substitute our known values to obtain

$$\frac{dA}{dt} = x'y + xy' = (1)(3) + (2)(2) = 7.$$

So at that moment, the area of the rectangle is increasing at 7 units per second.
Let us clarify a few points concerning this method. Notice that the relationship between the variables which we differentiate (in order to relate their rates of change) must be one that holds for a period of time around our moment of interest: for example, the fact that temporarily (*i.e.*, at the moment of interest) we also have the relationship $3x(t) = 2y(t)$ is completely irrelevant, as this is a transitory phenomenon. We write down all the long-term relationships, then take derivatives with respect to time (or some other "independent" variable), and *only then* do we substitute

the numerical values which we know, before solving our equations for the unknown rates. It is helpful to note that the equations we get will all consist of constants and terms in each of which only a single rate of change appears, multiplied by something that depends only on the values of the variables (not their rates of change); thus the equations will always be easy to solve algebraically (unlike, in general, any attempt to relate the variables directly before differentiating).
This discussion can be summarized in a **Strategy for Related Rates:**

0. Always begin by naming all (possibly) relevant **variables** (quantities of interest); **what do we know?** (*For example, we may know the values of some derivatives.*)

1. **What do we want? When?** (*The answer to the first question will be a derivative of one of the variables with respect to another one; the answer to the second question will be a specification of the* ***moment of interest****.*)

2. Write down all ***long-term*** **relations** among the variables. (*These often come from either the geometry or the physics of the situation.*)

3. **Differentiate** the long-term relations to obtain relations among the rates of change.

4. **Substitute** the values of variables and known rates *at the moment of interest* into these relations. (*We may have to solve some of the long-term relations for some variables in terms of others that we know.*)

5. **Solve** for the desired rate (= answer to "What do we want?")

Let us see how this strategy works in practice, with some other examples.
Example: *A ten foot long ladder leans against a wall, and its base is sliding away from the wall at 4 ft/min. How fast is the tip is moving down the wall when the height of the tip is twice the distance of the base from the wall?*
Let us go through the steps of the strategy:

0. **Variables:** We will denote the distance of the base of the ladder from the wall by x, and the height of the tip along the wall by y. We know from the problem that
$$\frac{dx}{dt} = 4.$$

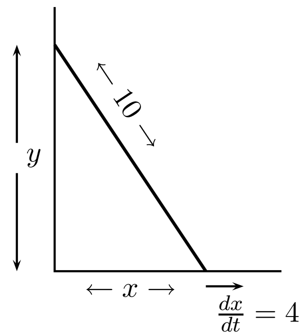


Figure 4.19: The sliding ladder

1. **What do we want? when?** We want to know $\frac{dy}{dt}$ when $y = 2x$.
2. **Long-term relations:** The Pythagorean Theorem tells us that

$$x^2 + y^2 = 10^2 = 100.$$

Note that the relation $y = 2x$ is *not* a *long-term* relation, since it holds at only one moment of time (it specifies the moment of interest). We do not introduce it before performing the next step.

3. **Relation between rates:** Differentiating the equation above, we have

$$2x\frac{dx}{dt} + 2y\frac{dy}{dt} = 0.$$

4. **Known values at moment of interest:** At the moment we are interested in, we have $y = 2x$ and $x^2 + y^2 = 100$; substituting the first into the second we get

$$\begin{aligned} x^2 + (2x)^2 &= 100 \\ 5x^2 &= 100 \\ x^2 &= 20 \\ x &= 2\sqrt{5} \end{aligned}$$

(since x must be positive)

$$\begin{aligned} y &= 2x \\ &= 4\sqrt{5}. \end{aligned}$$

5. **Solve for desired rate:** Substituting this, as well as the known rate $\frac{dx}{dt} = 4$, into the relation between rates, we have

$$\begin{aligned} 2(2\sqrt{5})(4) + 2(4\sqrt{5})\frac{dy}{dt} &= 0 \\ 8\sqrt{5}\frac{dy}{dt} &= -16\sqrt{5} \\ \frac{dy}{dt} &= -2. \end{aligned}$$

The tip of the ladder is moving down the wall at 2 ft/min.
We note in passing that we didn't really need the values of x and y at the moment of interest to solve this problem: once we differentiated the long-term relations, we could have simply substituted the momentary relation $y = 2x$ into the relation between rates to get

$$2x\frac{dx}{dt} + 2(2x)\frac{dy}{dt} = 0$$

so

$$\begin{aligned} \frac{dy}{dt} &= -\frac{1}{2}\frac{dx}{dt} \\ &= -\frac{1}{2}(4) \\ &= -2. \end{aligned}$$

Example: *A point moves along the curve $xy = 1$ with horizontal velocity 1 (see Figure 4.20). How fast is the distance between it and the point $(2, 3)$ changing when $x = 1$? When $x = 2$?*

0. **Variables:** The coordinates of the point are (x, y); let us denote the distance to $(2, 3)$ by ℓ. We know that

$$\frac{dx}{dt} = 1.$$

1. **What do we want? when?** We want to know $\frac{d\ell}{dt}$ at two different times: when $x = 1$ and when $x = 2$.

2. **Long-term relations:** The distance formula is most conveniently written in the indirect form

$$(x-2)^2 + (y-3)^2 = \ell^2,$$

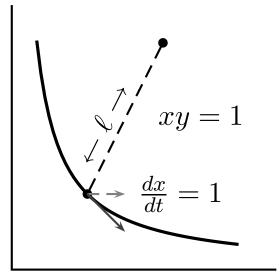


Figure 4.20: The moving point

and the fact that the point lies on the curve

$$xy = 1$$

is, in this case, a *long-term* relation.

3. **Relation between rates:** Differentiating the equations above, we have

$$2(x-2)\frac{dx}{dt} + 2(y-3)\frac{dy}{dt} = 2\ell\frac{d\ell}{dt}$$

or

$$(x-2)\frac{dx}{dt} + (y-3)\frac{dy}{dt} = \ell\frac{d\ell}{dt};$$

also, from the equation of the curve,

$$x\frac{dy}{dt} + \frac{dx}{dt}y = 0.$$

4. **Known values at moment of interest:** At the moments we are interested in, we have $x = 1$ (*resp.* $x = 2$); substituting into the equation of the curve gives us y; when $x = 1$,

$$y = 1$$

while when $x = 2$,

$$y = \frac{1}{2}.$$

Substituting each of these into the equation for the distance to $(2, 3)$, this time in the direct form, gives

$$\ell = \sqrt{(x-2)^2 + (y-3)^2}$$

which when $x = 1$ reads

$$\begin{aligned} \ell &= \sqrt{(-1)^2 + (-2)^2} \\ &= \sqrt{5} \end{aligned}$$

while when $x = 2$

$$\begin{aligned} \ell &= \sqrt{(0)^2 + \left(-\frac{3}{2}\right)^2} \\ &= \frac{3}{2} \end{aligned}$$

5. **Solve for desired rate:** Substituting the respective y-values, as well as the known rate $\frac{dx}{dt} = 1$, into the (second) rate relation, between $\frac{dx}{dt}$ and $\frac{dy}{dt}$, we have, when $x = 1$,

$$(1)\frac{dy}{dt} + (1)(1) = 0$$

or

$$\frac{dy}{dt} = 1$$

while when $x = 2$

$$(2)\frac{dy}{dt} + (1)(\frac{1}{2}) = 0$$

or

$$\frac{dy}{dt} = -\frac{1}{4}.$$

Now, substituting into the (first) rate relation, we have when $x = 1$

$$\begin{aligned} (1-2)(1) + (1-3)(-1) &= \left(\sqrt{5}\right)\frac{d\ell}{dt} \\ \frac{d\ell}{dt} &= \frac{1}{\sqrt{5}}, \end{aligned}$$

while when $x = 2$

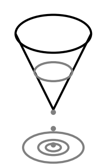
$$\begin{aligned} (0)(1) + \left(-\frac{3}{2}\right)\left(-\frac{1}{4}\right) &= \left(\frac{3}{2}\right)\frac{d\ell}{dt} \\ \frac{3}{8} &= \left(\frac{3}{2}\right)\frac{d\ell}{dt} \\ \frac{1}{4} &= \frac{d\ell}{dt}. \end{aligned}$$

Exercises for § 4.6

Answers to Exercises 1, 3, 6-12 (even only) are given in Appendix B.

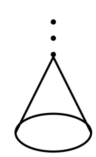
Practice problems:

1. A conical filter whose width at the top equals its depth passes water at the rate of 2 cubic centimeters per minute. How fast is the depth of the water changing when it equals 5 cm?

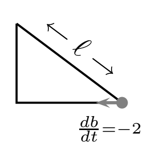
   

   (The volume of a cone with height h and *radius* r is $\pi r^2 h/3$.)

2. Rubble is being dumped at $1000\, ft^3/hr$ into a conical pile whose diameter is always the same as its height ; how fast is the height of the pile increasing when the pile is 10 feet high?

3. A right triangle has constant area 6; if the base is decreasing at 2 units/sec, how fast is the hypotenuse changing

   (a) when the base is 4 and the height is 3? Is the hypotenuse increasing or decreasing?

(b) What about when the *height* is 4 and the *base* is 3?

4. The volume of a spherical balloon is increasing at 8π units per second. How fast is its surface area changing? (The volume and surface area of a sphere of radius r are, respectively $V = 4\pi r^3/3$ and $S = 4\pi r^2$.)

5. The *Ideal Gas Law* says that

$$pV = nRT$$

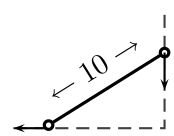
where V is the volume of a gas, T is its (absolute) temperature, p is its pressure, n is the number of "moles"(a measure of the number of molecules) of the gas, and R is a universal constant, known as the *gas constant*.

(a) Use implicit differentiation to show that if volume and pressure are both increasing, so is the temperature.

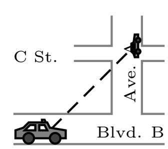
(b) What relation between the rate of change of volume and that of pressure is needed to make sure the gas is cooling?

(c) What relation between the rate of change of volume and that of temperature is needed to make sure the pressure is not increasing?

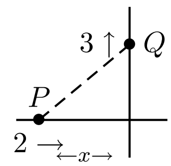
6. Two parts of a machine are connected by a ten-foot strap; the first part moves horizontally while the second moves along a vertical rail. If the first part is 8 feet from the foot of the rail and moving away from the rail at 2 feet per second, how fast is the second part moving?



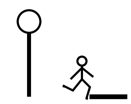
7. A police cruiser is driving along Boulevard B, 0.3 miles from the intersection with Avenue A, at 50 mph toward the avenue, when the officer sees a stolen car on Avenue A crossing C Street, which is 0.4 miles from Boulevard B; she determines that the distance between them is increasing at 26 mph. How fast is the stolen car traveling?



8. A point P on the x-axis is approaching the origin from the left at 2 units/sec, while point Q on the y-axis is moving up at 3 units/sec. If P is x units from the origin and the slope of the line segment PQ is m, how fast is the slope changing? (Give a formula in terms of x and m).



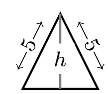
9. (a) A woman 5 feet tall is running away from a 12 foot high lamp post at 7 feet per second. How fast is the length of her shadow changing? Is it getting shorter or longer?



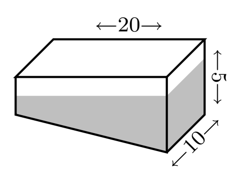
(b) In the situation described above, how fast is the tip of her shadow moving with respect to the lamppost?

(c) Show that if a 6 foot tall man walks away from the same lamppost, then the length of his shadow will increase at the same rate as his speed in walking.

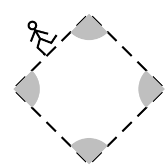
10. Consider an isosceles triangle whose two equal sides have length 5 inches; let h be the length of the line from the vertex between these sides and the midpoint of the third side (this is its height, if it rests on the third side). How fast is the area of the triangle changing if $h = 3$ and increasing at $1 in/sec$?



11. A swimming pool is 20 meters long, 10 meters wide, and has a depth increasing steadily from 1 meter at the shallow end to 5 meters at the deep end. Water is being poured into the pool at the rate of 150 cubic meters per hour. How fast is the depth of the water (measured at the deep end) increasing when it is (a) 4.5 meters (b) 3 meters?



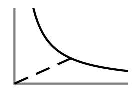
12. A baseball diamond is a square 90 feet on each side. If a player is running from second to third base at 6 feet per second, how fast is his distance from home plate changing when he is (a) one-third of the way there (b) halfway there?



13. If

$$xy = C$$

for some constant C, and m represents the slope of the line from the origin to the point, then the natural logarithm of m *decreases* twice as fast as the natural logarithm of x *increases.*



4.7 Extrema Revisited

We saw in § 3.3 that a continuous function whose domain includes a closed interval $I = [a, b]$ achieves its maximum and minimum values on I. In this section, we shall see how derivatives can help us find these values, which *a priori* involve comparing infinitely many possible inputs.

In § 3.3, we defined the notion of a maximum and minimum for a function f on a specified interval (or on any specific set S of real numbers). Here it will be useful to consider a related but somewhat less specific notion of maximum, minimum, and extremum. By a **neighborhood** of a point x_0 we shall mean any set which contains an open interval containing x_0; here we think particularly about an open interval of the form $(x_0 - \varepsilon, x_0 + \varepsilon)$, where $\varepsilon > 0$ is some small number—this constitutes the set of points "near" x_0. We call a point x_0 an **interior point** of a set $S \subset \mathbb{R}$ if S is a neighborhood of x_0.

Definition 4.7.1. *A function f has a **local maximum** (resp. **local minimum**) at $x = x_0$ if there is some neighborhood $\mathcal{N}$ of x_0 contained in the domain of f and $f(x_0) = \max_{x \in \mathcal{N}} f(x)$ (resp. $f(x_0) = \min_{x \in \mathcal{N}} f(x)$); we refer to x_0 as a **local maximum point** (resp. **local minimum point**) for f. To refer to either kind of point without specifying which it is, we speak of a **local extremum** for f at $x = x_0$, and refer to x_0 as a **local extreme point** of f.*

Note that these local definitions make no reference to the domain of f (except to require that it contain some neighborhood of x_0); to distinguish the maximum (*resp.* minimum) value (if any) of f on its whole domain, we refer to the **global maximum** (*resp.* **global minimum**) value of f.

Clearly, if f achieves its *global* maximum or minimum value at an interior point of its domain, then that point is also a *local* minimum (*resp.* local maximum) point for f. However, if for example the domain of the function is a closed interval, then it is possible that the *global* maximum or minimum occurs at an endpoint of the interval, which would *not* be a *local* extreme point for f. See Figure 4.21.

Our basic technical lemma is the following:

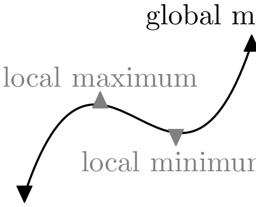


Figure 4.21: Local vs. global extrema

Lemma 4.7.2. *If f is differentiable at $x = x_0$ with nonzero derivative $f'(x_0) = m \neq 0$, then for any sequence of points $x_k < x_0 < x'_k$ with $x_k \to x_0$, $x'_k \to x_0$ we eventually have*

- *If $m > 0$*
$$f(x_k) < f(x_0) < f\left(x'_k\right);$$
- *If $m < 0$*
$$f(x_k) > f(x_0) > f\left(x'_k\right).$$

Proof. Let us suppose that $f'(x_0) = m > 0$. This means that

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = m$$

and in particular that for any sequence $x_k \to x_0$ with $x_k \neq x_0$, for k sufficiently large

$$\frac{f(x_k) - f(x_0)}{x_k - x_0} > \frac{m}{2}.$$

Since $\frac{m}{2} > 0$, this means that $f(x_k) - f(x_0)$ and $x_k - x_0$ have the same sign: if $x_k < x_0$ then $f(x_k) < f(x_0)$, and if $x_0 < x_k$, then $f(x_0) < f(x_k)$. Applying this to sequences converging to x_0 strictly monotonically from each side yields the desired result. A similar argument yields the desired result when $m < 0$. □

Lemma 4.7.2 says that a point where the derivative of f exists and is nonzero cannot be a local extreme point; thus, at any local extreme point, the derivative either fails to exist or equals zero. We call a point where $f'(x_0) = 0$ a **stationary point** of f and one where the derivative fails to exist a **point of non-differentiability** for f; the two kinds of points together are referred to as **critical points** of f. (The value of f at a critical point is called a **critical value** of f.) We have seen the following:

Theorem 4.7.3 (Critical Point Theorem)**.** *Local extrema can only occur at critical points.*

The observation that a point where a differentiable function $f(x)$ achieves an extremum must be a stationary point was first made by Johannes Kepler (1571-1630) in a 1615 work on the geometry of wine bottles [34]. In the late 1620's, Pierre de Fermat (1601-1665) worked out an algebraic method for finding the maximum of a polynomial which amounts, in modern terms, to setting the derivative equal to zero. See Exercise 15 for details. Because of this, Theorem 4.7.3 is sometimes called "Fermat's Theorem".[11]

Simple examples of both phenomena are provided by the two functions $f(x) = x^2$ and $f(x) = |x|$, both of which have a local (and even global) minimum at $x = 0$: in the first case, it is a stationary point, in the second, it is a point of non-differentiability. Note also that the theorem does *not* say that every critical point is a local extremum: for example, the function $f(x) = x^3$ has a stationary point at $x = 0$ but is strictly monotone on $(-\infty, \infty)$ and hence has no local extrema. We will deal with this question more in the next section.

It follows immediately from Theorem 4.7.3 that if f is defined on an interval I, then any point where it achieves a *global* extremum is either a critical point (if it is interior to I) or an endpoint. Since most everyday functions have only finitely many critical points in any closed interval, it is fairly easy to locate the (global) extrema of such a function on a closed interval I: locate the critical points and the endpoints of its domain, then compare the values of f there: the minimum of these is $\min_{x\in I} f(x)$ and their maximum is $\max_{x\in I} f(x)$.

For example, to find the (global) extreme values of $f(x) = x^3 + 6x^2$ on the interval $[-3, 1]$ (see Figure 4.22), we first find the critical points: since the function is differentiable everywhere, this means we want the stationary points: the equation

$$f'(x) = 3x^2 + 12x = 3x(x + 4) = 0$$

has solutions at $x = 0$ and $x = -4$; the second of these is not in our interval, so we ignore it, but at the stationary point $x = 0$ we have

$$f(0) = 0.$$

[11] This is of course different from the result proved recently by Andrew Wiles, which is known as "Fermat's Last Theorem"—but it *is* the same Fermat

The endpoints are also candidates for global extrema, so we evaluate the function there:

$$f(-3) = 27, \quad f(1) = 7.$$

The lowest of the three numbers $\{0, 27, 7\}$ is

$$\min_{x\in[-3,1]} f(x) = 0 = f(0)$$

while the highest is

$$\max_{x\in[-3,1]} f(x) = 27 = f(-3)\,.$$

Note that in this case one extremum occurs at an endpoint and the other at a critical point. However, it is possible to have both extrema at the endpoints, or both at critical points.

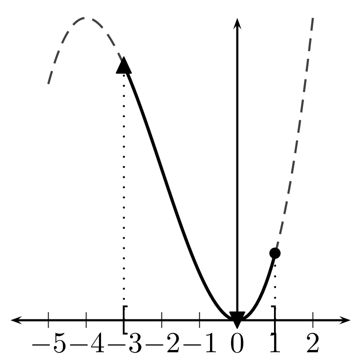


Figure 4.22: $y = x^3 + 6x$

The Extreme Value Theorem (Theorem 3.3.4) only guarantees the existence of global extreme points for a continuous function on an interval that includes both of its endpoints. Nonetheless, when a function has a (one-sided) limit at an endpoint that is *not* in the interval, we can still use this in place of the "value" of the function at the endpoint in our test; however, we need to be careful how we interpret this in the context of finding extreme values on the interval.

If the function approaches a value at an endpoint which is higher (*resp.* lower) than all of the values at critical points as well as that at the other endpoint, then there are points actually in the interval where the function takes values arbitrarily near this extreme value. Assuming that there are only finitely many critical points in the interval, this is the *only* place where we have a chance of achieving a maximum (*resp.* minimum); thus if

the extreme "value" is a limit at an endpoint *not* included in the interval, and this value is not achieved at any internal critical point, then this value *is* the supremum (*resp.* infimum) of the function on the interval, but this value *is not* achieved in the interval, and hence the function has no maximum (*resp.* minimum) there.
For example, on the open interval $(-1, 4)$, the function

$$f(x) = x^3 - 3x^2 + 2$$

has derivative

$$\begin{aligned} f'(x) &= 3x^2 - 6x \\ &= 3x(x-2) \end{aligned}$$

and hence its critical points are $x = 0$ and $x = 2$, with critical values

$$\begin{aligned} f(0) &= 2 \\ f(2) &= -2. \end{aligned}$$

The endpoint "values" are

$$\begin{aligned} \lim_{x \to -1^+} f(x) &= f(-1) \\ &= -2 \\ \lim_{x \to 4^-} f(x) &= f(4) \\ &= 18. \end{aligned}$$

The lowest of these four values is $f(2) = f(-1) = -2$; since this value is taken on at the point $x = 2$ which lies inside the interval, we have

$$\begin{aligned} \min_{x \in (-1,4)} (x^3 - 3x^2 + 2) &= f(2) \\ &= -2 \end{aligned}$$

but since the highest of these values, $f(4) = 18$, is taken on only at $x = 4$ which is an endpoint *not* in the interval, the function has *no maximum* on this interval. However,

$$\begin{aligned} \sup_{x \in (-1,4)} (x^3 - 3x^2 + 2) &= \lim_{x \to 4^-} (x^3 - 3x^2 + 2) \\ &= 18. \end{aligned}$$

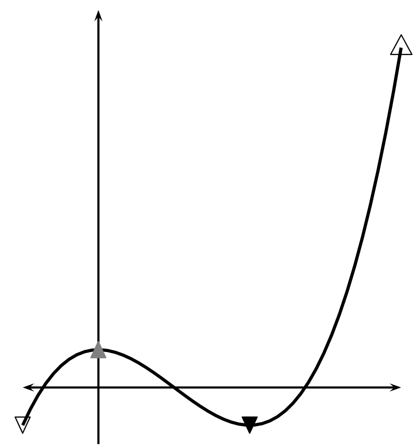


Figure 4.23: $y = x^3 - 3x^2 + 2$

See Figure 4.23

This reasoning works even when the "endpoint " in question is $\pm\infty$ (*i.e.*, the interval is unbounded) provided that there are finitely many critical points[12]. Of course, if the function diverges to ∞ (*resp.* $-\infty$) at an endpoint, then the function is not bounded above (*resp.* below), either. For example, let us find the extreme values of the function

$$f(x) = x + \frac{1}{x}$$

on $(0, \infty)$. We have

$$\begin{aligned} f'(x) &= 1 - \frac{1}{x^2} \\ &= \frac{x^2 - 1}{x^2}. \end{aligned}$$

This vanishes at $x = \pm 1$, but only the value $x = 1$ which lies in the interval concerns us: at $x = 1$, we have the critical value

$$f(1) = 2.$$

[12]or, if there are infinitely many, that all the critical values are at some positive distance from the limit at the "end"

The function diverges to ∞ at both endpoints: near zero, the second term blows up while the first goes to zero, and for large x the first term blows up while the second goes to zero. Thus the only critical value is $f(1) = 2$ and the only endpoint "value" is positive infinity. It follows that the function is *not bounded above:*

$$\sup_{0<x<\infty} \left(x + \frac{1}{x}\right) = \infty$$

while

$$\begin{aligned}\min_{x\in(0,\infty)} \left(x + \frac{1}{x}\right) &= f(1) \\ &= 2.\end{aligned}$$

See Figure 4.24

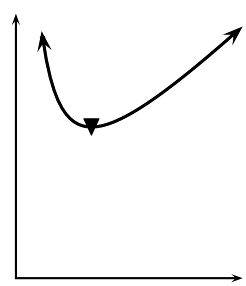


Figure 4.24: $y = x + \frac{1}{x}$

By contrast, the function

$$f(x) = \frac{1}{x^2 + 1}$$

has derivative

$$f'(x) = \frac{-2x}{(x^2 + 1)^2}$$

so the only critical point is $x = 0$, with critical value

$$f(0) = 1.$$

The endpoint "values" are

$$\begin{aligned}\lim_{x\to\infty}\frac{1}{x^2+1} &= 0\\ &= \lim_{x\to-\infty}\frac{1}{x^2+1};\end{aligned}$$

it follows that the higher value gives a maximum

$$\begin{aligned}\max_{x\in\mathbb{R}}\frac{1}{x^2+1} &= f(0)\\ &= 1\end{aligned}$$

while the lower value, which is never achieved but is approximated arbitrarily well, is

$$\inf_{x\in\mathbb{R}}\frac{1}{x^2+1} = 0;$$

the function *has no minimum.* See Figure 4.25

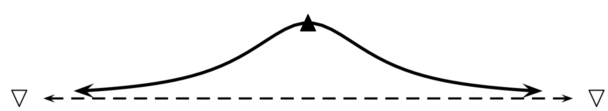


Figure 4.25: $y = \frac{1}{x^2+1}$

We can summarize the preceding observations in the following sophisticated version of a simple critical-point test for extreme values.

Proposition 4.7.4 (Critical Point Test for Global Extrema)**.** *Suppose I is an interval with left endpoint α and right endpoint β, where either or both endpoints may be included in or excluded from I (including the possibility of infinite—excluded—endpoints).*
If f is continuous on I and has only finitely many critical points interior to I, consider the set $\mathcal{CV}(f)$ of critical values from I, together with the endpoint "values" $\lim_{x\to\alpha^+} f(x)$ and $\lim_{x\to\beta^-} f(x)$ (where we assume that these one-sided limits either exist—as finite numbers—or are infinite, i.e., the function diverges to $\pm\infty$ at an endpoint).

Then since $\mathcal{CV}(f)$ is finite, $\max\mathcal{CV}(f)$ *and* $\min\mathcal{CV}(f)$ *both exist, and*

$$\sup_{x\in I} f(x) = \max\mathcal{CV}(f);$$
$$\inf_{x\in I} f(x) = \min\mathcal{CV}(f).$$

If either of these values is associated with a critical point or endpoint which belongs to I, then it gives $\max_{x\in I} f(x)$ *(resp.* $\min_{x\in I} f(x)$*); if either corresponds only to an endpoint* not *included in I, then f has no maximum (resp. minimum) in I.*

We consider a few more examples.
The function

$$f(x) = \frac{x^2 - 4x - 8}{x - 4} + 6\ln(4 - x)$$

is defined only for $x < 4$; let us look for its extreme values on $[-4, 4)$. The derivative is

$$\begin{aligned} f'(x) &= \frac{(2x-4)(x-4) - (x^2 - 4x - 8)}{(x-4)^2} - \frac{6}{4-x} \\ &= \frac{x^2 - 8x + 24}{(x-4)^2} + \frac{6}{x-4} \\ &= \frac{x^2 - 2x}{(x-4)^2} \end{aligned}$$

so the critical points are $x = 0$ and $x = 2$ with respective critical values

$$\begin{aligned} f(0) &= 2 + 6\ln 4 \\ &\approx 10.318 \\ f(2) &= 6 + 6\ln 2 \\ &\approx 10.159 \end{aligned}$$

while the endpoint values are

$$\begin{aligned} f(-4) &= -3 + 6\ln 8 \\ &\approx 9.477 \\ \lim_{x\to 4^-} f(x) &= \infty. \end{aligned}$$

In terms of the notation of Proposition 4.7.4,

$$\mathcal{CV}(f) = \{-3 + 6\ln 8,\ 6 + 6\ln 2,\ 2 + 6\ln 4,\ \infty\}$$

(written in ascending order). Thus, since the lowest value occurs inside the interval,

$$\begin{aligned}\min_{[-4,4)} f &= f(-4)\\ &= -3 + 6\ln 8\end{aligned}$$

while the highest "value" is infinite, so the function is not bounded above; it has no maximum, and

$$\sup_{[-4,4)} f = \infty.$$

As a second example, let us find the largest rectangle (*i.e.*, , of maximal area) which can be inscribed in a circle of radius 1. First, it is clear that the maximum area will involve a rectangle whose vertices all lie on the circle, and second, that by rotating we can assume that the sides of the rectangle are horizontal and vertical. Thus, the corners of the rectangle will be located at the four points $(\pm x, \pm y)$, where $x = \cos\theta$ and $y = \sin\theta$ for some $\theta \in \left[0, \frac{\pi}{2}\right]$.

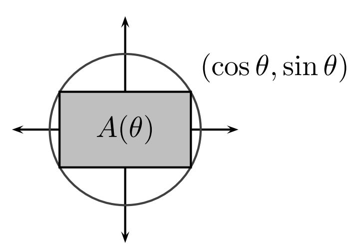


Figure 4.26: Rectangle inscribed in circle

The area of such a rectangle is

$$A(\theta) = (2\cos\theta)(2\sin\theta)$$

which we may recognize as $2\sin 2\theta$. Thus

$$A'(\theta) = 4\cos 2\theta;$$

the only critical point of this function lying in the interval $\left[0, \frac{\pi}{2}\right]$ is $\theta = \frac{\pi}{4}$. We see that

$$\mathcal{CV}(f) = \{0 = A(0) = A(\frac{\pi}{2}), 2 = A(\frac{\pi}{4})\}$$

so

$$\max \mathcal{CV}(f) = A(\frac{\pi}{4}) = 2$$

which is achieved when the sides of the rectangle are equal.
In the preceding problem, it was fairly straightforward to see how to proceed, but for some "word problems" involving maxima or minima, it would be helpful to have a strategy analogous to that for Related Rates in § 4.6. We formulate it as follows:
Strategy for Max-Min Word Problems:

0. Write down all **variables** and **constraints and relations**. (*Relations and constraints will usually come from the geometry or physics of the situation: for example, the relation between sides of a right triangle given by Pythagoras' theorem, or the constraint that lengths cannot be negative.*)

1. **What do we want?** (*The answer to this will have the form "Maximum [or minimum] of ....", where the blank is the name of the* ***variable to be optimized.***)

2. **Solve** the relations for the variable to be optimized as a function of *one* other variable; use the constraints to obtain the domain of this function.

3. Use the critical point method to optimize the variable:

   (a) Find the critical points *in the domain*, and evaluate to get the **critical values** of the function;

   (b) Find the **endpoint values**: if an endpoint belongs to the domain, evaluate the function there; if not, find the limit of the function as the endpoint is approached.

4. The lowest (*resp.* highest) of these values is the minimum (*resp.* maximum) of the desired quantity, *except* that if the extreme value only occurs as a limit at an endpoint, then the maximum (*resp.* minimum) does not exist.

5. **Reinterpret** your answer in terms of the problem as posed.

We illustrate the use of this strategy with two examples.
Example: *Find the point(s) P on the curve $xy^2 = 8$ closest to the point $Q(1, 0)$.*

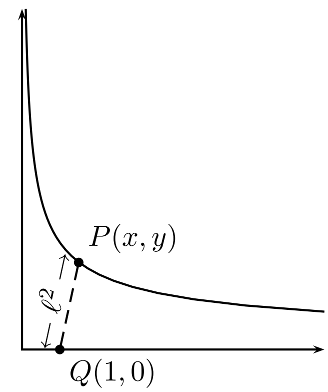


Figure 4.27: $xy^2 = 8$

We go through our strategy step by step:

0. **Variables, constraints, relations:**

   (a) *Variables:* The coordinates of the point(s) P are x and y. We also need to consider the distance PQ; note, however, that rather than minimize the *distance*, which involves a pesky square root, we can minimize the *square of the distance*, since this will be minimized at the same point as the distance itself. Just to be safe, we can give this a name, say ℓ.

   (b) *Constraints:* We note that the coordinates of points on the curve must satisfy

   $$x > 0$$
   $$y \neq 0.$$

   (c) *Relations:* The equation of the curve

   $$xy^2 = 8$$

   relates the two variables x and y, while ℓ is related to both by the distance formula

   $$\ell = (x - 1)^2 + y^2.$$

1. **What do we want?** We want the *minimum* of ℓ.

2. **Reduce to a function:** Substituting the equation of the curve into the distance formula, we see that the problem reduces to finding the minimum of the function

$$\ell(x) = (x-1)^2 + y^2 = (x-1)^2 + \frac{8}{x}$$

on the open interval $(0, \infty)$.

3. **Optimize:**

   (a) **Critical Points:** The derivative of our function is

$$\begin{aligned} \ell'(x) &= 2(x-1) - \frac{8}{x^2} \\ &= \frac{2}{x^2}\left(x^3 - x^2 - 4\right). \end{aligned}$$

   By trial and error, we find that one zero of the polynomial in parentheses is

$$x = 2$$

   and dividing the polynomial by $x - 2$ we obtain the factor $x^2 + x + 2$, which has no real roots, so there are no other critical points. The one critical value is

$$\ell(2) = 5.$$

   (b) **Ends:** The function is undefined at both endpoints, but we see that

$$\lim_{x \to 0^+} \ell(x) = \infty$$

   because the first term of $\ell(x)$ goes to 1 while the second diverges to ∞; similarly,

$$\lim_{x \to \infty} \ell(x) = \infty$$

   because the first term goes to ∞ while the second goes to zero.

4. **Extreme Values:** We see from the preceding that

$$\mathcal{CV}(f) = \{5, \infty\}$$

so the function is not bounded above, but achieves its minimum

$$\begin{aligned}\min_{x>0} \ell(x) &= \ell(2)\\ &= 5.\end{aligned}$$

5. **Interpretation:** We see that the minimum occurs when $x = 2$; there are two points on the curve $xy^2 = 8$ with this x-coordinate, corresponding to $y = \pm 2$. Thus the two points on the curve closest to $(1, 0)$ are $(2, -2)$ and $(2, 2)$.

Example: *James Bond is on one side of a canyon which is 12 miles wide, and needs to reach a secret military facility which is on the other side of the valley, 16 miles down the other side of the canyon. He has a hang-glider with which he can fly, at 3 miles per hour, in a straight line to a point on the opposite rim, after which he can walk the rest of the way along the rim at 5 miles per hour. Where on the opposite rim should he fly to, in order to get to the facility as fast as possible?*

0. **Variables, constraints, relations:** It might help in visualizing this to draw a picture (see Figure 4.28).

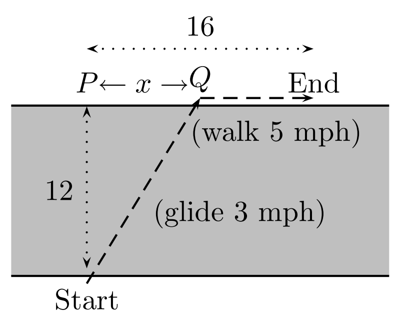


Figure 4.28: Bond's Mission

(a) *Variables:* We denote the distance (in miles) along the opposite rim from the point P directly opposite Bond to the point Q where he will land by x. To compute the time spent gliding, T_g and the time spent walking, T_w, we will need to find the corresponding *distances* D_g and D_w.

(b) *Constraints:* All quantities are non-negative, and it is clear that Bond should land somewhere between the opposite point and the facility, so we have the constraint

$$0 \le x \le 16.$$

(c) *Relations:* The distance gliding is given by Pythagoras' Theorem

$$D_g = \sqrt{x^2 + 12^2}$$

and the distance walking is

$$D_w = 16 - x.$$

The corresponding times are

$$\begin{aligned} T_g &= \frac{D_g}{3} \\ &= \frac{\sqrt{x^2 + 144}}{3} \\ T_w &= \frac{D_w}{5} \\ &= \frac{16 - x}{5}. \end{aligned}$$

1. **What do we want?** We want the *minimum* value of $T = T_g + T_w$.

2. **Reduce to a function:** From the preceding, we see that our problem reduces to finding the minimum over $[0, 16]$ of

$$T(x) = \frac{\sqrt{x^2 + 144}}{3} + \frac{16 - x}{5}.$$

3. **Optimize:**

(a) **Critical Points:** The derivative of our function is

$$\begin{aligned} T'(x) &= \frac{x}{3\sqrt{x^2 + 144}} - \frac{1}{5} \\ &= \frac{5x - 3\sqrt{x^2 + 144}}{5\sqrt{x^2 + 144}}. \end{aligned}$$

The numerator is zero when

$$\begin{aligned} 5x &= 3\sqrt{x^2+144} \\ 25x^2 &= 9\left(x^2+144\right) \\ 16x^2 &= 1296 \\ x^2 &= 81 \end{aligned}$$

so the lone critical point is

$$x = 9$$

and the corresponding critical value is

$$\begin{aligned} T(9) &= \frac{\sqrt{81+144=225}}{3} + \frac{16-9=7}{5} \\ &= \frac{15}{3} + \frac{7}{5} \\ &= 6.4. \end{aligned}$$

(b) **Ends:** At the two ends of our domain, we have the values

$$\begin{aligned} T(0) &= \frac{12}{3} + \frac{16}{5} \\ &= 7.2 \\ T(16) &= \frac{\sqrt{16^2+12^2}}{3} \\ &= \frac{20}{3} \\ &= 6.66... \end{aligned}$$

4. **Extreme Values:** We see from this that the critical value is the lowest, so our function is minimized when $x = 9$.

5. **Interpretation:** Bond should glide to a point nine miles down the rim and then walk the other seven miles.

Exercises for § 4.7

Answers to Exercises 1-2aceg, 3, 5, 7, 8, 12 are given in Appendix B.

Practice problems:

1. Find all critical points of the function f, and classify each as a stationary point or a point of non-differentiability.

   (a) $f(x) = x^3 - 3x^2 + 2x + 1$

   (b) $f(x) = 3x^4 - 20x^3 + 36x^2 + 3$

   (c) $f(x) = (x^2 - 1)^{2/3}$

   (d) $f(x) = \dfrac{1}{x^2 - 1}$

   (e) $f(x) = \dfrac{x}{x^2 + 1}$

   (f) $f(x) = \sin(\pi x^2)$

   (g) $f(x) = \begin{cases} 2 - x & \text{for } x < -1, \\ 4 - x^2 & \text{for } -1 \le x \le 2, \\ x^2 - 8x + 12 & \text{for } x > 2. \end{cases}$

2. For each given function f and interval I, find $\min_{x \in I} f(x)$ and $\max_{x \in I} f(x)$, or explain why it doesn't exist.

   (a) $f(x) = x^3 - 27x$, $I = [-2, 4]$

   (b) $f(x) = x^2$, $I = [-2, 3]$

   (c) $f(x) = x^2$, $I = (-1, 1)$

   (d) $f(x) = (x^2 - 16)^{2/3}$, $I = [-5, 6)$

   (e) $f(x) = \dfrac{x}{x^2 + 4}$, $I = (-3, \infty)$

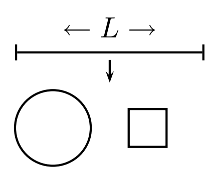
   (f) $f(x) = \dfrac{x^2}{x^2 + 1}$, $I = (-\infty, \infty)$

   (g) $f(x) = e^{-x}$, $I = (-\infty, \infty)$

   (h) $f(x) = \arctan x$, $I = (-\infty, \infty)$

3. A warrior lord has decided to build a rectangular compound on a cliff overlooking the river. He has 1000 feet of fencing, but does not need to fence in the edge of the cliff (which is straight). What dimensions will give him the largest area compound?

4. A wire of length L is to be cut in two, with one piece bent into a circle and the other into a square. How should the wire be cut to maximize the total area of the two figures?



5. The volume of a cylinder of height h and base radius r is $V = \pi r^2 h$, and its surface area is $S = 2\pi r(r + h)$.

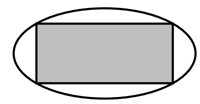
   (a) What are the dimensions of a closed-top cylindrical can with volume 2000π cubic inches that minimizes the amount of tin needed to make it?

(b) What is the maximum volume that can be enclosed in a closed cylindrical can made from 600π square inches of tin?

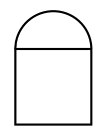
(c) Investigate how these answers change when the can has an open top.

6. Find the point on the curve $y = x^2$ which is closest to the point $(16, \frac{1}{2})$. (*Hint:* Minimize the *square* of the distance.)

7. A rectangle with horizontal and vertical sides is to have all four of its corners on the ellipse $x^2 + 4y^2 = 4$. Find the dimensions of the rectangle of maximal area.



8. A *Norman window* has the shape of a rectangle with a semicircle on top; the diameter of the semicircle exactly matches the width of the rectangle.



(a) Show that the maximal area for a Norman window whose perimeter is fixed at P is achieved when the height of the rectangle is half its width.

(b) How does the problem change if we fix the total amount of frame required, assuming that in addition to the outside frame we need a piece running along the top of the rectangle?

Theory problems:

9. Show for a triangle whose base has length 1 and whose perimeter is 4, the maximal area is $\frac{\sqrt{2}}{2}$. (*Hint:* If the other two sides are s_1 and s_2, we know that $s_1 + s_2$ is constant, and that the area of the triangle equals half the height—which in turn is determined by the shorter of the two sides. Alternatively, place the triangle so that the base is the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$ on the x-axis, and let (x, y) be the location of the vertex opposite the base; then the condition that the perimeter is 4 defines an ellipse of the form $ax^2 + by^2 = 1$ for appropriate constants a and b; use this to write the area of the triangle as a function of x.)

10. Give an example of a function f defined on $(-\infty, \infty)$ which has a local maximum at $x = a$ and a local minimum at $x = b$, but $f(b) > f(a)$.

11. Suppose that a curve $\mathcal{C}$ is given as the graph of a differentiable function f, and that $P(a, b)$ is a point *not* on $\mathcal{C}$. Show that, if Q is the point on $\mathcal{C}$ closest to P, then the line PQ is perpendicular to the tangent line to $\mathcal{C}$ at Q.

Challenge problems:

12. (M. Spivak) A hallway of width w_1 turns a (right-angled) corner and changes width, to w_2. How long a ladder can be taken around the corner? (See Figure 4.29)

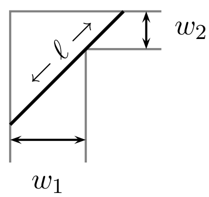


Figure 4.29: Ladder turning a corner

13. Show that among triangles inscribed in the circle $x^2 + y^2 = 1$, an equilateral one has maximum area:

    (a) After a rotation, we can assume that one vertex of the triangle is at $A(0, 1)$.

    (b) Show that we can assume the side opposite A is horizontal, as follows:

    Given $-1 < y_0 < 1$, consider all the triangles inscribed in the circle $x^2 + y^2 = 1$ with one vertex at $A(0, 1)$ for which the side opposite this vertex goes through the point $P(0, y_0)$ (see Figure 4.30).

    i. Verify that the line with slope m through $P(0, y_0)$ hits the circle at the points $B_m(x_1, y_0 + mx_1)$ and $C_m(x_2, y_0 + mx_2)$, where $x_1 < x_2$ are the solutions of the equation

$$x^2 + (y_0 + mx)^2 = 1.$$

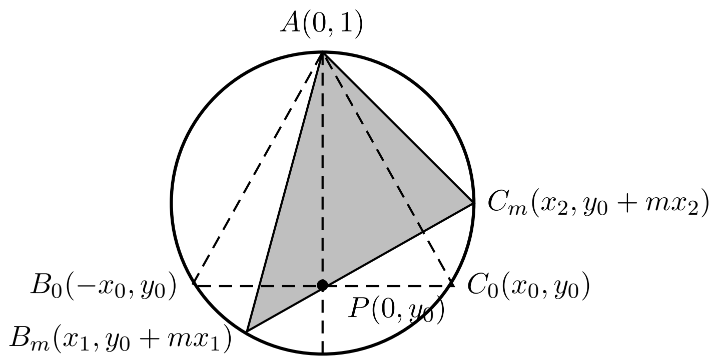


Figure 4.30: Exercise 13

ii. Show that the area of triangle $\triangle AB_mC_m$ is

$$\begin{aligned}\mathcal{A}(m) &= \frac{1}{2}\{(x_2 - x_1)(1 - y_0)\} \\ &= \frac{(1 - y_0)\sqrt{1 + m^2 - y_0^2}}{1 + m^2}.\end{aligned}$$

(*Hint:* Consider the two triangles $\triangle AB_mP$ and $\triangle APC_m$, noting that they share the common "base" AP; what are the corresponding "heights"?)

iii. Show that this quantity (for fixed $y_0 \in (0, 1)$) takes its maximum when $m = 0$.

(c) Now consider all triangles with vertex A and base the horizontal chord at height $-1 \le y \le 1$. Show that the area of such a triangle is given by

$$\mathcal{A}(y) = (1 - y)\sqrt{1 - y^2}.$$

Find the value of y which maximizes this area, and show that the resulting triangle is equilateral.

14. *Fermat's Principle*[13] in optics says that a light beam from point A to point B takes the path of least time. We investigate two consequences.

[13] Stated by Fermat in a letter in 1662, but apparently already known to Heron of Alexandria (ca. 75 AD)

(a) The *Law of Reflection* says that a light beam reflected (*e.g.*, off a mirror) leaves the mirror at the same angle as it hit: *the angle of incidence equals the angle of reflection.* Show that this follows from Fermat's Principle: (see Figure 4.31)

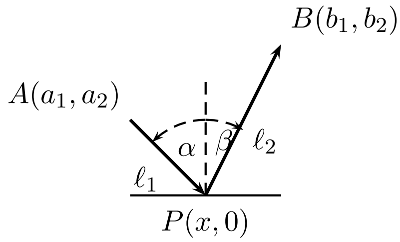


Figure 4.31: Law of Reflection

Show that among paths APB, where $A(a_1, a_2)$ and $B(b_1, b_2)$ are points above the x-axis and $P(x, 0)$ is on the axis, the shortest path (*i.e.*, minimizing $\ell_1 + \ell_2$) has

$$\alpha = \beta.$$

(b) *Snell's Law* (or the *Law of Refraction*) says that if a light ray passes from a medium in which the speed of light is v_1 to one in which it is $v_2 > v_1$, the ratio of the sines of the angles the incoming and outgoing beam make with the normal to the interface equals v_1/v_2. Show that this follows from Fermat's Principle: (see Figure 4.32)

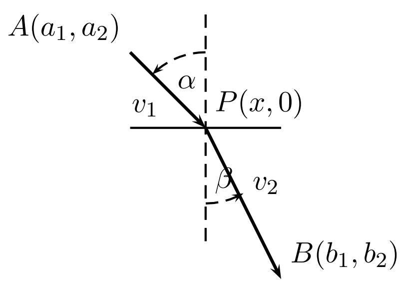


Figure 4.32: Law of Refraction

Show that among paths APB, where $A(a_1, a_2)$ and $B(b_1, b_2)$ are points on opposite sides of the x-axis and $P(x, 0)$ is on the axis, if travel along AP is at speed v_1 while travel along PB is at speed v_2, then the fastest route from A to B has

$$\frac{\sin\alpha}{\sin\beta} = \frac{v_1}{v_2}.$$

History note:

15. **Fermat on Maxima and Minima:** [33, pp. 470-3], [20, pp. 122-125], [49, pp. 610-2] Fermat, building on the work of Viète concerning the relation between the coefficients of a polynomial and its roots, worked out a method for finding the maximum of a polynomial by algebraic means. He illustrated it, in a 1638 letter to Descartes [51, pp. 222-5], via the problem of dividing a line segment AC at a point E so as to maximize the area of the rectangle with sides AE and EC:

    (a) Let b denote the length of AC and a the length of one of the two segments AE and EC; then the length of the other segment is $b - a$, so we are looking to maximize the quantity $ba - a^2$.

    (b) Consider the corresponding situation when a is replaced by $a + e$; then we have in place of $ba - a^2$ the quantity $ba - a^2 + be - 2ae - e^2$. Fermat now applies to these two quantities a process he calls **adequating** (a term he takes from Diophantus): he first eliminates common terms from the two, getting

    $$be \sim 2ae + e^2;$$

    he then divides by e, getting

    $$b \sim 2a + e$$

    and then throws away all terms involving e; this leaves him the equation to be solved

    $$b = 2a;$$

    that is, the maximum occurs when the segment is divided into equal parts.

    (c) Relate this procedure to Theorem 4.7.3. (Note that Fermat did *not* have our notion of a derivative. He did, however, recognize that a similar procedure could be used to determine the line tangent to a curve at a point.)

4.8 Geometric Application of Limits and Derivatives

Differentiation of a function, together with data on limits and a few analytic observations, can give us very good qualitative information about the shape of the curve given as a graph of a function. In this section we pull together these techniques and apply them to sketching the graphs of several rational functions.

Intercepts. One of the most obvious features of a curve in the plane is the places where it crosses or touches the coordinate axes. A curve has an ***x*-intercept** at $x = x_0$ if the point $(x_0, 0)$ lies on the curve, and a ***y*-intercept** at $y = y_0$ if the point $(0, y_0)$ lies on the curve. When the curve is defined by a function via $y = f(x)$, these are easy to characterize in terms of the function:

Remark 4.8.1 (Intercepts)**.** *The graph $y = f(x)$ of a function has*

- *a **y-intercept** at $y = f(0)$, provided $f(0)$ is defined;*
- *an **x-intercept** at each solution of the equation $f(x) = 0$.*

Note that the graph of a function has at most one y-intercept.
For example, the graph of

$$f(x) = x^3 - 3x$$

has its y-intercept at $y = 0$ (*i.e.*, at the origin) and x-intercepts at $x = -\sqrt{3}$, $x = 0$ (the origin again) and $x = \sqrt{3}$; the function

$$f(x) = \frac{1}{x^2 - 1}$$

has a y-intercept at $y = -1$, but no x-intercepts (because the numerator is never zero).

Asymptotes. Two curves are *asymptotic* if they approach each other "at infinity"; in particular, a line asymptotic to a curve is an **asymptote** to it. We shall concentrate on horizontal and vertical asymptotes.

Remark 4.8.2 (Asymptotes)**.** *The graph $y = f(x)$ is*

- *asymptotic to the horizontal line $y = y_0$ if either* $\lim_{x\to\infty} f(x) = y_0$ *or* $\lim_{x\to-\infty} f(x) = y_0$ *(we say the curve has a* ***horizontal asymptote*** *at $y = y_0$)*
- *asymptotic to the vertical line $x = x_0$ if $f(x)$ diverges to $\pm\infty$ as $x \to x_0$ from either the right or the left (we say the curve has a* ***vertical asymptote*** *at $x = x_0$).*

For example, graph of the function

$$f(x) = \frac{1}{x^2+1}$$

has a horizontal asymptote at $y = 0$, because

$$\lim_{x\to-\infty} \frac{1}{x^2+1} = 0 \text{ and } \lim_{x\to\infty} \frac{1}{x^2+1} = 0.$$

Furthermore, the graph has its y-intercept at $yf(0) = 1$ and no x-intercepts; from this information we can already get a pretty good idea that the graph should look like Figure 4.33.

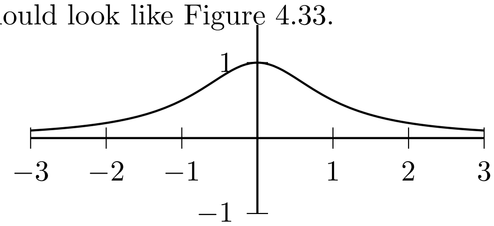


Figure 4.33: $y = \frac{1}{x^2+1}$

As another example, graph of the function

$$f(x) = \frac{1}{x^2-1}$$

has a y-intercept at $y = -1$ and no x-intercepts, is asymptotic to the x-axis ($y = 0$), since $\lim_{x\to\pm\infty} \frac{1}{x^2-1} = 0$, and vertical asymptotes at $x = \pm 1$, because

$$\lim_{x\to-1^-} \frac{1}{x^2-1} = +\infty \quad \lim_{x\to-1^+} \frac{1}{x^2-1} = -\infty$$
$$\lim_{x\to1^-} \frac{1}{x^2-1} = -\infty \quad \lim_{x\to1^+} \frac{1}{x^2-1} = +\infty.$$

The specific information about where the function diverges to $+\infty$ and where it diverges to $-\infty$ helps us sketch the curve. For example, since it diverges to $+\infty$ as x approaches $x = 1$ from the right and has no x-intercepts to the right of $x = 1$, it must stay above the x-axis as it approaches it with $x \to +\infty$. The sketch given in Figure 4.34 has some features which seem reasonable, but will be justified later in this section.

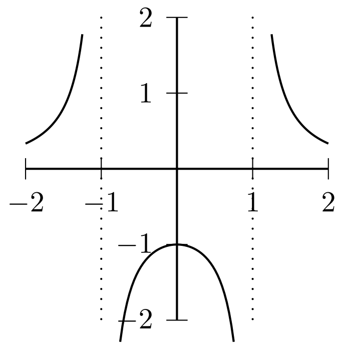


Figure 4.34: $y = \frac{1}{x^2-1}$

Monotonicity. Recall from Definition 3.2.2 that a function f is **strictly increasing** ($f\uparrow$) (*resp.* **strictly decreasing** ($f\downarrow$)) on an interval if for any two points $x_1 < x_2$ belonging to this interval we have $f(x_1) < f(x_2)$. We saw in Lemma 4.7.2 that if the derivative is nonzero at a point x_0 then we have some points on either side of x_0 satisfying similar inequalities with respect to x_0; the following extension of that argument gives us information about monotonicity of f on any interval where the function is differentiable and the derivative does not switch sign.

Proposition 4.8.3 (Monotonicity)**.** *Suppose f is differentiable*[14] *on the nontrivial*[15] *interval I.*
Then

- *if $f'(x) > 0$ at every interior point of I, then f is strictly increasing on I;*
- *if $f'(x) < 0$ at every interior point of I, then f is strictly decreasing on I.*

[14] Actually, at the endpoints we only need continuity: see Exercise 7, § 4.9.

[15] We are merely avoiding the silly exception that I is a single point, which is technically a closed interval.

Proof. We will prove the first case; the second is easy to deduce from the first.
We will show that if x_0 is an interior point of I, then for any $x > x_0$ in I, $f(x) > f(x_0)$. First, we will show that there exists $\varepsilon > 0$ so that this is true whenever $x_0 < x < x_0 + \varepsilon$: if not, we have a sequence of points x_k converging to x_0 from the right (which we can assume is strictly decreasing) with $f(x_k) \leq f(x_0)$ for $k = 1, 2, ...$ But then the secants joining x_0 to x_k all have nonpositive slope, and since these slopes converge to $f'(x_0)$, we have $f'(x_0) \leq 0$, contradicting our assumption[16]
Now, suppose that there is *some* point $x \in I$ with $x > x_0$ but $f(x) \leq f(x_0)$; then consider the infimum of the nonempty bounded set consisting of *all* such points in I for which $x > x_0$ but $f(x) \leq f(x_0)$:

$$x' := \inf\{x \in I \,|\, x_0 < x \text{ but } f(x_0) \geq f(x)\}.$$

By continuity of f, $f(x') \leq f(x_0)$, which implies that it cannot lie between x_0 and $x_0 + \varepsilon$. Now, consider the function f on the closed interval $[x_0, x']$: it is continuous on this interval, and hence achieves its maximum, and the maximum value is greater than either endpoint value (since there are points where $f(x) > f(x_0)$), so f has a relative maximum strictly between x_0 and x'. But since f is differentiable there, it must be a stationary point, contradicting the assumption that $f'(x) > 0$ everywhere in I.
Thus, we have shown that *every* point in I to the right of x_0 satisfies $f(x) > f(x_0)$. A similar argument (still assuming $f(x) > 0$ for all $x \in I$) shows that every point to the *left* of x_0 in I satisfies $f(x) < f(x_0)$. Finally, we note the whole case was made for an *arbitrary* interior point of I, so we have shown that if $x_1 < x_2$ are not both endpoints of I, then $f(x_1) < f(x_2)$. But if $I = [x_1, x_2]$ then we can apply the argument using an interior point x_3 to argue that $f(x_1) < f(x_3) < f(x_2)$, completing the argument, at least when $f'(x) > 0$ for every interior point of I. □

Note that this result does *not* require that the derivative be nonzero at the endpoints of I. Of course, the values of f at the endpoints are being compared only to values achieved inside I, in particular at points on only one side of the endpoint.
As an example, let us look again at the polynomial

$$f(x) = x^3 - 3x.$$

Its derivative is

$$f'(x) = 3x^2 - 3 = 3(x - 1)(x + 1).$$

[16] equivalently, contradicting Lemma 4.7.2.

We note that the derivative vanishes (*i.e.*, equals zero) at $x = \pm 1$, and thus has constant sign on each of the three intervals $(-\infty, -1)$, $(-1, 1)$, and $(1, \infty)$. In the first interval, both factors are negative, so $f'(x) > 0$ on $(-\infty, 1)$ and f is strictly increasing on the interval $(-\infty, 1]$; in the second interval, one factor is positive, the other negative, so $f'(x) < 0$ for $x \in (-1, 1)$. It follows that f is strictly decreasing on $[-1, 1]$; finally, in the third interval, both factors are positive, hence so is $f'(x)$ and f is strictly increasing on $[1, \infty)$. This gives us a rough picture of the graph of this function (Figure 4.35), which we will refine further with other information. We note from this figure that the function has a local maximum at $x = -1$

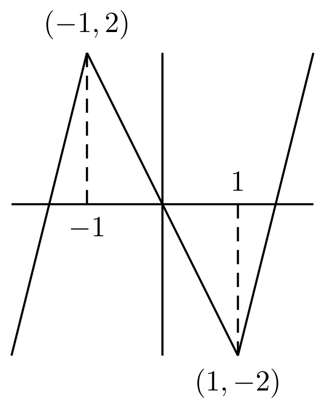


Figure 4.35: Intervals of monotonicity for $y = x^3 - 3x$

and a local minimum at $x = 1$. This happens at any point where $f'(x)$ *switches sign.*

Remark 4.8.4 (First Derivative Test for Local Extrema)**.** *Suppose f is differentiable in a neighborhood of $x = x_0$, except possibly at $x = x_0$, where it is at least continuous.*
Then

- *if $f'(x) > 0$ for $x < x_0$ and $f'(x) < 0$ for $x > x_0$ in the neighborhood, then f has a local maximum at $x = x_0$;*
- *if $f'(x) < 0$ for $x < x_0$ and $f'(x) > 0$ for $x > x_0$ in the neighborhood, then f has a local minimum at $x = x_0$;*
- *if $f'(x)$ has the same sign on both sides of $x = x_0$, then x_0 is not a local extreme point for f.*

As another example, consider

$$f(x) = \frac{x^2 - 1}{x}.$$

This has x-intercepts at $x = \pm 1$ but no y-intercept, and in fact has a vertical asymptote at $x = 0$:

$$\lim_{x \to 0^-} \frac{x^2 - 1}{x} = \infty \text{ and } \lim_{x \to 0^+} \frac{x^2 - 1}{x} = -\infty.$$

We note also that

$$\lim_{x \to -\infty} \frac{x^2 - 1}{x} = -\infty \text{ and } \lim_{x \to \infty} \frac{x^2 - 1}{x} = +\infty.$$

Finally, the derivative

$$f'(x) = \frac{x^2 + 1}{x^2}$$

is always positive. However, this does not imply that the function is strictly increasing on the whole real line: the fact that f (and f') are undefined at zero means we can only use Proposition 4.8.3 on the intervals $(-\infty, 0)$ and $(0, \infty)$; in fact, it is clear from Figure 4.36 that there are points $x_1 < 0 < x_2$ with $f(x_1) > f(x_2)$.

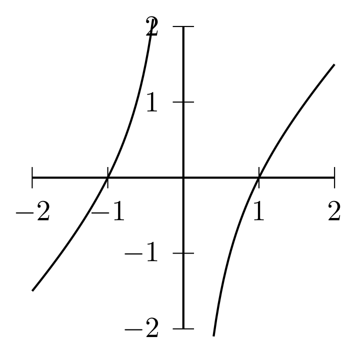


Figure 4.36: $y = \frac{x^2-1}{x}$

Concavity/Convexity. A curve is called **convex** if every one of its secant segments lies on one side of the curve. Of course, in terms of normal usage, a curve which is *convex* (*i.e.*, bulging out) from one side, is *concave* from the other. For the graph of a function f, we characterize a piece of its graph which is convex in terms of the direction from which it is concave.

Definition 4.8.5. *A differentiable function f is*

- ***concave up** on an interval I if its derivative f' is strictly increasing on I*
- ***concave down** [17] on an interval I if its derivative f' is strictly decreasing on I.*

In colloquial terms, the graph of a function which is concave up is "smiling", while the graph of a function which is concave down is "frowning". Each type of concavity can occur for either a strictly increasing or a strictly decreasing function (Figure 4.37).

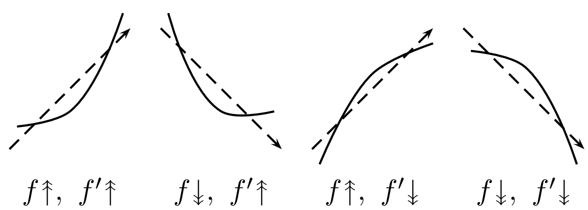


Figure 4.37: Concavity and Monotonicity

Recall that the **second derivative** of a function f is the derivative of its derivative

$$f''(x) := \frac{d}{dx}\left[f'(x)\right] = \left(f'\right)'(x).$$

A function is **twice differentiable** if it is differentiable and f' is also differentiable. Applying Proposition 4.8.3 to the derivative of a twice differentiable function, we have

Remark 4.8.6. *If $f''(x) > 0$ (resp. $f''(x) < 0$) on an interval I, then f is concave up (resp. concave down) on I.*

A point on the graph where *concavity switches* is called an **inflection point** of the graph. Since this means it is a local extreme point for f', it follows that at an inflection point of the graph, the *second* derivative must vanish; however, just as for local extrema of f, not every point with $f''(x) = 0$ is an inflection point. For example, it is easy to check that $f(x) = x^4$ has $f''(0) = 0$, but $f'(x) \geq 0$ everywhere. Near an inflection point, the graph is "S"-shaped; it curves one way, then the other.

[17] Sometimes a function which is concave down is called a **convex function**; we will stick to the clearer language of concavity

For example, the function

$$f(x) = x^3 - 3x$$

with

$$f'(x) = 3x^2 - 3$$
$$f''(x) = 6x$$

is concave down on the interval $(-\infty, 0]$ and concave up on the interval $[0, \infty)$, with an inflection point at the origin. This information allows us to "smooth out" the rough plot shown in Figure 4.35, to get Figure 4.38.

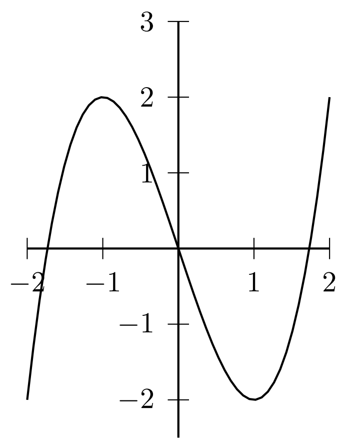


Figure 4.38: $y = x^3 - 3x$

The function

$$f(x) = \frac{1}{x^2 - 1}$$

with

$$f'(x) = \frac{-2x}{(x^2-1)^2}$$
$$f''(x) = \frac{6x^2+2}{(x^2-1)^3}$$

is concave up on each of the intervals $(-\infty, -1)$ and $(1, \infty)$ and concave down on the interval $(-1, 1)$; however, it has no inflection point: the regions of concavity are separated by vertical asymptotes.

Similarly, the function

$$f(x) = \frac{x^2 - 1}{x}$$

which has

$$f'(x) = \frac{x^2 + 1}{x^2}$$
$$f''(x) = \frac{-2}{x^3}$$

is concave up on $(-\infty, 0)$ and concave down on $(0, \infty)$, with these regions separated by a vertical asymptote.
We note also the following observation, which follows easily from Remark 4.8.4 and Remark 4.8.6:

Remark 4.8.7 (Second Derivative Test for Stationary Points)**.** *If a function f is twice differentiable in a neighborhood of a stationary point $x = x_0$ (so $f'(x_0) = 0$) then*

- *if $f''(x_0) < 0$ then f has a local* ***maximum*** *at $x = x_0$*
- *if $f''(x_0) > 0$ then f has a local* ***minimum*** *at $x = x_0$*

Note that if $f''(x_0) = 0$, we need more information before drawing any conclusions: the point may be a relative maximum a relative minimum, or neither.
Let us conclude with another example. The function

$$f(x) = \frac{x}{x^2 + 1}$$

has an x-intercept and y-intercept at the origin, no vertical asymptotes, and is asymptotic to the x-axis in both the positive and negative direction, since

$$\lim_{x\to\infty} \frac{x}{x^2 + 1} = \lim_{x\to-\infty} \frac{x}{x^2 + 1} = 0.$$

Also,

$$f'(x) = \frac{1 - x^2}{(x^2 + 1)^2} = \frac{(1 - x)(x + 1)}{(x^2 + 1)^2}$$
$$f''(x) = \frac{2x^3 - 6x}{(x^2 + 1)^3} = \frac{2x(x - \sqrt{3})(x + \sqrt{3})}{(x^2 + 1)^3}$$

so the function is strictly decreasing on $(-\infty, -1]$ and $[1, \infty)$, and strictly increasing on $[-1, 1]$; by either Remark 4.8.4 or Remark 4.8.7, it has a local minimum at $x = -1$ (where $f(-1) = -\frac{1}{2}$) and a local maximum at $x = 1$ (where $f(1) = \frac{1}{2}$). It is concave down on the intervals $\left[-\infty, -\sqrt{3}\right)$ and $\left[0, \sqrt{3}\right]$ and concave down on the intervals $\left[-\sqrt{3}, 0\right]$ and $\left(\sqrt{3}, \infty\right]$. A plot of the graph, with these features identified, is given in Figure 4.39

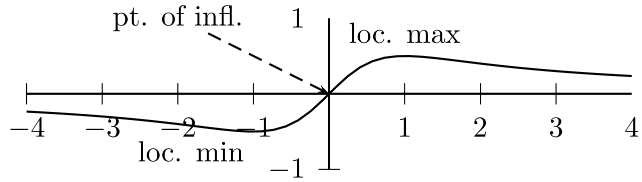


Figure 4.39: $f(x) = \frac{x}{x^2+1}$

Exercises for § 4.8

Answers to Exercises 1acegikm are given in Appendix B.

Practice problems:

1. For each function below, locate all (i) x- and y-intercepts; (ii) horizontal and vertical asymptotes; (iii) (maximal) intervals on which the function is strictly monotone (determine where it is strictly increasing and where it is strictly decreasing); (iv) (maximal) intervals of concavity (determine where it is concave up and where it is concave down); (v) local maxima and minima; (vi) points of inflection; and then *sketch the graph of the function, labelling all the features you have found above*:

   (a) $f(x) = x^2 - 4x$ (b) $f(x) = x^3 - 6x + 16$

   (c) $f(x) = x^4 - 8x^3 + 22x^2 - 24x+8$

   (d) $f(x) = x^4 - x^2$

   (e) $f(x) = \dfrac{x^{4/3}}{2} - 2x^{1/3} = (x-4)\dfrac{\sqrt[3]{x}}{2}$ (f) $f(x) = x^{5/3} + 5x^{-1/3} = x^{-1/3}(x^2+5)$

   (g) $f(x) = x - \dfrac{4}{x} = \dfrac{x^2-4}{x}$ (h) $f(x) = x + \dfrac{4}{x} = \dfrac{x^2+4}{x}$

   (i) $f(x) = \dfrac{x^2+x-2}{x^2-x-2} = 1 + \dfrac{2x}{(x+1)(x-2)}$

(j) $f(x) = \dfrac{x^2}{x^2 + x - 2} = 1 - \dfrac{x - 2}{(x - 1)(x + 2)}$

(k) $f(x) = e^{-x^2/2}$ (l) $f(x) = \arctan x$

(m) $f(x) = \begin{cases} x + 2 & \text{for } x < -1, \\ x^2 & \text{for } x \geq -1. \end{cases}$

2. Sketch the graph of a function f satisfying *all* of the conditions below, or explain why no such function exists:

   (a) f has local minima at $x = -1$, $x = 1$, and $x = 3$;

   (b) f has local maxima at $x = -3$, $x = 0$, $x = 2$, and $x = 5$;

   (c) $\max f(x) = f(-3) = 2$;

   (d) $\min f(x) = f(1) = -5$;

   (e) f is differentiable *except* at $x = 5$.

Theory problems:

3. For each phenomenon, either give an example or explain why none exists.

   (a) A point of inflection which is also a local extremum.

   (b) A function which is strictly increasing and concave down on $(-\infty, \infty)$.

   (c) A bounded function which is defined and concave up on $(-\infty, \infty)$.

4. (a) Show that if the graph $y = f(x)$ of a rational function has a horizontal asymptote at $y = L$, then

$$\lim_{x \to \infty} f(x) = \lim_{x \to -\infty} f(x) = L.$$

   In particular, the graph of a rational function has *at most one* horizontal asymptote. (*Hint:* See Exercise 9 in § 3.4.)

   (b) Find an example of a (non-rational) function whose graph has two *distinct* horizontal asymptotes.

4.9 Mean Value Theorems

In this section, we consider the Mean Value Theorem and some of its variants. The material in this section is relatively technical, and its use is theoretical rather than directly practical. Its consequences are particularly important in demonstrating L'Hôpital's rule, which we treat in the next section, and in the theory of formal integration which we treat in Chapter 5.

Several results in this section share the same hypothesis. Let us say that a function f is **differentiable interior to, and continuous at the ends of** the closed interval $[a,b]$ (for which we use the abbreviation "DICE on $[a,b]$") if it is continuous on $[a,b]$ and differentiable at every point of (a,b): since differentiablility implies continuity, what this really means is that f' exists at every point of $[a,b]$ with the possible exception of the endpoints, and that the one-sided limits (from within the interval) at the endpoints match the function values there:

$$\lim_{x\to a^+} f(x) = f(a)$$
$$\lim_{x\to b^-} f(x) = f(b)\,.$$

When stating this condition, we will also assume that the interval is **non-trivial**; that is, $a \neq b$. These conditions will be present in all of the Mean Value Theorems in this section.

Rolle's Theorem

Our first result is a special case of the Mean Value Theorem, known as *Rolle's Theorem*. Michel Rolle (1652-1719) included a statement of this theorem as an aside, in the middle of a discussion of approximate solution of equations, in an obscure book he published in 1691. It is not clear who named this result "Rolle's theorem", but in an amusing description preceding a translated selection from this [49, pp. 253-60], David Smith suggests that no one could find an appropriate reference to back up this attribution until someone discovered it in 1910 in a Paris library.

Proposition 4.9.1 (Rolle's Theorem)**.** *Suppose f is DICE on the nontrivial closed interval $[a,b]$, and in addition that the endpoint values are the same:*

$$f(a) = f(b)\,.$$

Then there exists at least one point $c \in (a,b)$ *where*

$$f'(c) = 0.$$

(See Figure 4.40.)

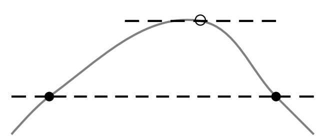


Figure 4.40: Proposition 4.9.1 (Rolle's Theorem)

Proof. If f is constant on $[a,b]$, there is nothing to prove, since then $f(c) = 0$ for *every* $c \in (a,b)$. So assume there is at least one point in (a,b) where $f(x) \neq f(a)$. We assume without loss of generality that $f(x) > f(a)$ somewhere in (a,b).
Now, since f is continuous on the closed interval $[a,b]$, it achieves its maximum there, by Theorem 3.3.4, and since the maximum value is strictly greater than the endpoint values, it is achieved at an interior point, $c \in (a,b)$. Since f is differentiable at c, Lemma 4.7.2 tells us that $f'(c) = 0$, as required. □

The content of this can be stated in the following terms: if a car goes out along a straight and narrow road and, staying on that road, returns to its original position, it had to turn around somewhere; at that moment, its motion in the direction of the road was equivalent to stopping. This is a bit like the Intermediate Value Theorem (Theorem 3.2.1), but for the velocity: at first, the car is moving east (velocity positive), and at the end, it is moving west (velocity negative), so at some time it had to be doing neither (velocity zero). However, this form of the argument would require that velocity is a *continuous* function of time—a reasonable assumption for cars, but in general, we would need to require that the *derivative function* f' be continuous on the *closed* interval $[a,b]$. Here, we use much weaker assumptions.

Mean Value Theorems

A more general result says that if the car travels sixty miles along the road in a period of exactly one hour, then at some moment it had to be going *precisely* sixty miles per hour. This is the content of the *Mean Value Theorem*: that the *instantaneous* rate of change of a function should match the *average* (or "mean") rate of change at least once. If the position at time t is given by the function $y = f(t)$, then the *average* velocity over the time interval $t \in [a, b]$ is the displacement $\triangle y = f(b) - f(a)$ divided by the time elapsed, $\triangle t = b - a$, or the slope of the secant line joining the points $(a, f(a))$ and $(b, f(b))$ on the graph $y = f(t)$:

$$v_{avg} = \frac{f(b) - f(a)}{b - a}.$$

The Mean Value Theorem says that if f is differentiable interior to and continuous at the ends of $[a, b]$, then there is some time $t \in (a, b)$ at which the *instantaneous* velocity is exactly equal to the *average* velocity over this period.

Proposition 4.9.2 (Mean Value Theorem)**.** *Suppose that f is DICE on $[a, b]$.*
Then there is at least one interior point of the interval where the tangent line is parallel to the secant joining the endpoints; that is, for at least one $c \in (a, b)$,

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{4.38}$$

(See Figure 4.41.)

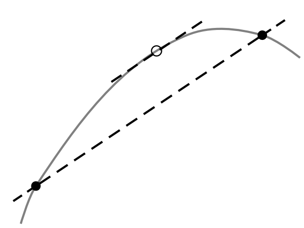


Figure 4.41: Proposition 4.9.2 (Mean Value Theorem)

According to [30, p. 240], this appeared in Lagrange's *Théorie des Fonctions Analytiques* (1797). The version of this given by Cauchy in his original treatment is contained in Exercise 14. The proof here, now standard, was apparently first discovered by Ossian Bonnet (1819-1892) [20, p. 314].

Proof. Let

$$m := \frac{f(b) - f(a)}{b - a}$$

be the slope of the secant line joining the points $(a, f(a))$ and $b, f(b))$; we are looking for a point $c \in (a, b)$ where $f'(c) = m$. Note that the secant line is the graph of the affine function

$$\alpha(x) := f(a) + m(x - a);$$

now take the vertical displacement between the secant and the graph of f:

$$h(x) = f(x) - \alpha(x) = f(x) - f(a) - m(x - a).$$

Since both the graph and the secant go through the two points $(a, f(a))$ and $(b, f(b))$, we have immediately that $h(a) = h(b) = 0$, so Rolle's Theorem (Proposition 4.9.1) applies to give $c \in (a, b)$ with $h'(c) = 0$, or

$$f'(c) = m.$$

□

Let us consider some quick consequences. The first will prove particularly useful in Chapter 5:

Corollary 4.9.3. *Suppose f is continuous on an interval I and $f'(x) = 0$ for all interior points x of I.*
Then f is constant on I.

Proof. Suppose $f(x) \neq f(x')$ for two points in I; then the corresponding secant line has nonzero slope m, and by the Mean Value Theorem some point x'' between x and x' has $f'(x'') = m$, contradicting the assumption that $f'(x) = 0$ everywhere in I. □

The following geometric lemma will be useful in proving the second result, but is also interesting in its own right. Suppose we have two lines with different slopes, say L_1 with slope m_1 and L_2 with slope $m_2 > m_1$; they intersect at some point P, and it is reasonably clear that L_2 is *above* L_1 to the *right* of P and *below* L_1 to the *left* of P. Our lemma says that the same happens if we replace one of the lines with a curve tangent to it at P, and restrict attention to a small neighborhood of P.

Lemma 4.9.4. *Suppose f is differentiable at $x = x_0$, and let*

$$m = f'(x_0)\,.$$

If L is a line through $(x_0, f(x_0))$ with slope $m' \neq m$, then the graph of $y = f(x)$ crosses L at $(x_0, f(x_0))$: furthermore,

1. *if $m' > m$ then the graph is* BELOW *L slightly to the* RIGHT *of the point, and* ABOVE *it slightly to the* LEFT
2. *if $m > m'$ then the graph is* ABOVE *L slightly to the* RIGHT *of the point, and* BELOW *it slightly to the* LEFT

(See Figure 4.42.)

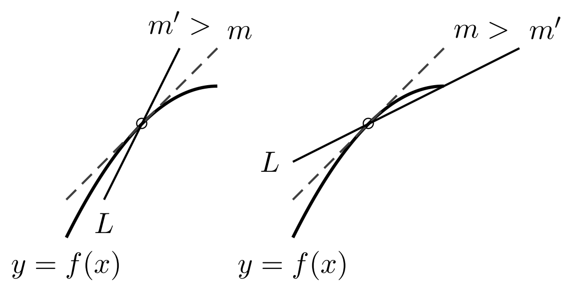


Figure 4.42: Lemma 4.9.4

Proof. The line L is the graph of a function of the form $\ell(x) = m'x + b$ with $\ell(x_0) = f(x_0)$. Consider the function

$$h(x) := \ell(x) - f(x)\,.$$

Clearly,

$$h(x_0) = 0$$
$$h'(x_0) = m' - m.$$

Then Lemma 4.7.2 in § 4.7 says that

1. if $m' > m$, then $h(x)$ is strictly positive for x slightly to the right of $x = x_0$ and strictly negative for x slightly to the left of $x = x_0$;
2. if $m' < m$, then $h(x)$ is strictly negative for x slightly to the right of $x = x_0$ and strictly positive for x slightly to the left of $x = x_0$.

But this is precisely what we wanted to prove. □

The following is the analogue of the Intermediate Value Theorem for derivatives—but notice that we are *not* assuming that the *derivative* function is continuous (examples do exist where the derivative exists on an interval but does not vary continuously with the point—for example, see Exercise 6 in § 4.5).

Corollary 4.9.5 (Darboux's Theorem). *If f is differentiable on a closed interval $[a, b]$, then for every number D strictly between $f'(a)$ and $f'(b)$, there exists a point $t \in (a, b)$ where*

$$f'(t) = D.$$

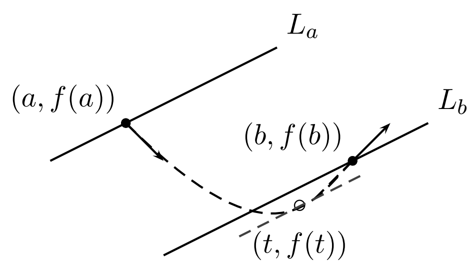


Figure 4.43: Corollary 4.9.5 (Darboux's Theorem)

Jean-Gaston Darboux (1842-1917) made distinguished contributions to the theory of singularities in differential equations, and particularly to the theory of surfaces; he also contributed improvements to Riemann's theory of integrals. I am not sure of the location of this theorem in the literature; however Darboux did isolate the conclusion of the Intermediate Value Theorem as an independent property (sometimes called *Darboux continuity*),[18] and showed that it does *not* imply continuity. This result shows that derivatives are always *Darboux* continuous.

Proof of Darboux's Theorem. Here, we sketch a geometric argument, using Figure 4.43. The idea is to find a line L of slope D which is crossed by the graph of $f(x)$ at two distinct points (corresponding to x-values in the interval $[a, b]$); this makes a segment of L a secant of the graph of $f(x)$,

[18]See the discussion following Theorem 3.2.1 concerning *discontiguity*.

and hence by the Mean Value Theorem the tangent to the graph is parallel to L (and hence has slope D) at some point $t \in (a, b)$.
Draw lines L_a through $(a, f(a))$ (*resp.* L_b through $(b, f(b))$) with slope D. If the two lines coincide, then we can take $L = L_a = L_b$, so assume they are different. Suppose that $f'(a) < D < f'(b)$. Lemma 4.9.4 tells us that for x slightly to the left of $x = b$, the graph of f lies below L_b, while for x slightly to the right of $x = a$, the graph lies below L_a. Let L be the lower of the two lines; then it separates points on the graph corresponding to $x \in (a, b)$ near one end of (a, b) from the point corresponding to the other end of $[a, b]$, and hence is crossed by the graph of $f(x)$ at an interior point of $[a, b]$ (what justifies this conclusion?).
We leave it to you to adapt the argument to the case when $f'(a) > D > f'(b)$ (Exercise 8). In Exercise 9 we consider an alternative, analytic approach to this argument. □

Next we will prove a more general version of Proposition 4.9.2, known as the *Cauchy Mean Value Theorem*, which talks about the ratio of two slopes. This was proved by Cauchy in his *Leçons sur le calcul differentiel* (1829), and used in establishing L'Hôpital's Rule as well as in proving a higher-order version of the Second Derivative Test for local extrema (Remark 4.8.7). Suppose we have *two* functions f and g, both satisfying the conditions of the Mean Value Theorem. The Mean Value Theorem says that there are points $c_f \in (a, b)$ (*resp.* $c_g \in (a, b)$) where the derivative of the function equals the slope of the corresponding secant, but of course there is no guarantee that these two points are the same. Cauchy's Mean Value Theorem does not assert the existence of such a common point, but it does assert the existence of a point c where the *ratio* of the derivatives equals the *ratio* of the secant slopes:

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Actually, this statement only makes sense if $g(b) \neq g(a)$, so we rewrite it in a form that does not require this, by multiplying through by $(g(b) - g(a))$ and $g'(c)$:

$$f'(c)\,(g(b) - g(a)) = g'(c)\,(f(b) - f(a))\,. \tag{4.39}$$

Theorem 4.9.6 (Cauchy Mean Value Theorem)**.** *Suppose f and g are DICE on $[a, b]$.*
Then there exists a point $c \in (a, b)$ satisfying Equation (4.39).

Proof. We consider the difference of the two sides of Equation (4.39) with $f'(c)$ (*resp.* $g'(c)$) replaced by f (*resp.* g):

$$h(x) = f(x)\,(g(b) - g(a)) - g(x)\,(f(b) - f(a))\,.$$

Evaluating at $x = a$, we have

$$h(a) = f(a)\,(g(b) - g(a)) - g(a)\,(f(b) - f(a)) = f(a)\,g(b) - g(a)\,f(b)$$

and similarly $h(b)$ gives the same value. Hence by Rolle's Theorem (Proposition 4.9.1) we have $h'(c) = 0$ for some $c \in (a, b)$. Note that h is just a linear combination of f and g, so we get

$$h'(x) = f'(x)\,(g(b) - g(a)) - g'(x)\,(f(b) - f(a))$$

and in particular, at $x = c$,

$$0 = h(c) = f'(c)\,(g(b) - g(a)) - g'(c)\,(f(b) - f(a))\,,$$

which is equivalent to Equation (4.39). □

Note that the Mean Value Theorem is really the special case of Theorem 4.9.6 where $g(x) = x$.

Exercises for § 4.9

Answers to Exercises 1aceg, 3ace, 4ac, 6, 10 are given in Appendix B.

Practice problems:

1. In each part below, you are given a function f and a closed interval $[a, b]$; for each, verify that the hypotheses of Rolle's Theorem (Proposition 4.9.1) hold, and identify all points c at which the conclusion holds:

   (a) $f(x) = x^2,\ [-1, 1]$
   (b) $f(x) = x^3 - 3x^2,\ \left[1 - \sqrt{3}, 1 + \sqrt{3}\right]$
   (c) $f(x) = 2x^3 - 3x^2,\ \left[0, \frac{3}{2}\right]$
   (d) $f(x) = 2x^3 - 3x^2,\ \left[-\frac{1}{2}, 1\right]$
   (e) $f(x) = e^x + e^{-x},\ [-1, 1]$
   (f) $f(x) = \frac{x^2 + 1}{x},\ \left[\frac{1}{2}, 2\right]$
   (g) $f(x) = \sin x + \cos x,\ \left[0, \frac{\pi}{2}\right]$
   (h) $f(x) = \sin x - \cos x,\ \left[\frac{\pi}{4}, \frac{5\pi}{4}\right]$

2. In each part below, you are given an equation and an interval; in each case, show that the equation has *precisely one* solution in the given interval:

   (a) $x^3 + x = 1$ on $[0, 1]$.
   (b) $x^5 + 3x^2 + 2x + 1 = 8$ on $[1, 2]$.
   (c) $x^3 - x^2 + x = -2$ on $(-1, 1)$.
   (d) $\cos x + x = 0$ on $[-2, 2]$.
   (e) $\cos x = x$ on $(0, \pi)$.

3. In each part below, you are given a function f and a closed interval $[a, b]$; for each, verify that the hypotheses of the Mean Value Theorem (Proposition 4.9.2) hold, and identify all points c at which the conclusion holds:

   (a) $f(x) = x^2,\ [0, 1]$
   (b) $f(x) = x^2,\ [-2, 1]$
   (c) $f(x) = x^3,\ [0, 2]$
   (d) $f(x) = x^3,\ [-1, 1]$
   (e) $f(x) = \dfrac{x^2 - 1}{x},\ \left[\dfrac{1}{2}, 2\right]$
   (f) $f(x) = \ln x,\ [1, e]$

4. In each part below, you are given two functions, f and g, and a closed interval $[a, b]$; for each, verify that the hypotheses of the Cauchy Mean Value Theorem (Theorem 4.9.6) hold, and identify all points c at which the conclusion holds; also, decide whether this point agrees with either of the points satisfying the conclusion of the Mean Value Theorem (Proposition 4.9.2) applied to f or g:

   (a) $f(x) = x^2,\ g(x) = x^3,\ [0, 1]$
   (b) $f(x) = x^2,\ g(x) = \dfrac{1}{x},\ [1, 2]$
   (c) $f(x) = \sin x,\ g(x) = \cos x,\ \left[\dfrac{\pi}{6}, \dfrac{\pi}{3}\right]$

Theory problems:

5. (M. Spivak) Show that the function $f(x) = x^3 - 3x + c$ cannot have two zeroes in $[0, 1]$ for any value of c.

6. A function is *even* if $f(-x) = f(x)$ for all x; it is *odd* if $f(-x) = -f(x)$ for all x. Show that every differentiable *even* function f has $f'(0) = 0$. Are there any restrictions on the derivative at $x = 0$ of an *odd* function?

7. **Differential Inequalities:**

   (a) Use Proposition 4.9.2 to show that in Proposition 4.8.3 we can replace the hypothesis that f is differentiable on I with the (weaker) assumption that f is DICE on I.

   (b) Use Proposition 4.9.2 to show that, if two functions f and g are both DICE on $[a,b]$ and satisfy

   $$\begin{aligned} f(a) &= g(a) \\ f'(x) &\leq g'(x) \text{ for all } x \in (a,b) \end{aligned}$$

   then

   $$f(x) \leq g(x) \text{ for all } x \in [a,b]$$

   (c) In particular, show that if

   $$\alpha \leq f'(x) \leq \beta$$

   for every $x \in (a,b)$, where f is DICE on $[a,b]$, then

   $$f(a) + \alpha(x-a) \leq f(x) \leq f(a) + \beta(x-b) \text{ for all } x \in [a,b]\,.$$

   (*Hint:* What are the derivatives of the two outer functions?)

8. Modify the proof of Darboux's Theorem (Corollary 4.9.5) to handle the case $f'(a) > D > f'(b)$.

9. **Another proof of Darboux's Theorem:** Here is another way to prove Corollary 4.9.5 from Lemma 4.9.4 [50, p. 199]: Suppose f is differentiable on $[a,b]$.

   (a) Show that if $f'(a) < 0$ and $f'(b) > 0$ then $f'(x) = 0$ for at least one point in (a,b). (*Hint:* Extrema!)

   (b) Show that if $f'(a) < c < f'(b)$ then $f'(x) = c$ for at least one point in (a,b). (*Hint:* Apply the previous result to the sum of f with an appropriately chosen function.)

10. Show that if a function f fails to be differentiable at just one point, its derivative can fail to be Darboux continuous.

Challenge problems:

11. (D. Benardete) Suppose f is twice-differentiable on $[-1, 1]$, and satisfies the following:

    - $f(-x) = f(x)$ for all x;
    - $f(0) = -1$;
    - $f(x) = 0$ for at least one point in $[-1, 1]$.

    Show that $f''(x) \geq 2$ for at least one point in $(-1, 1)$.
    (*Hint:* Find the function $g(x)$ satisfying all the hypotheses above with $g''(x) = 2$ for all $x \in [-1, 1]$. Use Problems 6 and 7, and/or the Mean Value Theorem (twice) to show that if $f(x)$ satisfies all the conditions above but $f''(x) < 2$ for all $x \in [-1, 1]$, then $f(x) \leq g(x)$ for $x \in [-1, 1]$ (with equality only at $x = 0$). This means $f(x) < 0$ for all $x \in [-1, 1]$, contrary to hypothesis.)

12. Show that if f is differentiable on $(0, 1)$, it is impossible for f' to have a jump discontinuity at $x = \frac{1}{2}$.

13. (a) Show that if f is continuously differentiable near $x = a$ and $f'(a) \neq 0$, then f is locally invertible: that is, there exist neighborhoods I of $x = a$ and J of $y = f(a)$ and a function $g(y)$ defined for $y \in J$ such that $g(f(x)) = x$ for all $x \in J$.

    (b) (M. Spivak) Consider the function

    $$f(x) = \begin{cases} \frac{x}{2} + x^2 \sin \frac{1}{x} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

    Show that $f'(0) \neq 0$, but f is not one-to-one on any neighborhood of $x = 0$.

History notes:

14. The first theorems proved by Cauchy using his rigorous definition of the derivative were the following. A translation of his treatment of both results is in [26, pp. 168-70]. Note that he assumed that the derivative of a continuous function always exists; we have inserted this as a hypothesis. Earlier proofs were given by Joseph Louis Lagrange (1736-1813), based on Taylor's Theorem (Theorem 6.1.7) and in 1806 by André-Marie Ampère (1775-1836) using related techniques (cf. [26, p. 122]).

(a) **Cauchy's Mean Value Inequality** If f is differentiable on $[a, b]$ then

$$\inf_{x\in[a,b]} f'(x) \leq \frac{f(b)-f(a)}{b-a} \leq \sup_{x\in[a,b]} f'(x)\,.$$

i. Prove this using Proposition 4.9.2 or Exercise 7.

ii. Historically, Cauchy proved this *before* he had the Mean Value Theorem. Here is an outline of his proof (based on the translation in [26]).

First, he says, given $\varepsilon > 0$, find $\delta > 0$ so small that whenever $|h| < \delta$ we can guarantee that

$$f'(x) - \varepsilon \leq \frac{f(x+h)-f(x)}{h} \leq f'(x) + \varepsilon.$$

for every $x \in [a, b]$. (*Note:* here, Cauchy is assuming a uniformity in the rate of convergence of secants to tangents—*a priori* δ will depend on both ε and x.)

Now, divide the interval $[a, b]$ into n equal pieces $[x_i, x_{i+1}]$, $i = 0, \ldots, n-1$ (with $x_0 = a$ and $x_n = b$) and consider each of the secant slopes

$$m_i = \frac{f(x_{i+1})-f(x_i)}{x_{i+1}-x_i},$$

picking n so that $h = x_{i+1} - x_i < \delta$ for all i. Then we have

$$f'(x_i) - \varepsilon \leq m_i \leq f'(x_i) + \varepsilon$$

for each i.

At this point, Cauchy invokes an obscure algebraic fact:

Suppose we have a family of fractions

$$\frac{a_1}{b_1} \leq \frac{a_2}{b_2} \leq \cdots \leq \frac{a_k}{b_k}$$

with all the denominators positive.
Then

$$\frac{a_1}{b_1} \leq \frac{a_1+a_2+\cdots+a_k}{b_1+b_2+\cdots+b_k} \leq \frac{a_k}{b_k}.$$

See Exercise 16 in § 1.2.
Use this to show that $\frac{f(b)-f(a)}{b-a}$ lies between $\min m_i$ and $\max m_i$, and hence between $\min f'(x_i) - \varepsilon$ and $\max f'(x_i) + \varepsilon$. (*Hint:* Write $f(b) - f(a) = (f(x_n) - f(x_{n-1})) + \cdots + (f(x_1) - f(x_0))$ with a corresponding "telescoping sum" for $b - a$.)
Since each $f'(x_i)$ lies between $\inf f'(x) - \varepsilon$ and $\sup f'(x) + \varepsilon$, we have

$$\inf f'(x) - \varepsilon \leq \frac{f(b)-f(a)}{x-a} \leq \sup f'(x) + \varepsilon.$$

But we can take $\varepsilon > 0$ arbitrarily small, proving the required inequality.

(b) Using this, show that if f' is continuous on $[a, b]$, then there exists some value of θ between 0 and 1 for which, setting $h = b - a$,

$$\frac{f(a+h)-f(a)}{h} = f'(a + \theta h) .$$

How does this differ from Proposition 4.9.2?

4.10 L'Hôpital's Rule and Indeterminate Forms

Guillaume Françoise Antoine, Marquis de L'Hôpital (1661-1704) is known as the author of the first printed calculus textbook, *Analyse des infiniment petits* (1696), which was very influential throughout the eighteenth century. In 1692, L'Hôpital, a good amateur mathematician, engaged Johann Bernoulli (1667-1748) (who, along with his brother Jacob Bernoulli (1654-1705) was one of the leading mathematicians of the time and a major contributor to the early development of the calculus) to instruct him in the new calculus in return for a regular salary; part of the deal was that L'Hôpital would be free to use Bernoulli's mathematical results as he pleased. What is now forever known as L'Hôpital's Rule was one of these results, which L'Hôpital included in his textbook. In the preface to this book, L'Hôpital acknowledged his indebtness to Leibniz and the Bernoullis, without however saying anything explicit about which results were due to whom.[19] L'Hôpital also wrote a textbook of analytic geometry,

[19]One's first reaction to this arrangement is outrage, and after L'Hôpital's death Bernoulli more or less accused L'Hôpital of plagiarism, but as Dirk Struik is reported to have said, "Let the good Marquis keep his elegant rule; he paid for it."[18, p. 36]

Traité analytique des sections coniques, published posthumously in 1707, which was equally influential [9].

L'Hôpital's Rule is very useful in evaluating certain kinds of limits, known as **indeterminate forms**. Recall that we have the general rule

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}.$$

This only makes sense if the denominator on the right is nonzero. However, if the denominator is zero but the numerator is a non-zero number, we can use common sense to help determine the limit: when x is near $x = a$, the numerator is near the nonzero number $\lim_{x\to a} f(x)$, while the denominator is *small*, so we can predict that the ratio is *large*; if we can somehow determine the *sign* of the ratio, then we can predict divergence to $\pm\infty$, and in any case we know the *absolute value* of the ratio diverges to ∞. However, if *both* $\lim_{x\to a} f(x) = 0$ *and* $\lim_{x\to a} g(x) = 0$, then we are trying to analyze an expression in which both numerator and denominator are small, and the behavior of the ratio is unpredictable without further information; we refer to this as the *indeterminate form* $\mathbf{\frac{0}{0}}$. We shall discuss other indeterminate forms later in this section.

Formally, L'Hôpital's Rule says that we can replace the limit of the ratio of the *functions* $\frac{f(x)}{g(x)}$ with the ratio of their *derivatives* $\frac{f'(x)}{g'(x)}$ when taking the limit. The version for $\frac{0}{0}$ (and only for $x \to b$ finite) is the one in L'Hôpital's book.

Proposition 4.10.1 (L'Hôpital's Rule for $\mathbf{\frac{0}{0}}$). *Suppose f and g are differentiable on the open interval (a, b) and*

$$\lim_{x\to b^-} f(x) = \lim_{x\to b^-} g(x) = 0 \tag{4.40}$$

$$\lim_{x\to b^-} \frac{f'(x)}{g'(x)} = L. \tag{4.41}$$

Then

$$\lim_{x\to b^-} \frac{f(x)}{g(x)} = L. \tag{4.42}$$

The same conclusions hold if the left-hand limit at b is replaced with the right-hand limit at a, as well as if convergence to L in Equation (4.41) *is replaced by divergence to $\pm\infty$.*

Proof. We will prove the main statement, and leave the variations mentioned in the last paragraph to you (Exercise 3, Exercise 4).

We separate two cases, depending on whether $b \in \mathbb{R}$ or $b = \infty$.
When $b \in \mathbb{R}$, Equation (4.40) means that by defining f and g to be zero at $x = b$, we have functions which satisfy the hypotheses of Theorem 4.9.6 on any closed interval of the form $[a', b]$ where $a < a' < b$. Suppose we have a sequence $\{x_n\}$ converging to b; without loss of generality, we assume it is strictly increasing. For each x_n, the Cauchy Mean Value Theorem with $a' = x_n$ insures the existence of a point $c_n \in (x_n, b)$ such that

$$\frac{f'(c_n)}{g'(c_n)} = \frac{f(b) - f(x_n)}{g(b) - g(x_n)} = \frac{f(x_n)}{g(x_n)}$$

where the last equality follows from Equation (4.40). But then $c_n \to b$, and so by Equation (4.41)

$$\lim \frac{f(x_n)}{g(x_n)} = \lim \frac{f'(c_n)}{g'(c_n)} = L$$

as required.
For the case $b = \infty$, we substitute $x = -\frac{1}{t}$ to consider the functions

$$F(t) = f\left(-\frac{1}{t}\right)$$
$$G(t) = g\left(-\frac{1}{t}\right)$$

for which we can apply the preceding argument with $t \to 0^-$; note that since the derivatives of these new functions are $t^{-2}f'\left(\frac{1}{t}\right)$ and $t^{-2}g'\left(\frac{1}{t}\right)$, respectively, their *ratio* is the same as the ratio of the derivatives of f and g, with $x = -\frac{1}{t}$, and the result follows. □

As an example, consider the limit

$$\lim_{x \to 0^+} \frac{\ln(\cos x)}{x}.$$

We check that since $\cos x \to 1$ as $x \to 0$, $\ln(\cos x) \to \ln 1 = 0$ and we have the indeterminate form $\frac{0}{0}$. L'Hôptial's rule says that this limit is equal to

$$\lim_{x \to 0^+} \frac{-(\sin x)\frac{1}{\cos x}}{1} = \frac{-(0)\frac{1}{1}}{1} = 0.$$

The indeterminate form $\frac{\infty}{\infty}$ arises when the numerator and denominator both diverge to infinity. The relevant version of L'Hôpital's Rule[20] is

[20] I am not sure where this first appeared.

Proposition 4.10.2 (L'Hôpital's Rule for $\frac{\infty}{\infty}$)**.** *Suppose f and g are differentiable on the open interval (a, b) and*

$$\lim_{x \to b^-} f(x) = \pm \lim_{x \to b^-} g(x) = \pm\infty \tag{4.43}$$

$$\lim_{x \to b^-} \frac{f'(x)}{g'(x)} = L. \tag{4.44}$$

Then

$$\lim_{x \to b^-} \frac{f(x)}{g(x)} = L. \tag{4.45}$$

The same conclusions hold if the left-hand limit at b is replaced with the right-hand limit at a, as well as if convergence to L in Equation (4.44) *is replaced by divergence to $\pm\infty$.*

Proof. Surprizingly, the proof of this variant is significantly different from that of Proposition 4.10.1. We begin with the following observation:

> **Claim:** *If $|\alpha|, |\beta| < \varepsilon < \frac{1}{3}$, then*
>
> $$\left| \frac{1+\alpha}{1+\beta} - 1 \right| < 3\varepsilon. \tag{4.46}$$

To see this, note that

$$\frac{1-\varepsilon}{1+\varepsilon} < \frac{1+\alpha}{1+\beta} < \frac{1+\varepsilon}{1-\varepsilon},$$

and

$$0 < \frac{1+\varepsilon}{1-\varepsilon} - 1 = \frac{2\varepsilon}{1-\varepsilon} < \frac{2\varepsilon}{2/3} = 3\varepsilon,$$

while

$$0 < 1 - \frac{1-\varepsilon}{1+\varepsilon} = \frac{2\varepsilon}{1+\varepsilon} < \frac{2\varepsilon}{1-\varepsilon}.$$

$\diamondsuit$

Now, suppose we have four numbers, a, A, b, B such that

$$\left|\frac{a}{A}\right|, \left|\frac{b}{B}\right| < \varepsilon < \frac{1}{3}.$$

Then multiplying Equation (4.46) with $\alpha = \frac{a}{A}$ and $\beta = \frac{b}{B}$ by $\left|\frac{A}{B}\right|$, we obtain the estimate

$$\left| \frac{A+a}{B+b} - \frac{A}{B} \right| < 3\varepsilon \left| \frac{A}{B} \right|. \tag{4.47}$$

Using this estimate, we prove L'Hôpital's Rule when f and g both diverge to ∞ as follows. Given a sequence of points $x_k \to b$, we can, passing to subsequences if necessary, assume without loss of generality that for $k = 1, 2, \ldots,$

$$x_k < x_{k+1} < b$$

and

$$f(x_{k+1}) > (k+3)f(x_k)\,, \;\; g(x_{k+1}) > (k+3)g(x_k)\,.$$

By the Cauchy Mean Value Theorem (Theorem 4.9.6) applied to $[x_k, x_{k+1}]$, there exist points c_k between x_k and x_{k+1} such that

$$\frac{f'(c_k)}{g'(c_k)} = \frac{f(x_{k+1}) - f(x_k)}{g(x_{k+1}) - g(x_k)}.$$

Note that in particular $c_k \to b$, so

$$\frac{f'(c_k)}{g'(c_k)} \to L.$$

Now, Equation (4.47) with $A = f(x_{k+1})$, $a = -f(x_k)$, $B = g(x_{k+1})$, $b = -g(x_k)$, and $\varepsilon = \frac{1}{k+3}$ reads

$$\left|\frac{f'(c_k)}{g'(c_k)} - \frac{f(x_{k+1})}{g(x_{k+1})}\right| < \frac{3}{k+3}\left|\frac{f(x_{k+1})}{g(x_{k+1})}\right|$$

or

$$\frac{f'(c_k)}{g'(c_k)} = (1+\delta_k)\frac{f(x_{k+1})}{g(x_{k+1})},$$

where

$$|\delta_k| < \frac{3}{k+3} \to 0,$$

which shows that

$$\lim \frac{f(x_{k+1})}{g(x_{k+1})} = \lim \frac{f'(c_k)}{g'(c_k)} = L,$$

which is what we needed to prove. □

For example, the limit

$$\lim_{x\to\infty} \frac{x}{e^x}$$

can be evaluated using L'Hôpital's Rule: Since both e^x and x diverge to infinity as $x \to \infty$, we can consider the limit of the ratio of their derivatives, which is

$$\lim_{x\to\infty} \frac{1}{e^x} = 0$$

(since the denominator diverges to infinity while the numerator is 1).
Sometimes, we can use L'Hôpital's Rule several times in succession. For example, the limit

$$\lim_{x\to\infty} \frac{x^2}{e^x}$$

leads, after replacing numerator and denominator by their derivatives, to

$$\lim_{x\to\infty} \frac{2x}{e^x}$$

which we might recognize as twice the limit we just finished doing via L'Hôpital's Rule. In fact, by induction we can establish

Lemma 4.10.3. *Let f be a polynomial. Then for any $a > 0$*

$$\lim_{x\to\infty} \frac{f(x)}{e^{ax}} = 0.$$

Proof. We go by induction on the degree d of f. For $d = 0$, f is a constant, and the result is trivial. Suppose we already know the result for all polynomials of degree $d - 1$; then applying L'Hôpital's Rule, we have

$$\lim_{x\to\infty} \frac{f(x)}{e^{ax}} = \lim_{x\to\infty} \frac{f'(x)}{ae^{ax}}$$

and since f' (and hence $\frac{1}{a}f'(x)$) is a polynomial of degree $d - 1$, we know that this last limit is zero, as required. □

Lemma 4.10.3 encodes in precise terms our intuition that *exponentials grow much faster than polynomials*. We introduce a notation which allows us to state this very efficiently.

Definition 4.10.4 ($\mathcal{O}$-notation)**.** *Suppose two functions f and g are defined on an open interval I with a an endpoint (possibly $a = \pm\infty$). We say that*

- $f(x) = \mathcal{O}(g(x))$ *(pronounced "$f(x)$* is big-oh of *$g(x)$") in I if $f(x)/g(x)$ is bounded on I;*
- $f(x) = \mathfrak{o}(g(x))$ *(pronounced "$f(x)$* is little-oh of *$g(x)$") as $x \to a$ (from the left) if*

$$\frac{f(x)}{g(x)} \to 0 \text{ as } x \to a \text{ (in } I\text{)}.$$

This constitutes an "abuse of notation", since we are not really defining $\mathcal{O}(g(x))$ or $\mathfrak{o}(g(x))$ as a quantity, but it is a standard shorthand. Of course, $f(x) = \mathfrak{o}(g(x))$ implies $f(x) = \mathcal{O}(g(x))$, but not vice versa (see Exercise 5). As easy examples, we have clearly

$$\sin x = \mathcal{O}(1) \text{ on } (-\infty, \infty)$$
$$x = \mathfrak{o}(1) \text{ as } x \to 0.$$

The following are easy observations (Exercise 6). If $m < n$ are real numbers, then

$$x^m = \mathfrak{o}(x^n) \text{ as } x \to +\infty$$

but

$$x^n = \mathfrak{o}(x^m) \text{ as } x \to 0^+.$$

Similarly, if $0 < b_1 < b_2$, then

$$b_1^x = \mathfrak{o}(b_2^x) \text{ as } x \to +\infty$$

but

$$b_2^x = \mathfrak{o}(b_1^x) \text{ as } x \to -\infty.$$

However, there is a standard hierarchy of basic functions for which L'Hôpital's rule gives us quick information.

Proposition 4.10.5. *1. For any $n > 0$,*

$$\ln x = \mathfrak{o}(x^n) \textit{ as } x \to +\infty$$
$$\ln x = \mathfrak{o}(x^{-n}) \textit{ as } x \to 0^+.$$

2. For any $n > 0$,

$$x^n = \mathfrak{o}(e^x) \textit{ as } x \to +\infty.$$

3. For any $b > 0$,

$$b^x = \mathfrak{o}(x^x) \textit{ as } x \to +\infty.$$

Proof. To compare $\ln x$ and x^n, we take the ratio

$$\frac{f(x)}{g(x)} = \frac{\ln x}{x^n}.$$

For $n > 0$, x^n and $\ln x$ both diverge to $+\infty$ as $x \to +\infty$, so theorem 4.10.2 gives

$$\lim_{x\to+\infty} \frac{\ln x}{x^n} = \lim_{x\to+\infty} \frac{x^{-1}}{nx^{n-1}} = \lim_{x\to+\infty} nx^{-n} = 0.$$

Replacing n with $-n$ and taking $x \to 0^+$, we have $\ln x \to -\infty$, $x^{-n} \to +\infty$ as $x \to 0^+$, so

$$\lim_{x\to 0^+} \frac{\ln x}{x^{-n}} = \lim_{x\to 0^+} \frac{x^{-1}}{-nx^{-n-1}} = \lim_{x\to 0^+} -\frac{1}{n}x^{-n} = 0.$$

To compare x^n and e^x, we look at the ratio

$$\frac{f(x)}{g(x)} = \frac{x^n}{e^x}.$$

For n positive, numerator and denominator both diverge to infinity as $x \to +\infty$, so again

$$\lim_{x\to+\infty} \frac{x^n}{e^x} = \lim_{x\to+\infty} \frac{nx^{n-1}}{e^x}.$$

For $0 < n < 1$, the new numerator converges to 0 while the denominator diverges to infinity, so the right hand limit is zero. Similarly for $n = 1$, the new numerator is bounded, so the ratio still goes to zero.

What if $n > 1$? Then the new numerator is a constant times x^{n-1} while the denominator is unchanged. In particular, if $1 < n \leq 2$, we know from the preceding argument that $x^{n-1} = \mathfrak{o}(e^x)$ as $x \to +\infty$, so the new ratio must go to zero, and hence also $x^n = \mathfrak{o}(e^x)$ as $x \to +\infty$. Inductively, if $N > 0$ is the integer for which $N < n \leq N + 1$, we can apply the version of Proposition 4.10.1 for the indeterminate form $\frac{\infty}{\infty}$ N times to obtain a ratio which we can prove goes to zero, and hence we can conclude that $x^n = \mathfrak{o}(e^x)$ for every $n > 0$.

To compare b^x with x^x as $x \to +\infty$, it is a bit difficult to use L'Hôpital's rule (the differentiation gets messy), but a different approach is easy: fixing $b > 0$, just notice that once $x > 2b$, we have

$$\frac{b^x}{x^x} = \left(\frac{b}{x}\right)^x < \left(\frac{1}{2}\right)^x \to 0.$$

□

Other Indeterminate Forms

A related limit is

$$\lim_{x\to 0^+} x \ln x$$

which yields the indeterminate form[21] $\mathbf{0} \cdot \boldsymbol{\infty}$. To apply L'Hôpital's Rule, we need to rewrite this as a ratio. This could be done in two ways:

$$x \ln x = \frac{x}{\frac{1}{\ln x}} = \frac{\ln x}{\frac{1}{x}};$$

the first version leads to the indeterminate form $\frac{0}{0}$ while the second leads to $\frac{-\infty}{\infty}$ (which can be handled just like $\frac{\infty}{\infty}$). However, if we differentiate numerator and denominator in the first version, we get

$$\frac{1}{-\frac{1}{x(\ln x)^2}}$$

which is worse, rather than better, as a limit to evaluate. The second version leads to

$$\lim_{x \to \infty} \frac{\frac{1}{x}}{-\frac{1}{x^2}}$$

which may seem just as bad, but then we can do the algebra to rewrite the fraction as

$$\frac{\frac{1}{x}}{-\frac{1}{x^2}} = -\frac{x^2}{x} = -x$$

which clearly goes to 0, so

$$\lim_{x \to 0^+} x \ln x = 0.$$

To handle indeterminate limits of the form $\boldsymbol{\infty} - \boldsymbol{\infty}$, like

$$\lim_{x \to 0^+} \left(\frac{1}{\sin x} - \frac{1}{x} \right)$$

we need to transform them into ratios, in this case by combining the fractions over a common denominator

$$\frac{1}{\sin x} - \frac{1}{x} = \frac{x - \sin x}{x \sin x}$$

whose limit at 0 yields the indeterminate form $\frac{0}{0}$. L'Hôpital's Rule yields the limit

$$\lim_{x \to 0^+} \frac{1 - \cos x}{\sin x + x \cos x};$$

[21] This notation is used whenever one factor goes to zero and the other becomes large, whether positive or negative.

this is also of the form $\frac{0}{0}$, and we can handle it in one of two ways. One is to apply L'Hôpital's Rule again, to get

$$\lim_{x\to 0^+} \frac{\sin x}{2\cos x - x\sin x}$$

whose numerator goes to zero, but whose denominator goes to 2—hence the last limit is $\frac{0}{2} = 0$. Or, we might notice that dividing numerator and denominator by x yields

$$\lim_{x\to 0^+} \frac{\frac{1-\cos x}{x}}{\frac{\sin x}{x} + \cos x}$$

whose numerator goes to zero 0 (by Exercise 7 in § 3.4) and denominator goes to $1 + 1$ (the first by Lemma 3.4.8), so again the ratio goes to 0.
It is important to check, before using L'Hôpital's Rule, that the ratio we are looking at does represent an indeterminate form. For example, if we had not noticed that the last limit above is *not* indeterminate and just applied L'Hôpital's Rule again, we would have obtained a ratio that diverges to ∞.
Another, and perhaps more subtle, indeterminate form is $\mathbf{0^0}$, for example the limit

$$\lim_{x\to 0^+} x^x.$$

Some of us might be tempted to reason that raising any number to the power 0 yields 1, while others might reason that 0 raised to any power is 0, and of course this conundrum must be resolved by more careful analysis.
Again, we need to convert this to a different form. One obvious move is to apply the natural logarithm, and look at the limit

$$\lim_{x\to 0^+} \ln(x^x) = \lim_{x\to 0^+} x\ln x.$$

We have already seen that this last limit is 0; it follows that

$$\lim_{x\to 0^+} (x^x = e^{x\ln x}) = e^{\lim_{x\to 0^+} x\ln x} = e^0 = 1.$$

Another example along these lines is

$$\lim_{x\to\infty} \left(1 + \frac{p}{x}\right)^x$$

where $p > 0$ is an arbitrary constant, which is an example of the indeterminate form $\mathbf{1^\infty}$. This time, while 1 to any power is 1, a number slightly above 1 raised to a large enough power is arbitrarily large. This

particular limit comes up for example in connection with compound interest: an annual percentage rate of $100p$ percent can be calculated on principal at varying frequencies. If it is calculated annually and then added to the principal, then at the end of a year the principal is multiplied by $1+p$. However, if it is calculated (and added to the principal) monthly, then each month the principal is multiplied by $1+\frac{p}{12}$; this occurs 12 times in the course of one year, so at the end of one year the principal has been multiplied by $\left(1+\frac{p}{12}\right)^{12}$, which is our expression above with $x=12$. This can be applied to interest you earn in a bank account, or interest you owe, say on a credit card bill. If you borrowed from a loan shark, the compounding would occur far more frequently: weekly compounding would correspond to $x=52$, daily to $x=365$, and so on. What would happen if the ultimate loan shark, using contemporary technology, decided to compound *instantaneously*? Would you owe a gazillion dollars at the end of a year for each dollar you borrowed at the beginning?
Again, the given limit cannot be attacked head-on; rather we try to determine (the limit of) its natural logarithm:

$$\lim_{x\to\infty} x\ln\left(1+\frac{p}{x}\right)$$

which has the indeterminate form $\infty\cdot 0$. As usual, we turn this into a ratio

$$\lim_{x\to\infty}\frac{\ln\left(1+\frac{p}{x}\right)}{\frac{1}{x}}$$

which has the form $\frac{0}{0}$, hence we can apply L'Hôpital's Rule to reduce it to

$$\lim_{x\to\infty}\frac{-\frac{p}{x^2}\cdot\frac{x}{x+p}}{-\frac{1}{x^2}}=\lim_{x\to\infty}\frac{px}{x+p}=p.$$

This gives us the limit of the logarithm of our expression, which by continuity is the logarithm of the limit, so we have that

$$\lim_{x\to\infty}\left(1+\frac{p}{x}\right)^x=e^p.$$

At a 100% APR ($p=1$), this means the amount more than doubles, but less than triples, over one year. This limit (with $p=1$) is sometimes used as a definition for the number e; note that our definition is different.
This problem was proposed by Jacob Bernoulli (1654-1705) in his *Ars Conjectandi* (1713). Euler derived the limit above in [22, Chap. 7]

Similar situations give rise to the indeterminate form $\infty^{\mathbf{0}}$, and are handled in a similar way.

Exercises for § 4.10

Answers to Exercises 1acegikmoqs, 4, 5 are given in Appendix B.

Practice problems:

1. Evaluate each limit below, justifying your conclusions:

   (a) $\lim_{x\to 0} \frac{x^2 - x}{x^2 - 2x}$ (b) $\lim_{x\to 2} \frac{\sqrt[3]{x} - \sqrt[3]{2}}{x - 2}$ (c) $\lim_{x\to 0} \frac{x^2}{1 - \cos x}$

   (d) $\lim_{x\to 0} \frac{1 - \cos x}{\sin x}$ (e) $\lim_{x\to 0} \frac{1 - \cos x}{x \sin x}$ (f) $\lim_{x\to 1} \frac{\ln x}{1 - x}$

   (g) $\lim_{x\to 1} \frac{x^2 + 1 - 2e^{x-1}}{(x-1)^3}$ (h) $\lim_{x\to\infty} x \arctan \frac{1}{x}$

   (i) $\lim_{x\to \frac{\pi}{2}} \frac{\tan x - 10}{\sec x + 10}$ (j) $\lim_{x\to\infty} (\sqrt{x+1} - \sqrt{x-1})$

   (k) $\lim_{x\to\infty} (\sqrt{x^2 + x} - x)$ (l) $\lim_{x\to\infty} x(e^{1/x} - 1)$

   (m) $\lim_{x\to 0} (2x)^x$ (n) $\lim_{x\to\infty} x^{1/x}$ (o) $\lim_{x\to 0^+} x^{1/x}$

   (p) $\lim_{x\to 0} (1+x)^{2/x}$ (q) $\lim_{x\to 0} (1+2x)^{1/x}$

   (r) $\lim_{x\to\infty} \frac{e^x - e^{-x}}{e^x + e^{-x}}$ (*Hint:* This problem and the next require some algebraic manipulation.)

   (s) $\lim_{x\to 1^-} [(\ln x)(\ln(1-x))]$

2. Show that the function

$$f(x) = \begin{cases} \frac{\sin x}{x} & \text{for } x \neq 0, \\ 1 & \text{for } x = 0 \end{cases}$$

   is differentiable at $x = 0$, and find its derivative there.

Theory problems:

3. (a) Given two real numbers $a < b$, find a function of the form $t(x) = \alpha x + \beta$ such that $t(a) = b$ and $t(b) = a$.

(b) Show that if f is differentiable on the open interval (a, b) then the function defined by

$$\tilde{f}(x) := f(t(x))$$

is also differentiable on (a, b), with

$$\tilde{f}'(x) = -f'(t(x))$$
$$\lim_{x\to a^+} \tilde{f}(x) = \lim_{x\to b^-} f(x)$$
$$\lim_{x\to a^+} \tilde{f}'(x) = -\lim_{x\to b^-} f'(x)\,.$$

(c) Apply this trick to the functions f and g to prove the version of L'Hôpital's Rule (Proposition 4.10.1) for right-hand limits at $a \in \mathbb{R}$.

(d) What trick can you use to take care of the case $a = -\infty$?

4. (a) We know from Lemma 3.4.10 that if $\frac{f(x)}{g(x)}$ diverges to infinity as $x \to b^-$ then $\frac{g(x)}{f(x)}$ converges to 0. Show that the converse statement does not hold, by giving an example of functions f and g for which $\lim_{x\to b^-} \frac{g(x)}{f(x)} = 0$ but $\lim_{x\to b^-} \frac{f(x)}{g(x)} \neq +\infty$.

(b) What further condition on $\frac{g(x)}{f(x)}$ is needed to guarantee that $\lim_{x\to b^-} \frac{f(x)}{g(x)} = +\infty$?

(c) Use this to prove the case of L'Hôpital's rule when the limit L in Equations (4.41) and (4.42) is replaced with $+\infty$.

(d) What should you do to prove the case when $L = -\infty$?

(e) What about the case when both b and L are $+\infty$?

5. Give an example of two functions f and g such that $f(x) = \mathcal{O}(g(x))$ but $f(x) \neq \mathfrak{o}(g(x))$.

6. (a) Suppose $m < n$ are real numbers. Show that

$$x^m = \mathfrak{o}(x^n) \text{ as } x \to +\infty$$

but

$$x^n = \mathfrak{o}(x^m) \text{ as } x \to 0^+.$$

(*Hint:* Express the ratio as a single power of x.)

(b) Suppose $0 < b_1 < b_2 < \infty$. Show that

$$b_1^x = \mathfrak{o}(b_2^x) \text{ as } x \to +\infty$$

but

$$b_2^x = \mathfrak{o}(b_1^x) \text{ as } x \to -\infty.$$

Challenge problems:

7. Consider the limit

$$\lim_{x \to 0^+} \frac{e^{-1/x}}{x}$$

which is indeterminate of type $\frac{0}{0}$.

(a) Show that repeated application of L'Hôpital's Rule to the given ratio cannot resolve the evaluation of this limit.

(b) Make the substitution $t = \frac{1}{x}$ and evaluate the limit using L'Hôpital's Rule.

(c) Show how by reinterpreting the original ratio as $\frac{1/x}{e^{1/x}}$ we could also have resolved the limit.

(d) Perform a similar analysis to evaluate

$$\lim_{x \to 0^+} \frac{e^{-1/x}}{x^n}$$

where n is any integer (you may want to separate the cases of n positive and n negative).

8. Consider the function $f(x) = e^{-1/x^2}$ defined for $x > 0$.

(a) Show that for each $k = 1, 2, ...$, the k^{th} derivative of f has the form

$$f^{(k)}(x) = \frac{p_k(x)}{x^{3k}} e^{-1/x^2}$$

where $p_k(x)$ is a polynomial.

(b) Substituting $y = x^2$, use Exercise 7 to show that for any polynomial $p(x)$ and any positive power x^m of x,

$$\lim_{x \to 0^+} \frac{p(x)}{x^m} e^{-1/x^2} = 0.$$

(c) Use this to show that the function defined by

$$f(x) = \begin{cases} e^{-x^2} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

has derivatives of all orders at $x = 0$.

This example was given by Cauchy in his *Résumé des lecons sur le calcul infinitesimal* (1823) to counter Lagrange's idea [37] of basing calculus on power series.

4.11 Continuity and Derivatives (Optional)

We saw in § 4.1 (Remark 4.1.3) that any differentiable function is automatically continuous. We have also seen examples of continuous functions (like the absolute value) which are *not* differentiable at some points. Here we will go further and construct an example of a function which is continuous on $[0, 1]$, but is *not* differentiable *anywhere*.
The first such example was constructed by Bernhard Bolzano (1781-1848) in the 1830's but like much of his work it went unnoticed; in fact, Bolzano's manuscript was not published until 1930. An English translation has appeared very recently [7]. We will discuss this example in Exercise 5.
The first example to be widely discussed was given by Karl Theodor Wilhelm Weierstrass (1815-1897) in lectures around 1861, and in a paper presented to the Royal Prussian Academy of Sciences in 1872; an English translation of this is given in [54]. The example was first published in 1875 by Paul Du Bois-Reymond (1818-1896), a follower of Weierstrass.
Weierstrass's construction involved a series of functions, and is slightly beyond the scope of our course: just for the record, here is the formula:

$$f(x) = \sum_{n=0}^{\infty} b^n \cos(a^n \pi x)$$

where a is an odd (positive) integer, and $b \in (0, 1)$ satisfies $ab > 1 + \frac{3\pi}{2}$.
We shall study a more geometric construction, based in part on [42], which is closer in spirit to Bolzano's approach.[22]
The construction of our example is reminiscent of the definition of the function 2^x in § 3.6: we begin by defining $f(x)$ for certain rational numbers

[22] Another construction, treated more analytically, is given in [35]. The graphs of all of these functions are examples of *fractals*.

$x \in [0,1]$ and then extend the definition to all real numbers in $[0,1]$ by means of sequences of approximations.

Definition at triadic rationals:

We start by defining f at the endpoints of $[0,1]$ by

$$f(0) = 0,$$
$$f(1) = 1.$$

From here, we proceed in stages. At each stage, we have already defined f at a finite set of x-values; we divide each "gap" between successive values into three equal intervals and define f at the two new division points. At the first stage, we start with the single gap $[0,1]$ and introduce the division points $x = \frac{1}{3}$ and $x = \frac{2}{3}$. At the second stage, the three gaps $\left[0, \frac{1}{3}\right]$, $\left[\frac{1}{3}, \frac{2}{3}\right]$ and $\left[\frac{2}{3}, 1\right]$ are divided by the respective pairs of points $\{\frac{1}{9}, \frac{2}{9}\}$, $\{\frac{4}{9}, \frac{5}{9}\}$, and $\{\frac{7}{9}, \frac{8}{9}\}$. It is easy to see that at the k^{th} stage we start with f defined at all x-values which can be expressed as fractions with denominator 3^{k-1}, and we extend the definition to all the new values obtained by allowing the denominator to be 3^k. The rationals which can be written with denominator a power of 3 are called **triadic rationals**, and we shall refer to the stage at which a given triadic rational t first appears as its **order**: this is just the exponent appearing in the denominator when t is written in lowest terms.

The process of definition at the k^{th} stage goes as follows. If $0 \leq t_0 < 1$ is a triadic rational of order at most k and $t_0' = t_0 + \frac{1}{3^k}$ is the *next* such rational[23], we consider the line segment joining the known points $(t_0, f(t_0))$ and $(t_0', f(t_0'))$ on the graph of $f(x)$. This is divided into three equal subsegments by the two points (t_1, y_1) and (t_2, y_2), where

$$t_1 := t_0 + \frac{1}{3^{k+1}}$$
$$t_2 := t_0 + \frac{2}{3^{k+1}} = t_0' - \frac{1}{3^{k+1}}.$$

We obtain two new points on the graph of f by interchanging the y-coordinates of the two division points (Figure 4.44), giving

$$f(t_1) = y_2$$
$$f(t_2) = y_1.$$

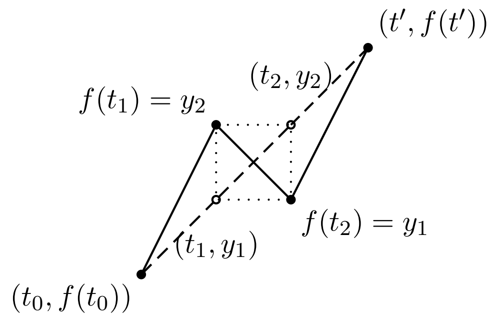


Figure 4.44: Basic Step

This process can be described geometrically, as follows. At the beginning of the k^{th} stage, we draw a broken line through the $3^{k-1}+1$ known points on the graph of f. The segment joining two adjacent points, $(t_0, f(t_0))$ and $(t_0', f(t_0'))$, has slope

$$m_k(t_0) := \frac{f(t') - f(t_0)}{t' - t_0}$$

and forms the diagonal of a rectangle

$$R_k(t_0) = I_k(t_0) \times J_k(t_0)$$

where $I_k(t_0) := [t_0, t_0']$ and $J_k(t_0)$ (the projection of $R_k(t_0)$ on the y-axis) is the closed interval with endpoints $f(t_0)$ and $f(t_0')$. The diagonal of $R_k(t_0)$ is divided into three subsegments by $(t_1, f(t_1))$ and $(t_2, f(t_2))$. The middle subsegment forms the diagonal of a subrectangle $R_{k+1}(t_1)$; to form the broken line for the next stage, we replace this with the other diagonal of $R_{k+1}(t_1)$, whose slope is

$$m_{k+1}(t_1) = -m_k(t_0)$$

and the other two subsegments of the diagonal of $R_k(t_0)$ with segments of slope

$$m_{k+1}(t_0) = m_{k+1}(t_2) = 2m_k(t_0)\,.$$

These are diagonals of two other subrectangles,

$$R_{k+1}(t_0)\,, R_{k+1}(t_2) \subset R_k(t_0)$$

[23]Note that at least one of t_0 or t_0' has order *precisely* k

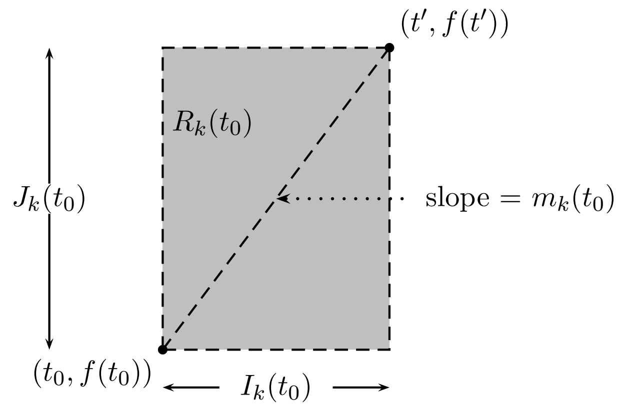


Figure 4.45: $R_k(t_0)$

whose width (*resp.* height) is one-third (*resp.* two-thirds) that of $R_k(t_0)$ (Figure 4.46). Note that the union of the k^{th} order rectangles contains all known points on the graph at the beginning of the k^{th} stage.
The relation between the geometry at the beginning of the k^{th} stage and that at the end is summarized in

Lemma 4.11.1. *With the notation above,*

1. $m_0(t) = 1$ *for all* t*; if* t_0 *is a triadic rational of order at most* k*, and* $t_i = t_0 + \frac{i}{3^k}$ *for* $i = 1, 2$*, then*

$$m_{k+1}(t_0) = m_{k+1}(t_2) = 2m_k(t_0)$$
$$m_{k+1}(t_1) = -m_k(t_0)$$

2. *For* $i = 0, 1, 2$*,*

$$R_{k+1}(t_i) \subset R_k(t_0)$$

is a subrectangle whose width $\|I_{k+1}(t_i)\|$ *and height* $\|J_{k+1}(t_i)\|$ *satisfy*

$$\|I_{k+1}(t_i)\| = \frac{1}{3}\|I_k(t_0)\|$$
$$\|J_{k+1}(t_i)\| \leq \frac{2}{3}\|I_k(t_0)\|.$$

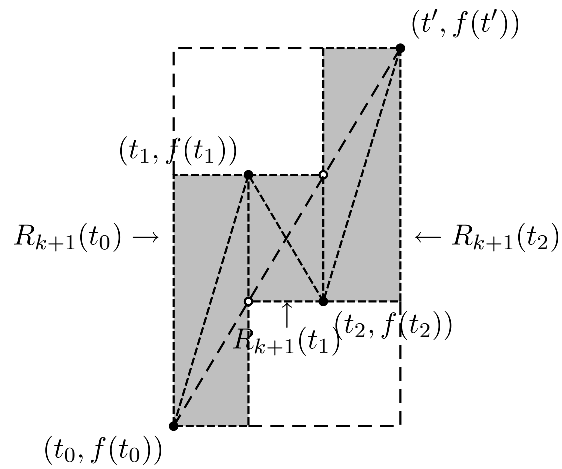


Figure 4.46: $R_{k+1}(t_i)$, $i = 0, 1, 2$ inside $R_k(t_0)$

3. *For every triadic rational t between t_0 and t_0' (of any order*[24]*) the point $(t, f(t))$ belongs to $R_k(t_0)$, and*

$$R_n(t) \subset R_k(t_0)$$

for all $n \geq$ order of t.

The proof of each of these observations is straightforward, and left to you (Exercise 2).
In Figure 4.47 we sketch the first four stages of this definition.

Definition at other reals:

Having defined our function for every triadic rational $t \in [0, 1]$, we need to define it for all real $x \in [0, 1]$. To this end, note that given $x \in [0, 1)$ and $n = 1, 2, ...$, there is a unique pair $t_n(x)$, $t_n'(x)$ of adjacent triadic rationals of order $\leq n$ bracketing x:

$$t_n(x) \leq x < t_n'(x) := t_n(x) + \frac{1}{3^n}.$$

For x fixed, the sequence $\{t_n(x)\}$ is nondecreasing, and if $x = t$ is itself a triadic rational of order k, then $t_n(x) = t$ for $n \geq k$. Thus it is natural to use the known values $f(t_n(x))$ as approximations to f.

[24]necessarily exceeding that of t_0

Figure 4.47: First four steps

By Lemma 4.11.1(3), the intervals $J_k(x)$ are nested, and by Lemma 4.11.1(2) their lengths go to zero, so by Corollary 2.6.2 there is a unique point $f(x)$ in their intersection, and it is the limit of the endpoints $f(t_k(x))$. In particular, when $x = t$ is itself triadic, then this definition yields the same value $f(x) = f(t)$ as the earlier one.

Continuity:

Lemma 4.11.2. *f as defined above is continuous on $[0, 1]$; that is,*

$$\lim_{x' \to x^+} f(x') = f(x) \qquad \text{for } x \in [0, 1)$$
$$\lim_{x' \to x^-} f(x') = f(x) \qquad \text{for } x \in (0, 1] \, .$$

Proof. If $x_n \downarrow x$ in $[0, 1]$, then for each $k = 1, 2, ...$ we must eventually have $t_k(x) < x_n < t'_k(x)$, so $(x_n, f(x_n)) \in R_k(t_k(x))$ and hence (eventually) $f(x_n) \in I_k(x)$ by Lemma 3. It follows immediately that

$$\lim_{x' \to x^+} f(x') = f(x) \text{ for } x \in [0, 1) \, .$$

If $x_n \uparrow x$ and x is *not* a triadic rational, then since $t_n(x) < x < t'_n(x)$, the same argument shows that $f(x_n) \to f(x)$. If $x = t$ *is* a triadic rational, then for k exceeding the order of t, we can consider the k^{th}-order rectangle $R_k(t''_0)$, where $t''_0 = t - 3^{-k}$ is the triadic rational immediately to the *left* of t. For n sufficiently large, $t''_0 < x_n < t$, and hence $f(x_n) \in J_k(t''_0)$: this clearly shows that $f(x_n) \to f(t)$, as required. □

Non-differentiability:

Finally, we need to show that f is not differentiable at any point $x \in (0,1)$. This will be straightforward for x a triadic rational, but for other x we will need the following easy observation (Figure 4.48; see Exercise 3):

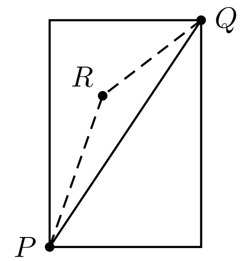


Figure 4.48: Comparing slopes of secants

Remark 4.11.3. *Suppose P and Q are distinct points on a line of slope m and R has x-coordinate between those of P and Q. Then one of the two line segments PR and RQ has slope $\geq m$ while the other has slope $\leq m$.*

Using this, we can prove

Proposition 4.11.4. *The function f defined above is continuous on $[0,1]$ but not differentiable at any point of $(0,1)$.*

Proof. We have established continuity in Lemma 4.11.2.
To establish nondifferentiability, fix $x_0 \in (0,1)$ and consider the sequence of slopes $m_k(t_k(x_0))$. By Lemma 4.11.1(1), for each k

$$m_{k+1}(t_{k+1}(x_0)) = \lambda_k m_k(t_k(x_0))$$

where λ_k is either -1 or 2. Since $m_1(t_1(x_0)) \neq 0$, this sequence is divergent (Exercise 1).
If $x_0 = t$ is a triadic rational, then eventually $t_k(x_0) = t$, so we have found a sequence of points $\{t'_k(t)\}$ converging to t for which the corresponding secants have divergent slopes $m_k(t)$, and consequently $f'(t)$ is undefined.
If x_0 is *not* a triadic rational, then we can, for each k, use Remark 4.11.3 to pick x'_k one of the two points $t_k(x_0)$ or $t'_k(x_0)$ in such a way that the slope $m_k(x_0)$ of the secant joining $x = x_0$ and $x = x'_k$ has the same sign as $m_k(t_k(x_0))$ and $|m_k(x_0)| \geq |m_k(t_k(x_0))|$. Again, this guarantees divergence of the slopes $m_k(x_0)$, so $f'(x_0)$ is undefined. $\square$

Exercises for § 4.11

1. Suppose $\{a_n\}$ is a sequence of nonzero numbers with the property that for $n = 1, 2, ...$ $\frac{a_{n+1}}{a_n}$ is either 2 or -1.

   (a) Show that for all n, $|a_{n+1} - a_n| \geq |a_1| > 0$.

   (b) Show that the sequence has no limit. (*Hint:* One way is to use Exercise 9 in § 2.5.)

   (c) Suppose $\{b_n\}$ is another sequence such that for each n a_n and b_n have the same sign, and $|b_n| \geq |a_n|$. Show that $\{b_n\}$ diverges.

2. Prove Lemma 4.11.1. (In the first two parts, you need to simply examine the process of going from one stage to the next with some care. The third part should be easy.)

3. Prove Remark 4.11.3. (*Hint:* Let S be the point on PQ with the same x-coordinate as R. Compare the slope of PR (*resp.* RQ) with that of PS (*resp.* SQ) using a comparison of the y-coordinates of R and S.)

Challenge problem:

4. **Triadic expansions:** The definition of our example f can be understood in terms of the **triadic expansion** (or base three expansion) of a number $x \in [0, 1]$. By analogy with the decimal and binary expansions (§ 2.4), we write

$$x \sim .\tau_1\tau_2\tau_3 \ldots$$

where each τ_i is either 0, 1, or 2, to mean that x is the sum of the series

$$\sum_{i=1}^{\infty} \frac{\tau_i}{3^i} = \frac{\tau_1}{3} + \frac{\tau_2}{3^2} + \frac{\tau_3}{3^3} + \ldots$$

Every number has a triadic expansion, which is unique unless x is a dyadic rational (of order k), in which case there is one expansion with $\tau_i = 0$ for all $i > k$ (which we write as the finite expansion $.\tau_1\tau_2 \ldots \tau_k$) and another in which τ_k is lowered by one and followed by an infinite string of 2's.

In our example, when t is a triadic rational, the interval $I_k(t)$ consists precisely of those points $x \in [0,1]$ which have a triadic expansion agreeing with the "finite" expansion of t to at least k places (if t has order $\ell < k$, then τ_i is zero for $i = \ell+1, \ldots, k$).

Let us define two functions $\sigma(\tau)$ and $\rho(\tau)$ for $\tau = 0, 1, 2$ by

$$\begin{array}{ll} \sigma(0) = 2 & \rho(0) = 0 \\ \sigma(1) = -1 & \rho(1) = 2 \\ \sigma(2) = 2 & \rho(2) = 1. \end{array}$$

(a) Show that if $t \sim .\tau_1\tau_2\ldots\tau_k$ is a triadic rational of order at most k, then

$$m_k(t) = m_k(.\tau_1\tau_2\ldots\tau_k) = \prod_{j=1}^{k} \sigma(\tau_j) = (-1)^s 2^{k-s},$$

where s is the number of 1's among the τ_i.

(b) Show that if $t \sim .\tau_1\tau_2\ldots\tau_k$ has order *precisely* k (*i.e.*, $\tau_k \neq 0$), then

$$f(t) = \sum_{j=1}^{k} \frac{\mu_j \rho(\tau_j)}{3^j}$$

where $\mu_1 = 1$ and for $j = 2, \ldots, k$, $\mu_j = m_{j-1}(.\tau_1\ldots\tau_{j-1})$.

(c) Suppose x is not a triadic rational, and its (unique) triadic expansion is

$$x \sim .\tau_1\tau_2\tau_3\ldots.$$

Show that the series

$$\sum_{j=1}^{\infty} \frac{\mu_j \rho(\tau_j)}{3^j}$$

has all terms of the form $\pm\frac{2^i}{3^j}$ with $i \leq j$. Use this to show that the partial sums

$$S_N := \sum_{j=1}^{N} \frac{\mu_j \rho(\tau_j)}{3^j}$$

satisfy

$$|S_N - S_M| \leq \sum_{i=N+1}^{\infty} \left(\frac{2}{3}\right)^i \text{ for } M > N,$$

and use Exercise 9 in § 2.5 to show that the series is convergent.

(d) Show that if $t \sim .\tau_1 \ldots \tau_k$ is a triadic rational of order precisely k, then the series as above obtained from the infinite expansion of t

$$t \sim .\tau_1 \ldots \tau_{k-1}(\tau_k - 1)222...$$

converges to $f(t)$.

(e) Suppose $x_n \in [0, 1]$ are a sequence of numbers with respective triadic expansions

$$x_n \sim \tau_1(n)\tau_2(n)\tau_3(n) \ldots.$$

Show that

$$\lim x_n = L$$

precisely if given k there exists N_k such that whenever $n \geq N_k$ the expansion of x_n agrees to at least k places with some expansion of L.

(f) Use this to show that f is continuous.

I have not seen a clear way to show non-differentiability directly from the triadic expansion point of view; perhaps this can be posed as an *Xtreme Challenge*!

History note:

5. **Bolzano's example:** [46, pp. 351-2, 487-9, 507-508] Bolzano's example was constructed in a way analogous to our construction.

   (a) Given a line segment from (a, A) to (b, B) with slope

   $$m = \frac{B - A}{b - a}$$

   let

   $$(c, C) = (\frac{a + b}{2}, \frac{A + B}{2})$$

   be its midpoint. We take the right quarter of each of the segments AC and CB and reflect it about the horizontal through its right endpoint; this moves two points to new positions:

   $$\left(a + \frac{3}{8}(b - a), A + \frac{3}{8}(B - A)\right) \text{ moves to } \left(a + \frac{3}{8}(b - a), A + \frac{5}{8}(B - A)\right)$$
   $$\left(a + \frac{7}{8}(b - a), A + \frac{7}{8}(B - A)\right) \text{ moves to } \left(a + \frac{7}{8}(b - a), A + \frac{9}{8}(B - A)\right).$$

Moving these points also rotates the first three-quarters of each segment to a new position, so we end up with a broken line consisting of four segments of respective slopes $\frac{5}{3}m$, $-m$, $\frac{5}{3}m$, and $-m$. We can think of this as moving each point of the line segment vertically to a new position. The highest point on the new broken line is the next-to-last breakpoint, whose new height is

$$A + \frac{9}{8}(B - A) = B + \frac{B - A}{8}$$

while the lowest is the left endpoint, which stands still at height B. Thus the vertical distance between the lowest and highest points of the new broken line is 9/8 times the vertical distance between the endpoints of the original line segment.

(b) In general, at the k^{th} stage we have 4^{k-1} segments; each has slope of the form $\pm(\frac{5}{3})^i m$, where $i \leq k$; its width (that is $\triangle x$ across it) is at most $(\frac{3}{8})^k(b - a)$. After the rotation, the height (*i.e.*, $\triangle y$) of the whole segment is multiplied by $\frac{9}{8}$, but the height of each of the four *sub*segments *after* the rotation is at most $\frac{5}{8}$ of the height of the *whole* segment *before* the rotation. Show that this means that any point is moved, at the k^{th} stage, by at most $(\frac{5}{8})^k(B - A)$. (Bolzano brings in the factor $\frac{9}{8}$, but it is not clear to me that this is necessary.) It follows that between the k^{th} and any later stage, the total vertical motion is bounded by

$$\sum_{r=1}^{\infty} \left(\frac{5}{8}\right)^{k+r} (B - A) = \frac{5^k}{3 \cdot 8^{k-1}}$$

which can be made as small as possible by choosing k large. This shows that the positions of a given point at successive stages form a Cauchy sequence and hence converge by Exercise 9 in §2.5.

(c) Initially, Bolzano produced this example to show that a continuous function need not be monotone on any interval. Prove this for his example, and for ours.

(d) Bolzano only showed that his example failed to be differentiable at a dense set of points (that is, for any point x and any $\varepsilon > 0$, there is a point of non-differentiability within ε of x).[25] Show

[25] The non-differentiability everywhere of this example was proved by Jarník in 1922; this was published in Czech, and translated into English in 1981. See [46] for bibliographic details.

that his example is not differentiable at any of the "break" points in any of the subdivisions, and that these are dense.

To make a start, out of particulars and make them general, rolling up the sum, by defective means...

William Carlos Williams
Patterson

...these were the days of Cavalieri and Roberval, when an area was looked upon as the same thing as an infinite number of line segments, a very helpful if dangerous definition.

J. L. Coolidge [16]

5 Integration

We have spent a good deal of time and energy studying differentiation; we now turn to the complementary notion of integration. Integration arises in connection with the calculation of areas, volumes, and length of curves, as well as averaging processes and related problems, particularly in probability and statistics.

The study of areas and volumes goes back to the Greeks; in the exercises for the first section of this chapter we will explore the results given by Euclid and Archimedes concerning the area of a circle. Both sources did much more, using versions of the method of exhaustion. For example, Archimedes found the area under a parabola, and volumes of cones, spheres, cylinders, and even paraboloids and hyperboloids. In the eleventh century the Islamic scholar Alhazen (*ca.*965-1039) extended some of these results, for example finding the volume obtained when a parabola is rotated about its base (rather than its axis). In the early seventeenth century Johannes Kepler (1571-1630) and Bonaventura Cavalieri (1598-1647) found formulas for other areas and volumes by heuristic methods. More general methods, closer to ones we now recognize as rigorous, were developed by Pierre de Fermat (1601-1665), Evangelista Torricelli (1608-1647), Gilles Personne de Roberval (1602-1675) and Blaise Pascal (1623-1662) in Europe, and John Wallis (1616-1703) and Isaac Barrow (1630-1677) in England.

Then, in the 1660's, the stunning result, known as the *Fundamental*

Theorem of Calculus, was discovered: the process of finding areas (integration) is the inverse of the operation of finding tangents (differentiation). This was observed in a special case by Torricelli [51, p. 253] and in more general terms (but in a geometric language that to a modern reader is rather obscure) by James Gregory (1638-1675)—a brilliant Scots mathematician whose early death diminished his historical influence—and Isaac Barrow—Newton's teacher and predecessor in the Lucasian Chair of Mathematics at Cambridge. The full power of this observation was harnessed by Newton and Leibniz (independently) and is at the heart of much of the power of the calculus.

After this breakthrough, calculus developed at a breathtaking pace; by 1700, many of the results and techniques of what we call first-year calculus (and much more) had already been developed by the English and Continental schools. The eighteenth century saw a further blossoming of the field, particularly through power series methods in the hands of the masters Leonard Euler (1701-1783) and Joseph Louis Lagrange (1736-1813). During this period, integration was seen (as it is by many beginning calculus students now) primarily as *antidifferentiation*, and its foundation in the theory of area was somewhat eclipsed.

Then in the early nineteenth century, in connection with problems concerning the wave equation and heat equation, the foundations of integration were re-examined. Augustin-Louis Cauchy (1789-1857) gave a careful formulation of the definite integral (making more rigorous Leibniz's view of a definite integral as a sum of infinitesimals) and used his formulation to discuss integrals in the complex domain, as well as convergence of series. Euler had done some investigating of approximation of integrals by finite sums, but for him this was a property of antiderivatives; for Cauchy it was the definition: in fact, the Fundamental Theorem of Calculus was obtained by him as a consequence of his definition. Cauchy dealt primarily with integrals of continuous functions, but the emerging study of Fourier series required going beyond the assumption of continuity. Cauchy's definition was further refined by Bernhard Georg Friedrich Riemann (1826-1866) in 1854 to a form close to what we use today, although the recognition that some of his results required the notion of uniform continuity really came later, from Karl Theodor Wilhelm Weierstrass (1815-1897). Riemann also gave necessary and sufficient conditions for his integral to exist for a given function; we explore this in the last section of this chapter.

The notion of the integral continued to be extended: Thomas-Jean Stieltjes (1856-1894) introduced the idea of integrating with respect to a mass distribution in 1894, and a revolutionary new theory of integration was produced by Henri Léon Lebesgue (1875-1941) in his doctoral thesis (1902); the description of the Lebesgue integral is beyond the scope of this book, but in its ability to handle really wild functions and its introduction of new ideas it has had a profound effect on the development of analysis in the last century; even more abstract and general formulations of the integral proliferated throughout that century.

5.1 Area and the Definition of the Integral

The idea of area, like that of velocity, begins with common-sense intuition but ends up requiring sophisticated limit processes to work in any but the simplest situations.

Our basic premise is that area is *additive*: if a region S consists of two pieces S_1 and S_2 which can share boundary points, but not interior points (*i.e.*, , they don't *overlap*) then the area $A(S)$ of the whole region should be the sum of the areas of the pieces:

$$A(S) = A(S_1) + A(S_2)\,. \tag{5.1}$$

We use this in two ways: *subdivision* and *tiling*. Starting with a standard square (sides of length 1) as our unit of area ($A = 1$), we can subdivide it into n^2 subsquares with sides of length $1/n$, so each of these subsquares must have area $1/n^2$. Now, a rectangle whose sides have rational length,[1] say width $w = p/n$ and length $\ell = q/n$, can be tiled by q rows consisting of p squares each with side $1/n$, so we obtain the area formula

$$A = \ell w$$

for (rational) rectangles; it is natural to extrapolate this formula to rectangles with rational or irrational sides.

It is also possible to use these kinds of subdivision and tiling arguments to deduce the area formula for triangles and other polygons. However, this approach collapses when we attempt to calculate the area of regions with curved sides.

To understand this more general situation, we investigate the area of the plane region (see Figure 5.1) S bounded below by the x-axis, on the left and right by two vertical lines ($x = a$ and $x = b$) and above by the graph

[1] with common denominator n

of a function $y = f(x)$, which we assume is non-negative and continuous on the interval $x \in [a, b]$:

$$S = \{(x, y) \,|\, a \leq x \leq b,\ 0 \leq y \leq f(x)\}.$$

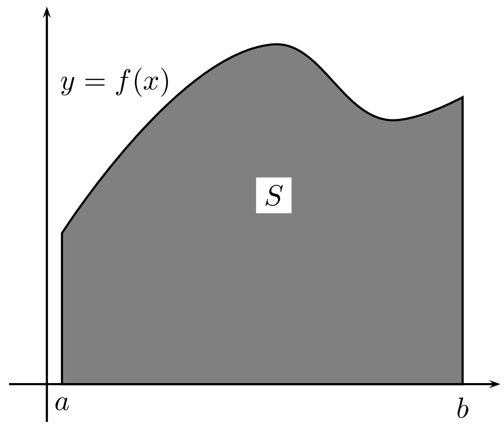


Figure 5.1: Area under a graph

We will illustrate our general discussion using the example

$$f(x) = x^3,\ a = 1,\ b = 2.$$

The idea is to locate the number $A(S)$ (giving the area of S) by a process of approximation analogous to our approximation of $\sqrt{2}$ in § 2.2. Our estimates use the *monotonicity* of area: the area of any subregion is less than the area of the whole:

$$S_1 \subset S_2 \Rightarrow A(S_1) \leq A(S_2)\,. \tag{5.2}$$

(This actually follows from additivity of area: see Exercise 3.)
To obtain *lower* bounds for $A(S)$, we fit non-overlapping rectangles *inside* S (Figure 5.3). Start by subdividing the interval $[a, b]$ into adjacent subintervals: this is specified by a finite list of division points, together with the two endpoints, numbered left-to-right and starting with index 0:

$$a = p_0 < p_1 < \ldots < p_n = b.$$

We refer to the collection

$$\mathcal{P} = \{p_0, p_1, \ldots, p_n\}$$

as a **partition** of $[a, b]$. The **component intervals** of $\mathcal{P}$

$$I_j = [p_{j-1}, p_j] \quad j = 1, \ldots, n$$

are also numbered left to right; the length of the j^{th} component interval is

$$\triangle x_j = \|I_j\| = p_j - p_{j-1} \quad j = 1, \ldots, n.$$

As an example, we consider, for $n = 2, 3, ...$, the partition $\mathcal{P}_n$ of the interval $[1, 2]$ into n intervals, *all of the same length*—that is, the points p_i, $i = 0, \ldots, n$ are equally spaced in $[1, 2]$, so

$$\triangle x_j = \frac{2-1}{n} = \frac{1}{n}$$

for $j = 1, \ldots, n$, and we easily see that

$$p_j = 1 + \frac{j}{n}, \quad i = 0, \ldots, n$$

or

$$\mathcal{P}_n = \{1, 1 + \frac{1}{n}, 1 + \frac{2}{n}, ..., 1 + \frac{n-1}{n}, 2\}.$$

Now, we use each component interval I_j (sitting on the x-axis) as the base of a rectangle of height h_j (and width $w_j = \triangle x_j$). This rectangle is

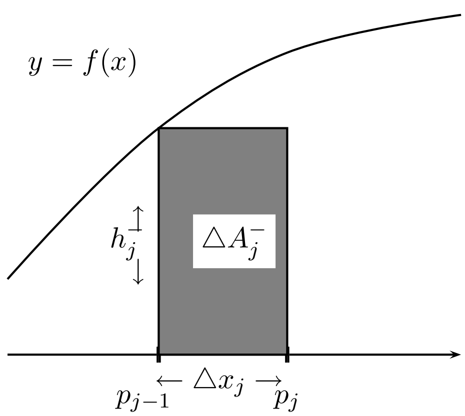


Figure 5.2: Inside Rectangle over $I_j = [p_{j-1}, p_j]$

contained inside S precisely if h_j is a lower bound for $f(x)$ on I_j, so the *tallest* rectangle with base I_j that fits inside S (Figure 5.2 has height

$$h_j = h_j^- := \inf_{x \in I_j} f(x)$$

and hence area

$$\triangle A_j^- = h_j^- w_j = \left[\inf_{x \in I_j} f(x)\right] \triangle x_j.$$

Thus, starting from the component intervals of the partition $\mathcal{P}$, the largest union of rectangles we can fit inside S (Figure 5.3) has total area

$$\sum_{j=1}^{n} \triangle A_j^- = \sum_{j=1}^{n} \left[\inf_{x \in I_j} f(x)\right] \triangle x_j.$$

The expression on the right is called the **lower sum** for f associated to the partition $\mathcal{P}$ of $[a, b]$; we write

$$\mathcal{L}(\mathcal{P}, f) = \sum_{j=1}^{n} \inf_{I_j} f \triangle x_j.$$

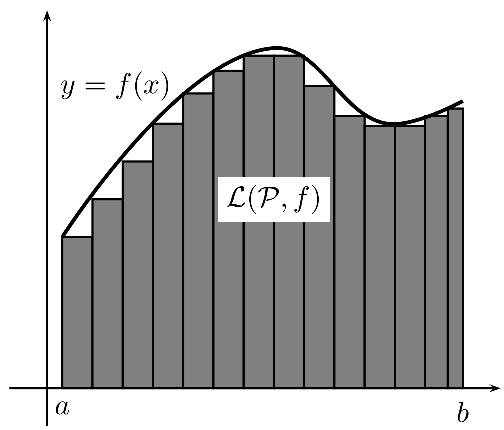


Figure 5.3: $\mathcal{L}(\mathcal{P}, f)$

Monotonicity of area then says that, for every partition $\mathcal{P}$,

$$\mathcal{L}(\mathcal{P}, f) \le A(S)\,.$$

In our example, $f(x) = x^3$ is strictly increasing over $[1, 2]$, so for each j,

$$h_j^- = \inf_{x \in I_j} x^3 = p_{j-1}^3$$
$$= \left(1 + \frac{j-1}{n}\right)^3$$

and the lower sum for x^3 associated to the partition $\mathcal{P}_n$ of $[1,2]$ into n equal parts is

$$\mathcal{L}(\mathcal{P}_n, x^3) = \sum_{j=1}^{n} \left[1 + \frac{(j-1)}{n}\right]^3 \cdot \frac{1}{n}$$
$$= \frac{1}{n} \sum_{i=j-1=0}^{n-1} \left[1 + \frac{i}{n}\right]^3$$
$$= \frac{1}{n} \sum_{i=0}^{n-1} \left[1 + 3\frac{i}{n} + 3\frac{i^2}{n^2} + \frac{i^3}{n^3}\right].$$

This formidable looking summation can be evaluated using the following summation formulas (Exercise 5)

$$\sum_{k=1}^{n} 1 = n \tag{5.3}$$
$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \tag{5.4}$$
$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \tag{5.5}$$
$$\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4} = \left(\sum_{k=1}^{n} k\right)^2. \tag{5.6}$$

Using these (and noting that we are applying them with n replaced by $n-1$ in all but the first sum, since the terms for $i=0$ drop out), we can write

$$\mathcal{L}(\mathcal{P}_n, x^3)$$
$$= \frac{1}{n}\left[n + \frac{3}{n}\frac{(n-1)(n)}{2} + \frac{3}{n^2}\frac{(n-1)(n)(2n-1)}{6} + \frac{1}{n^3}\frac{(n-1)^2(n)^2}{4}\right]$$
$$= 1 + \frac{3}{2}\left(\frac{n-1}{n}\right) + \frac{3}{6}\left(\frac{(n-1)(2n-1)}{n^2}\right) + \frac{1}{4}\left(\frac{(n-1)^2}{n^2}\right). \tag{5.7}$$

To obtain an *upper* bound for $A(S)$, we use a similar construction to enclose S inside a union of rectangles built on the same bases. Given $I_j = [p_{j-1}, p_j]$ a component of the partition $\mathcal{P} = \{p_0, \ldots, p_n\}$ of $[a,b]$, the rectangle with base I_j and height h_j contains all the points $(x,y) \in S$ with

$x \in I_j$ precisely if h_j is an upper bound for $f(x)$ on I_j; the *shortest* such rectangle has height[2]

$$h_j = h_j^+ :== \sup_{I_j} f.$$

This rectangle has area

$$\triangle A_j^+ = \sup_{I_j} f \triangle x_j.$$

The total area of these rectangles (Figure 5.4) gives an upper bound for $A(S)$,

$$A(S) \leq \sum_{j=1}^{n} \triangle A_j^+ = \sum_{j=1}^{n} \sup_{I_j} f \triangle x_j.$$

The sum on the right is called the **upper sum** for f associated to the partition $\mathcal{P}$, and denoted

$$\mathcal{U}(\mathcal{P}, f) = \sum_{j=1}^{n} \sup_{I_j} f \triangle x_j.$$

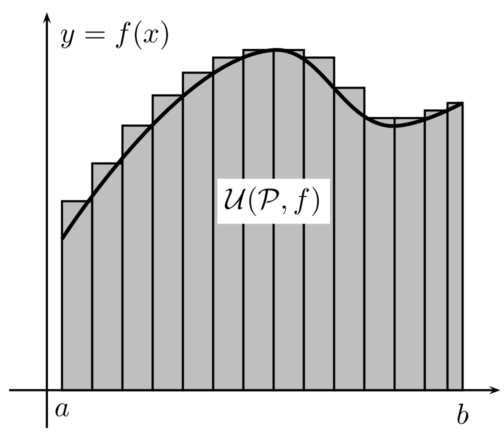


Figure 5.4: $\mathcal{U}(\mathcal{P}, f)$

In our example again, since $f(x) = x^3$ is strictly increasing on $[1, 2]$, we have

$$h_j^+ = \sup_{x \in I_j} x^3 = p_j^3$$

[2]The supremum here exists since f is continuous on I_j, hence bounded by the Extreme Value Theorem (Theorem 3.3.4)

so that the upper sum is

$$\begin{aligned}\mathcal{U}(\mathcal{P}_n, x^3) &= \frac{1}{n}\cdot\sum_{j=1}^{n}\left[1+3\frac{j}{n}+3\frac{j^2}{n^2}+\frac{j^3}{n^3}\right]\\ &= 1+\frac{3}{2}\left(\frac{n+1}{n}\right)+\frac{3}{6}\left(\frac{(n+1)(2n+1)}{n^2}\right)+\frac{1}{4}\left(\frac{(n+1)^2}{n^2}\right). \quad (5.8)\end{aligned}$$

We see, then, that the area $A(S)$ of the plane region

$$S = \{(x,y)\,|\,a \le x \le b,\ 0 \le y \le f(x)\}$$

is bounded by the lower and upper sums associated to any partition $\mathcal{P}$ of $[a,b]$:

$$\sum_{j=1}^{n}\inf_{I_j} f\triangle x_j = \mathcal{L}(\mathcal{P},f) \le A(S) \le \mathcal{U}(\mathcal{P},f) = \sum_{j=1}^{n}\sup_{I_j} f\triangle x_j.$$

We note in particular that the "trivial partition" $\mathcal{P}_0 = \{a,b\}$ has one component interval ($I_1 = [a,b]$) with $\triangle x_1 = b-a$, leading to the estimate

$$\inf_{a\le x\le b} f(x)\,(b-a) = \mathcal{L}(\mathcal{P}_0,f) \le A(S) \le \mathcal{U}(\mathcal{P}_0,f) = \sup_{a\le x\le b} f(x)\,(b-a).$$

Any other partition $\mathcal{P} = \{a = p_0 < p_1 < \ldots < p_n = b\}$ can be obtained from $\mathcal{P}_0$ by adding the new division points $p_1,\ldots,p_{n-1}$ to $\mathcal{P}_0$, or equivalently by subdividing the component $[a,b]$ of $\mathcal{P}_0$ into several adjacent intervals, to form the component intervals of $\mathcal{P}$. For each component interval I_j of $\mathcal{P}$, we clearly have

$$\inf_{I_j} f \ge \inf_{[a,b]} f \text{ and } \sup_{I_j} f \le \sup_{[a,b]} f$$

and, since

$$\sum_{j=1}^{n}\triangle x_j = b-a,$$

we can conclude that

$$\begin{aligned}\mathcal{L}(\mathcal{P},f) &= \sum_{j=1}^{n}\inf_{I_j} f\triangle x_j\\ &\ge \sum_{j=1}^{n}\inf_{[a,b]} f\triangle x_j = \inf_{I_j} f\sum_{j=1}^{n}\triangle x_j = (\inf_{I_j} f)(b-a)\\ &= \mathcal{L}(\mathcal{P}_0,f)\end{aligned}$$

and, similarly,

$$\mathcal{U}(\mathcal{P}, f) = \sum_{j=1}^{n} \sup_{I_j} f \triangle x_j \leq (\sup_{[a,b]} f)(b-a) = \mathcal{U}(\mathcal{P}_0, f).$$

An analogous argument (Exercise 6) shows that adding new division points to *any* partition $\mathcal{P}$ leads to a new partition $\mathcal{P}'$ whose lower and upper sums are bounded by those of $\mathcal{P}$. We call $\mathcal{P}'$ a **refinement** of $\mathcal{P}$ if, as sets of points,

$$\mathcal{P} \subset \mathcal{P}';$$

that is, every partition point of $\mathcal{P}$ is also a partition point of $\mathcal{P}'$ (or, in different terms, if every component interval of $\mathcal{P}'$ is entirely contained in a single component interval of $\mathcal{P}$). We have

Remark 5.1.1. *If $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then*

$$\mathcal{L}(\mathcal{P}, f) \leq \mathcal{L}(\mathcal{P}', f)$$

and

$$\mathcal{U}(\mathcal{P}, f) \geq \mathcal{U}(\mathcal{P}', f).$$

So far, we have been proceeding as if the area $A(S)$ were a well-defined number satisfying the estimates

$$\mathcal{L}(\mathcal{P}, f) \leq A(S) \leq \mathcal{U}(\mathcal{P}, f) \tag{5.9}$$

for every partition $\mathcal{P}$ of $[a, b]$. But all we *really* know are the lower and upper sums $\mathcal{L}(\mathcal{P}, f)$ and $\mathcal{U}(\mathcal{P}, f)$; we would like to *define* the area $A(S)$ as the number satisfying all the inequalities above. For such a definition to make sense, we need two conditions. First, there must be *at least one* number satisfying all of these inequalities. This boils down to the requirement that *every* lower sum is less than (or equal to) *every* upper sum, even for different partitions. When this holds, we are able to conclude that

$$\sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, f) \leq \inf_{\mathcal{P}} \mathcal{U}(\mathcal{P}, f),$$

(where the supremum and infimum are taken over all possible finite partitions $\mathcal{P}$ of $[a, b]$) and any number between these two automatically satisfies all the inequalities. But then, for these inequalities to actually *define* a number $A(S)$, we need to insure that *only one* number works—that is, we really need the *equality*

$$\sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, f) = \inf_{\mathcal{P}} \mathcal{U}(\mathcal{P}, f).$$

The first condition can be established using Remark 5.1.1. Given two partitions $\mathcal{P}_1$ and $\mathcal{P}_2$, we can form their **mutual refinement** $\mathcal{P}$ by taking all division points from either partition; that is, as sets of points,

$$\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2.$$

Now, $\mathcal{P}_i \subset \mathcal{P}$ ($\mathcal{P}$ is a refinement of $\mathcal{P}_i$) for $i = 1, 2$, and hence

$$\mathcal{L}(\mathcal{P}_i, f) \leq \mathcal{L}(\mathcal{P}, f) \leq \mathcal{U}(\mathcal{P}, f) \leq \mathcal{U}(\mathcal{P}_i, f), \quad i = 1, 2. \tag{5.10}$$

In particular, for arbitrary partitions $\mathcal{P}_1$ and $\mathcal{P}_2$,

$$\mathcal{L}(\mathcal{P}_1, f) \leq \mathcal{U}(\mathcal{P}_2, f).$$

But then the upper sum of *any* partition is an upper bound for *all* lower sums, and hence for their least upper bound

$$\sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, f) \leq \mathcal{U}(\mathcal{P}_2, f),$$

and the *left* side is a lower bound for *all* upper sums, hence for their greatest lower bound:

$$\sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, f) \leq \inf_{\mathcal{P}} \mathcal{U}(\mathcal{P}, f).$$

We state this formally, for future use:

Remark 5.1.2. *For any two partitions $\mathcal{P}_1$ and $\mathcal{P}_2$,*

$$\mathcal{L}(\mathcal{P}_1, f) \leq \mathcal{U}(\mathcal{P}_2, f)$$

and hence

$$\sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, f) \leq \inf_{\mathcal{P}} \mathcal{U}(\mathcal{P}, f).$$

So there is always *some* number bounded below by all lower sums and above by all upper sums. Its uniqueness takes further work. Before proceeding, we will enlarge our context. Note that in the construction of lower and upper sums, as well as in the proofs of the various inequalities so far, we have used the assumption that f is continuous *only* to establish that f is bounded on $[a, b]$ (so that the numbers involved in defining $\mathcal{L}(\mathcal{P}, f)$ and $\mathcal{U}(\mathcal{P}, f)$ are all well-defined). Go back over these arguments to convince yourself that everything so far works if we just assume that f is *bounded* (but not necessarily continuous) on $[a, b]$. That is, if there exists $M \in \mathbb{R}$ such that

$$0 \leq f(x) \leq M$$

for all $x \in [a, b]$, then for every partition $\mathcal{P}$ of $[a, b]$ we can define the upper and lower sums

$$\mathcal{L}(\mathcal{P}, f) = \sum_{j=1}^{n} \inf_{I_j} f \triangle x_j$$
$$\mathcal{U}(\mathcal{P}, f) = \sum_{j=1}^{n} \sup_{I_j} f \triangle x_j$$

and these satisfy the inequality

$$\mathcal{L}(\mathcal{P}_1, f) \leq \mathcal{U}(\mathcal{P}_2, f)$$

for any pair of partitions $\mathcal{P}_1$ and $\mathcal{P}_2$. It follows that

$$\sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, f) \leq \inf_{\mathcal{P}} \mathcal{U}(\mathcal{P}, f).$$

So for any bounded non-negative function f, we will say that the set of points in the plane

$$S = \{(x, y) \,|\, a \leq x \leq b,\ 0 \leq y \leq f(x)\}$$

has a well-defined area if

$$\sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, f) = \inf_{\mathcal{P}} \mathcal{U}(\mathcal{P}, f)$$

—that is, only one number can satisfy Equation (5.9)—and in that case the **area** $A(S)$ is precisely this (well-defined) number.
In general, it is a daunting task to try to either maximize (*resp.* minimize) the lower (*resp.* upper) sums over *all* partitions. In our example, however, we can get at this indirectly. You can verify directly from Equation (5.7) and Equation (5.8) that

$$\lim_{n\to\infty} \mathcal{L}(\mathcal{P}_n, x^3) = \lim_{n\to\infty} \mathcal{U}(\mathcal{P}_n, x^3) = 1 + \frac{3}{2} + \frac{6}{6} + \frac{1}{4} = 3\frac{3}{4}.$$

But since we know that *every* lower sum is less than *every* upper sum, we see that

$$3\frac{3}{4} = \lim_{n\to\infty} \mathcal{L}(\mathcal{P}_n, x^3) \leq \sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, x^3) \leq \inf_{\mathcal{P}} \mathcal{U}(\mathcal{P}, x^3) \leq \lim_{n\to\infty} \mathcal{U}(\mathcal{P}_n, x^3) = 3\frac{3}{4}$$

and so the area of $\{(x,y) \,|\, 0 \leq y \leq x^3,\ 1 \leq x \leq 2\}$ is well-defined and equals $3\frac{3}{4}$.

We note briefly that once we drop continuity of f, the area of S *need not* be well-defined. For example, the function

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

has, for *every* nontrivial interval I_j,

$$\inf_{I_j} f = 0, \qquad \sup_{I_j} f = 1$$

so that for *any* finite partition $\mathcal{P}$ of a non-trivial interval $[a,b]$ $(a < b)$,

$$\mathcal{L}(\mathcal{P}, f) = 0 < b - a = \mathcal{U}(\mathcal{P}, f)$$

and hence the area under the graph of f is *not* well defined in the sense above[3]. We shall return to this phenomenon later.

Actually, we can extend the context further. Our calculations were motivated by the problem of finding the area of the region S below the graph of f, and we needed f to be non-negative on $[a,b]$ to even define S. But the calculations *themselves* make formal sense even when f takes negative values, and lead to a useful object.

So we formulate the more general concept of the **definite integral** of f over $[a,b]$ as follows:

Definition 5.1.3. *Suppose f is bounded on the closed interval $[a,b]$. For every partition of $[a,b]$*

$$\mathcal{P} = \{a = p_0 < p_1 < \ldots < p_n = b\}$$

[3]The Lebesgue theory does handle this example.

with component intervals

$$I_j = [p_{j-1}, p_j] \quad j = 1, \ldots, n$$

of length $\|I_j\| = p_j - p_{j-1} = \triangle x_j$*,*
define the ***lower sum*** *(resp.* ***upper sum****) associated to* $\mathcal{P}$

$$\mathcal{L}(\mathcal{P}, f) = \sum_{j=1}^{n} \inf_{I_j} f \triangle x_j$$
$$\mathcal{U}(\mathcal{P}, f) = \sum_{j=1}^{n} \sup_{I_j} f \triangle x_j.$$

Then f *is* ***(Riemann) integrable*** *over* $[a, b]$ *if*

$$\sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, f) = \inf_{\mathcal{P}} \mathcal{U}(\mathcal{P}, f)$$

(the supremum and infimum are taken over all finite partitions $\mathcal{P}$ *of* $[a, b]$*).*
In this case, we call the number above the ***definite (Riemann) integral*** *of* f *over* $[a, b]$[4]*, denoted*

$$\int_a^b f(x)\ dx = \sup_{\mathcal{P}} \mathcal{L}(\mathcal{P}, f) = \inf_{\mathcal{P}} \mathcal{U}(\mathcal{P}, f).$$

This is, as it stands, a highly abstract definition, and with few exceptions it would be foolhardy to try to calculate the integral (or even verify integrability) directly from the definition. In the next section, we will continue in an abstract vein to deduce some consequences of this definition which will help us better handle the concept. In particular, we will find a large class of functions which are guaranteed to be integrable. Then, in § 5.3, we will establish the Fundamental Theorem of Calculus, which leads to a practical way of calculating the integrals of many functions, a process we will explore at length in § 5.4 and § 5.5.

Exercises for § 5.1

Answers to Exercises 1, 2ac are given in Appendix B.

Practice problems:

[4] or "from a to b"

1. Let $S = \{(x, y) \mid 1 \leq x \leq 2,\ 0 \leq y \leq x^2\}$ be the region bounded below by the x-axis, above by the graph $y = x^2$ of the function $f(x) = x^2$ and on the sides by the lines $x = 1$ and $x = 2$. Let $\mathcal{P}_1 = \{1, \frac{3}{2}, 2\}$ be the partition of $[1, 2]$ into two equal parts, and $\mathcal{P}_2 = \{1, \frac{4}{3}, \frac{5}{3}, 2\}$ the partition of $[1, 2]$ into three equal parts.

   (a) Calculate the upper and lower sums $\mathcal{L}(\mathcal{P}_1, f)$ and $\mathcal{U}(\mathcal{P}_1, f)$. (*Hint:* The function f is strictly increasing on $[1, 2]$.)

   (b) Calculate the upper and lower sums $\mathcal{L}(\mathcal{P}_2, f)$ and $\mathcal{U}(\mathcal{P}_2, f)$.

   (c) Let $\mathcal{P}$ be the mutual refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Calculate $\mathcal{L}(\mathcal{P}, f)$ and $\mathcal{U}(\mathcal{P}, f)$ and verify Equation (5.10) in this case.

   (d) Let $\mathcal{P}' = \{1, 1.2, 1.6, 2\}$. Calculate the upper and lower sums $\mathcal{L}(\mathcal{P}', f)$ and $\mathcal{U}(\mathcal{P}', f)$

2. For each function f and interval $[a, b]$ below, (i) write down a formula for the points x_i of the partition $\mathcal{P}_n$ of $[a, b]$ into n equal pieces; (ii) write an expression for $\mathcal{L}(\mathcal{P}_n, f)$ and $\mathcal{U}(\mathcal{P}_n, f)$ as an explicit sum; (iii) express each sum as a combination of sums of the form $A(n) + B(n)\sum_{i=1}^{n} i + C(n)\sum_{i=1}^{n} i^2$; (iv) use Equations (5.3)-(5.5) to write these as functions of n alone; (v) find the limits of these as $n \to \infty$ (*Hint:* Your answer for lower and upper sums should be the same; then the argument we used in the text for $f(x) = x^3$ shows that this limit equals the definite integral $\int_a^b f(x)\,dx$.)

   (a) $f(x) = x, [a, b] = [0, 1]$

   (b) $f(x) = x^2 + 1, [a, b] = [0, 1]$

   (c) $f(x) = 2x^2 - 1, [a, b] = [-2, 0]$

   (d) $f(x) = 1 - x^2, [a, b] = [-2, 2]$ (*Hint:* Consider the sums over $[-2, 0]$ and $[0, 2]$ separately.)

Theory problems:

3. Suppose we know that the area of a union of disjoint regions equals the sum of the areas of the pieces (Equation (5.1)). Use this to show that the area of a subregion is at most equal to the area of the whole region (Equation (5.2)).

4. Starting with the assumption that the area of a square of side 1 inch is 1 square inch, and the additive property of area (Equation (5.1)), derive

   (a) The formula for the area of a rectangle with length and width an *integer* number of inches;

   (b) the formula for the area of a rectangle with length and width a *rational* number of inches;

   (c) the formula for the area of a triangle with rational base and height;

   (d) the formula for the area of a trapezoid with both parallel sides and height a rational number of inches.

5. Use induction (see Appendix A, or p. 31) to prove formulas (5.3)-(5.6).

6. Mimic the argument on p. 323 to prove Remark 5.1.1.

History notes:

Areas of circles in Euclid: The next three exercises take you through Euclid's proof (*Elements* (ca 300 BC)[32, Book XII, Prop. 2] of the following: *(The areas of) two circles are to each other as the (areas of the) squares on their diameters.* In our terminology, if C_1 and C_2 are circles with respective diameters d_1 and d_2, then

$$\frac{A(C_1)}{A(C_2)} = \frac{d_1^2}{d_2^2}.$$

(Here we are assuming the formula $A = d^2$ for the area of a square with side d.) The discovery of the ratio statement is probably due to Hippocrates of Chios (460-380 BC), who gave a different argument from Euclid.

7. (a) Prove (*Elements*, Book XII, Prop. 1): *(The areas of) similar polygons inscribed in circles are to each other as the squares of the diameters.* (*Hint:* Use Exercise 4, together with the fact that every polygon can be tiled into triangles.)

   (b) (This is not explicit in Euclid, but is an interesting exercise in his style.) Suppose P is a regular polygon (*i.e.*, all sides have the same length, s). Then the perpendicular bisectors of all the sides cross at a single point, which we can refer to as the *center* of the polygon. (*Hint:* If P is inscribed in a circle, this point is the center of the circle.) Let h be the distance from the center to (the midpoint of) any side. Show that the area of P equals half of the circumference of P times h.

8. Let P_n be a regular 2^n-sided polygon inscribed in a circle (so P_2 is a square, P_3 an octagon, and so on). Let Δ_n be the difference between the area enclosed by the circle and that enclosed by P_n. Clearly $\{\Delta_n\}$ form a nonincreasing sequence. Show that for each $n = 2, ...$

$$\Delta_n - \Delta_{n+1} > \frac{1}{2}\Delta_n. \tag{5.11}$$

(*Hint:* Figure 5.5 shows the detail near one edge AB of P_n; the corresponding edges of P_{n+1} are AC and CB. So Δ_n consists of 2^n copies of the area D_n bounded by the chord AB and the circular arc $\widehat{ACB}$, while Δ_{n+1} consists of 2^n copies of the union D_{n+1} of the two shaded areas. In particular, $\Delta_n - \Delta_{n+1}$ is 2^n times the area of the triangle ABC, which is half the area of the rectangle $ABED$, and this rectangle in turn contains D_n.)

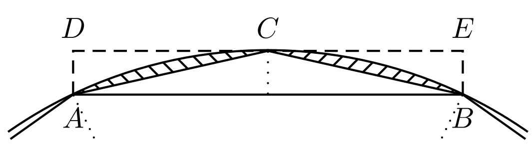


Figure 5.5: Equation (5.11)

9. Use Exercise 8 together with Exercise 34 in § 2.4 to show that if C is a circle, then for any number $A < A(C)$ we can find a regular inscribed 2^n-gon P_n such that $A(P_n) > A$.

10. Now we prove the main result. Suppose it is false; that is, suppose we have two circles C_i, $i = 1, 2$, such that $\frac{A(C_1)}{A(C_2)} \neq \frac{d_1^2}{d_2^2}$. We will derive a contradiction from this supposition.

    Pick numbers A_1 and A_2 so that

$$\frac{A_1}{A(C_2)} = \frac{d_1^2}{d_2^2} = \frac{A(C_1)}{A_2}.$$

Note that if $\frac{A(C_1)}{A(C_2)} > \frac{d_1^2}{d_2^2}$ then $A_1 < A(C_1)$ while if $\frac{A(C_1)}{A(C_2)} < \frac{d_1^2}{d_2^2}$ then $A_2 < A(C_2)$.

Suppose $A_1 < A(C_1)$. Find an inscribed regular 2^n-gon in C_1 such that

$$A(P_n) > A_1.$$

(Why can we do this?) Now consider the corresponding 2^n-gon P'_n inscribed in C_2. Show that

$$\frac{A(P_n)}{A(P'_n)} = \frac{A_1}{A(C_2)}$$

or

$$\frac{A(P'_n)}{A(C_2)} = \frac{A(P_n)}{A_1}.$$

Finally, show that the left side must be less than 1 while the right side must be more than 1.

How does this argument change if $\frac{A(C_1)}{A(C_2)} < \frac{d_1^2}{d_2^2}$?

Areas of circles in Archimedes: Euclid's Prop. 2 in Book XII shows, in our notation, that the area of a circle is given by a formula of the form

$$A = Kd^2$$

where d is the diameter of the circle and K is some constant; of course we can replace the diameter d in this formula with the radius r (or any other quantity which is proportional to the diameter), provided we adjust the value of the constant K. However, note that Euclid's result does *not* allow us to determine the value of K in any of these cases. A generation or so later than Euclid, the great Archimedes of Syracuse (*ca.* 212-287 BC) established the value of this constant, by means of two results, contained in his *Measurement of a Circle*[2] . We consider these in the following three exercises.

11. Archimedes assumes without proof that if P and P' are polygons respectively inscribed in and circumscribed about a circle S, then the circumference C of S lies between the circumference of P and that of P'. (See Exercises 15-18 for a (modern) proof of this.)

    Combine this with Exercise 9 to show that for any numbers $C_1 < C < C_2$, we can find regular 2^n-gons P_1 inscribed in (*resp.* P_2

circumscribed about) S so that the circumference of P_1 exceeds C_1, and the circumference of P_2 is less than C_2. (Note that $P_1 < P_2$ for example, by Exercise 7.)

12. Proposition 1 of the *Measurement of the Circle* is Archimedes' formula for the area of a circle:

    > *Proposition: Given a circle S, let K be a right triangle, one leg of which equals the circumference C of the circle and the other its radius, r. Then (the area of) S equals (the area of) K.*

    Verify that this agrees with our usual formula for the area of a circle. (By definition, π is the ratio between the circumference and the diameter of a circle.)

    Prove the proposition as follows:

    (a) Suppose the area of S is greater than that of K. Let Δ be the excess of the area of S over that of K. By Exercise 9, we can find a regular 2^n-gon P inscribed in S whose area differs from that of S by less than Δ. Show that the area of P is greater than that of K. However, using Exercise 7, and noting that $h < r$ and the circumference of P is less than C, show that the area of P is *less* than that of K. This is a contradiction.

    (b) Suppose the area of S is less than that of K. Let Δ be the excess of the area of K over that of S. Give a similar argument to the one above, to obtain a contradiction again. Since the area of S is neither greater than nor less than that of K, the areas are equal.

13. The rest of Archimedes' *Measurement of a Circle* is devoted to estimating the value of π—that is, the ratio between the circumference and the radius of a circle. His proof, based on a combination of estimates and ratio arguments, takes six pages (*cf* [31, pp. 93-98]); the crux of the argument can be stated more concisely using modern notation and algebra [20, 33-35] as follows. Archimedes will estimate the circumference of a circumscribed 96-gon P_{96} and an inscribed 96-gon P'_{96} [5].

[5] Our notation here differs a bit from that of the previous problems, as 96 is not a power of 2.

(a) For the circumference of P_{96}, Archimedes makes repeated use of the following [32, vol. 2, p. 195][6]

> **Lemma** *(Euclid, Book VI, Prop. 3): If an angle of a triangle be bisected and the straight line cutting the angle cut the base also, the segments of the base will have the same ratio as the remaining sides of the triangle.*

In terms of Figure 5.6, this says that

$$\angle COD = \angle DOA \Rightarrow CD : DA = CO : AO.$$

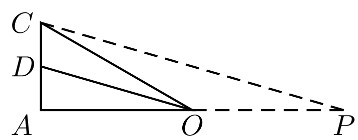


Figure 5.6: The Lemma

To use this lemma, imagine that O is the center of a circle with radius AO, which we can conveniently take to be of length 1, and the vertical line segment AC is tangent to the circle at A. Imagine also that AC is half a side of a regular circumscribed n-gon; denote the length of AC by t_n (so this is *half* the length of a side of P). If the angle $\angle COA$ is bisected by the line segment OD, then AD is half the side of a circumscribed regular $2n$-gon, with length t_{2n}.

Now, suppose P is picked so that $\angle AOC = 2\angle APC$. Then you can check that CP is parallel to DO, and that $CO{=}OP$. Similarity of the triangles $\triangle ADO$ and $\triangle ACP$ says in particular that

$$\frac{AD}{AO} = \frac{AC}{AO+OP} = \frac{AC}{AO+OC}.$$

But the first fraction equals t_{2n}, and writing the last fraction in terms of t_n we get the recurrence relation

$$t_{2n} = \frac{t_n}{1+\sqrt{1+t_n^2}}.$$

[6]We quote only the first half of the proposition, as this is all that Archimedes uses here

If we start with $n = 6$, then $AD = \tan \frac{\pi}{6} = \frac{1}{\sqrt{3}}$. Archimedes starts with the estimate

$$\sqrt{3} > \frac{265}{153}$$

and iterates the relation above, carefully rounding up at each step,[7] to get that half the circumference of $P_{96=2^4 \cdot 6}$ as at most $3\frac{1}{7}$ so

$$3\frac{1}{7} > 96t_{96} > \pi.$$

(Note that π is the ratio of the circumference to the *diameter*, and the latter in our case is 2.) For purposes of carrying out this exercise, you should simply calculate (electronically!) the values t_6, t_{12}, t_{24}, t_{48} and t_{96}, and then check the inequality above.

(b) For the circumference of P'_{96}, he uses the fact (Euclid, Book III, Prop. 31) that the triangle formed by a diameter of a circle together with the two lines joining another point on the circle is a right triangle with the diameter of the circle as hypotenuse. In particular, in Figure 5.7, the angles $\angle BDA$ and $\angle BCA$ are both right angles.

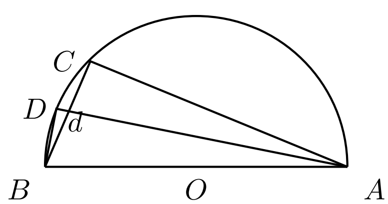


Figure 5.7: Inscribed polygons

Let us suppose that $OB = OA = 1$ in Figure 5.7 and that BC is a side of P'_n, of length s_n, and take BD to be a side of P'_{2n}, of length s_{2n}. The triangles $\triangle BdD$ and $\triangle AdC$ are similar, and under the assumption above AD bisects $\angle BAC$ (right?), so also $\triangle ABD$ is similar to the other two (why?). From these

[7] Archimedes works directly with the ratios that lead to this recurrence. Details are in [2, pp. 93-98, lxxx-xc]

similarities we have

$$\frac{AB}{AD} = \frac{Bd}{BD}$$
$$\frac{AC}{AD} = \frac{dC}{BD}$$

and hence

$$\frac{AB + AC}{AD} = \frac{Bd + dC}{BD} = \frac{BC}{BD}.$$

Rewritten in terms of s_n and s_{2n}, this reads

$$\frac{2 + \sqrt{4 - s_n^2}}{\sqrt{4 - s_{2n}^2}} = \frac{s_n}{s_{2n}}$$

which after some algebra becomes

$$s_{2n}^2 = \frac{s_n^2}{2 + \sqrt{4 - s_n^2}}.$$

Again, Archimedes starts with

$$s_6 = 1$$

and iterates the relation above,[8] rounding down and using the estimate

$$\sqrt{3} < \frac{1351}{780}.$$

This leads him to the estimate (since s_n is the side of P', we need to divide the circumference by 2)

$$\pi > 48 s_{96} > 3\frac{10}{71}.$$

(Again, for purposes of carrying out this exercise, calculate s_j for $j = 6, 12, 24, 48, 96$ and check the inequality above.)

(c) Check how closely these estimates find π (using a calculator).

[8] Again, Archimedes works directly with the ratios that lead to this recurrence; see [2, pp. 93-98, lxxx-xc]

Challenge Problems:

Length of Convex Curves: Archimedes' calculation of the area of a circle, as well as Newton's proof of the equation $\lim_{\theta\to 0}(\sin\theta)/\theta = 1$, depend on an assumption that Archimedes makes explicit, not in his *Measurement of a Circle* [2], but in another treatise, *On the Sphere and Cylinder* . We reproduce the relevant statements below[3, pp.2-4].
First, Archimedes defines what he means by a **convex curve** [3, p.2]:

> *1. There are in a plane certain terminated bent lines, which either lie wholly on the same side of the straight lines joining their extremities, or have no part of them on the other side.*
>
> *2. I apply the term **concave in the same direction** to a line such that, if any two points on it are taken, either all the straight lines connecting the points fall on the same side of the line, of some fall on one and the same side while others fall on the line itself, but none on the other side.*

14. We shall deal with these definitions when the curves are graphs of functions defined on a closed interval $[a, b]$ First, we show that if $f(x)$ is twice-differentiable on $[a, b]$, then convexity is the same as saying that the function does not switch concavity on $[a, b]$, and this implies that the graph is on one side of each of its tangent lines :
    (a) **Show** that if for all $a < c \leq b$ the graph $y = f(x)$ is (strictly) above the line L joining $(a, f(a))$ to $(c, f(c))$, then there exists a sequence of points $x_k \downarrow a$ for which $f'(x_k) < f'(a)$; in particular, $f''(a) \leq 0$. Replacing a with any $a' \in [a, b)$, conclude that if the graph is convex on this side, then $f(x)$ is concave down on $[a, b]$.
    (b) **Show** that if $a < x_1 < x_2 < b$ are points such that $(x_1, f(x_1))$ lies on *one* side of the line L and $(x_2, f(x_2))$ lies on the *other*, then $f(x)$ has an inflection point between x_1 and x_2. Thus, if $f(x)$ does not switch concavity on $[a, b]$ then its graph is convex.

After some other definitions, Archimedes states several assumptions, of which the first two interest us [3, pp.3-4]:

> *1. Of all lines which have the same extremities the straight line is the least.*
>
> *2. Of other lines in a plane and having the same extremities, [any two] such are unequal whenever both are concave in the*

> *same direction and one of them is either wholly included between the other and the straight line which has the same extremities with it, or is partly included by, and is partly common with, the other; and that [line] which is included is the lesser [of the two].*

The following sequence of exercises takes you through a proof of this, in modern terms. We shall continue to assume that the curves in question are the graphs of functions on a closed interval $[a, b]$.

15. **Length of a Convex Curve:** First, we need to *define* what we mean by the **length** of a convex curve. We shall deal with the case that $f(x)$ is concave down ($f' \downarrow$) on $[a, b]$; the upward case is analogous. A *polygonal curve* will be the graph of a continuous function, defined in pieces, such that its graph over each piece is a straight line segment. Note that specifying the endpoints of the line segments specifies the polygonal curve. The length of a polygonal curve is, naturally, the sum of the lengths of its line segments. We will say that the polygonal curve is *inscribed* in the graph of $y = f(x)$ if the endpoints of each line segment lie on the graph: thus it consists of a sequence of secant lines. Our intuition is that an inscribed polygonal curve should be shorter than the graph $y = f(x)$; we shall **define** the length of the graph $y = f(x)$ over $[a, b]$ (which we denote $\mathcal{C}$) by[9]

    $$\ell(\mathcal{C}) := \sup_{\text{polygonal curves } P \text{ inscribed in } \mathcal{C}} \ell(P).$$

    We need to show that the lengths of the inscribed polygonal curves are bounded above, so that their supremum is finite. To this end, note that (i) an inscribed polygonal curve is itself convex, and (ii) the function $f(x)$, as well as any function whose graph is an inscribed polygonal curve, achieves its maximum either at a unique point, or along a horizontal line segment.

    See Figure 5.8. Consider the curve Γ consisting of a vertical line from $(a, f(a))$ to $(a, \max f(x))$, a horizontal line to $(b, \max f(x))$, and a vertical line down to $(b, f(b))$. (If the maximum is achieved at an endpoint, then one of the vertical "lines" is a single point.)

    Show that any inscribed polygonal line is no longer than Γ.

[9] This definition works for many non-convex curves as well, but we will avoid the technical difficulties this entails.

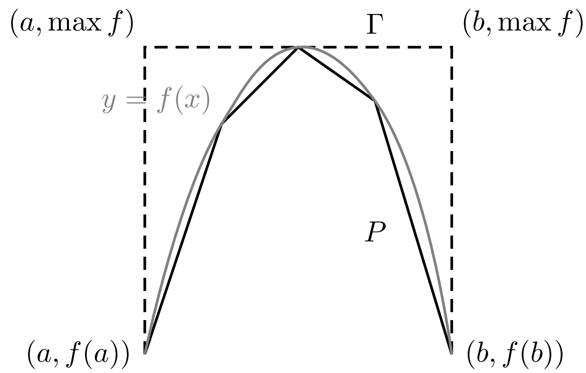


Figure 5.8: Γ

16. Given a convex polygonal curve, use the triangle inequality to **show** that if two points lie on adjacent segments, the curve obtained by replacing the piece of the curve joining them with a single straight line segment (*i.e.*, the secant line between the points) is convex and shorter. By induction on the number of interface points between, show that replacing the part of the curve between any two points with the secant joining them results in a shorter convex curve. See Figure 5.9.

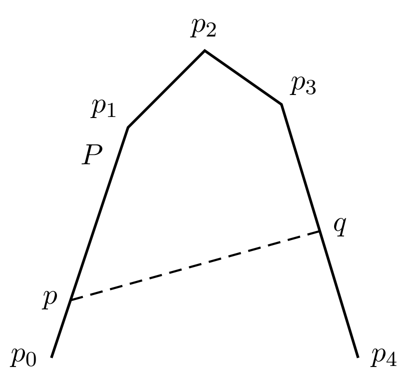


Figure 5.9: Cutting a curve with a secant.

17. Now suppose P_0 and P_1 are two convex polygonal curves with the

same endpoints, and that all points of P_1 are above or on P_0. Show that P_0 is shorter than P_1 as follows: Consider the first line segment of P_1; if it is not a subset of the first segment of P_2, extend the first segment of P_0 until it hits P_1, and replace the part of P_1 between these points with this extension (see Figure 5.10). This results in a new convex polygonal curve P_2, still lying above P_0, which agrees with P_0 over its first segment. Proceeding inductively, if P_0 has at least k segments and P_k agrees with P_0 over its first $k-1$ segments, then extend the k^{th} segment of P_0 until it hits P_k to get a new convex polygonal curve P_{k+1} which agrees with P_0 over its first k segments. **Show** that P_{k+1} is shorter than P_k for $k = 1, ..., n-1$, where P_0 has n segments, and P_n is P_0.

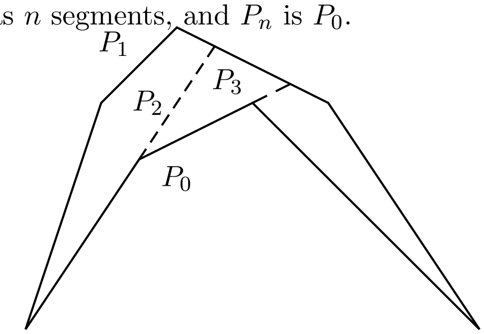


Figure 5.10: Comparing polygonal paths

18. Finally suppose $f(x)$ and $g(x)$ are both concave down over $[a, b]$, $f(a) = g(a)$ and $f(b) = g(b)$, and $f(x) > g(x)$ for $a < x < b$. **Show** that, given any polygonal curve inscribed in the graph of $g(x)$, there is a polygonal curve inscribed in the graph of $f(x)$ which lies above it. (*Hint:* If both endpoints of a line segment lie above a given line, then the whole segment lies above the line.)

    Use this to show that Archimedes' assumption holds.

19. (a) Show that if the line segment L from (x_1, y_1) to (x_2, y_2) has slope m, then
$$\ell(L) = \sqrt{1 + m^2}\triangle x$$
where $\triangle x = |x_1 - x_2|$.

    (b) Show that, if $\mathcal{C}$ is the graph over $[a, b]$ of a function $f(x)$ for which $\sqrt{1+(f'(x))^2}$ is integrable, then

$$\ell\,(\mathcal{C}) = \int_a^b \sqrt{1+(f'(x))^2}\,dx.$$

5.2 General Theory of the Riemann Integral

In this section, we shall look at general properties of the integral, including results that give us a long list of functions guaranteed to be integrable.
First, let us try to understand the integrability condition in Definition 5.1.3 better. We saw in § 2.3 that if a collection of numbers is bounded above (*resp.* below) by β (*resp.* by α) then its sup (*resp.* inf) exists, and equals β (*resp.* α) precisely if there is a sequence in the collection that converges to β (*resp.* α).
In our case, we can take the set to be $\mathcal{L} \subset \mathbb{R}$, the collection of all the possible values of $\mathcal{L}(\mathcal{P}, f)$ for different partitions of $[a, b]$. We know $\mathcal{L}$ is bounded above (for example by $\mathcal{U}(\mathcal{P}_0, f)$). So $\mathcal{L}$ has a supremum, which is the limit of some sequence in $\mathcal{L}$—that is, abstractly, there exists some sequence of partitions $\mathcal{P}_i^L$ whose corresponding lower sums converge to the supremum:

$$\mathcal{L}(\mathcal{P}_i^L, f) \to \sup \mathcal{L}.$$

Analogously, the set $\mathcal{U}$ of upper sums is bounded below (*e.g.*, by $\mathcal{L}(\mathcal{P}_0, f)$), and so some sequence of partitions $\mathcal{P}_i^U$ satisfies

$$\mathcal{U}(\mathcal{P}_i^U, f) \to \inf \mathcal{U}.$$

Using Remark 5.1.1, we see that refining any of these partitions does not change the convergence statements, so taking $\mathcal{P}_i$ the mutual refinement of $\mathcal{P}_i^L$ and $\mathcal{P}_i^U$ we have a sequence of partitions for which

$$\mathcal{U}(\mathcal{P}_i, f) - \mathcal{L}(\mathcal{P}_i, f) \to \inf \mathcal{U} - \sup \mathcal{L}.$$

By definition, f is integrable precisely if the right side of this is zero, so we have the following condition to characterize integrable functions:

Remark 5.2.1. *f is integrable over $[a, b]$ precisely if there exists a sequence of partitions $\mathcal{P}_i$ of $[a, b]$ for which the* ***convergence condition***

$$\mathcal{U}(\mathcal{P}_i, f) - \mathcal{L}(\mathcal{P}_i, f) \to 0$$

holds.

In this case,

$$\int_a^b f(x)\ dx = \lim \mathcal{U}(\mathcal{P}_i, f) = \lim \mathcal{L}(\mathcal{P}_i, f).$$

An easy consequence of this observation is the integrability of monotone functions.

Proposition 5.2.2. *Suppose f is monotone on $[a, b]$. Then f is integrable on $[a, b]$.*

Proof. We will prove the result assuming f is increasing on $[a, b]$. Let $\mathcal{P}_n$ be the partition of $[a, b]$ into n intervals of *equal* length: for $j = 0, \ldots, n$,

$$p_j = a + j\triangle x$$

where

$$\triangle x = \frac{b-a}{n} = \triangle x_j, \quad j = 1, \ldots, n.$$

Since f is increasing on each component $I_j = [p_{j-1}, p_j]$, we have

$$\inf_{I_j} f = \min_{I_j} f = f(p_{j-1})$$
$$\sup_{I_j} f = \max_{I_j} f = f(p_j)\,.$$

Then

$$\begin{aligned} \mathcal{U}(\mathcal{P}_n, f) - \mathcal{L}(\mathcal{P}_n, f) &= \sum_{j=1}^{n} (f(p_j) - f(p_{j-1}))\triangle x_j \\ &= [(f(p_1) - f(p_0)) + (f(p_2) - f(p_1)) + \ldots \\ &\ldots + (f(p_n) - f(p_{n-1}))]\,\triangle x. \end{aligned}$$

Note that the sum in brackets is a *telescoping sum*: everything cancels except

$$f(p_n) - f(p_0) = f(b) - f(a)\,.$$

Thus, we have the convergence condition for $\mathcal{P}_n$

$$\begin{aligned} \mathcal{U}(\mathcal{P}_n, f) - \mathcal{L}(\mathcal{P}_n, f) &= [f(b) - f(a)]\,\triangle x \\ &= [f(b) - f(a)]\,\frac{(b-a)}{n} \to 0 \end{aligned}$$

and f is integrable by Remark 5.2.1.
(For a geometric interpretation of this, see Figure 5.11.)

$\square$

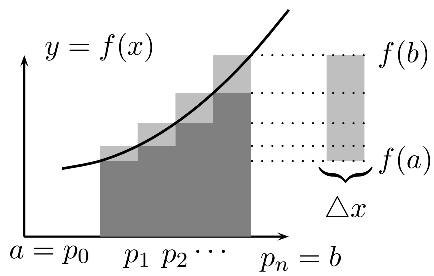


Figure 5.11: Telescoping sum

For general continuous functions, the argument is a bit more subtle, involving the notion of *uniform continuity* (see § 3.7). It turns out that, when f is continuous, it is fairly easy to find a sequence $\mathcal{P}_n$ of partitions satisfying the convergence condition. For any partition $\mathcal{P}$, we define the **mesh size** to be the length of its longest component:

$$\text{mesh}(\mathcal{P}) = \max_j \triangle x_j.$$

Then the following result, whose proof (using Theorem 3.7.6) is outlined in Exercise 8, gives the integrability of continuous functions:

Proposition 5.2.3. *If f is continuous on $[a, b]$ and $\mathcal{P}_n$ is any sequence of partitions with mesh($\mathcal{P}_n$) $\to 0$, then $\mathcal{P}_n$ satisfies the convergence condition*

$$\mathcal{U}(\mathcal{P}_n, f) - \mathcal{L}(\mathcal{P}_n, f) \to 0.$$

In particular, every continuous function is integrable.

In principle, the calculation of upper and lower sums involves a max-min problem on each component interval I_j, making the task of trying to calculate integrals even more daunting. However, if we know that a sequence of partitions $\mathcal{P}_n$ satisfies the convergence condition

$$\mathcal{U}(\mathcal{P}_n, f) - \mathcal{L}(\mathcal{P}_n, f) \to 0,$$

then the Squeeze Theorem (Theorem 2.4.7) gives us a way to calculate the integral without actually worrying about maxima or minima.
Given a partition $\mathcal{P}$, we can construct an analogue of the lower and upper sums by replacing the extreme value $\inf_{I_j} f$ or $\sup_{I_j} f$ in each component

I_j with the value of f at some sample point $x_j \in I_j$. This gives us what is called a **Riemann sum** associated to $\mathcal{P}$,

$$\mathcal{R}(\mathcal{P}, f) = \mathcal{R}(\mathcal{P}, f, \{x_j\}) = \sum_{j=1}^{n} f(x_j) \triangle x_j, \quad \text{where } x_j \in I_j \text{ for } j = 1, \ldots, n.$$

Of course, $\mathcal{R}(\mathcal{P}, f)$ depends on our particular choice of sample points $x_j \in I_j$, but whatever choice we make, we can guarantee the inequalities

$$\mathcal{L}(\mathcal{P}, f) \leq \mathcal{R}(\mathcal{P}, f) \leq \mathcal{U}(\mathcal{P}, f).$$

But then if we start with a sequence of partitions $\mathcal{P}_n$ satisfying the convergence condition, the above inequality squeezes *any* Riemann sums $\mathcal{R}(\mathcal{P}_n, f)$ toward the integral. This is a very useful technical observation.

Remark 5.2.4. *Suppose f is integrable on $[a, b]$ and $\mathcal{P}_n$ is a sequence of partitions satisfying the convergence condition*

$$\mathcal{U}(\mathcal{P}_n, f) - \mathcal{L}(\mathcal{P}_n, f) \to 0.$$

For each partition $\mathcal{P}_n$, form a Riemann sum $\mathcal{R}(\mathcal{P}_n, f)$.
Then, whatever choice of sample points we take, we are guaranteed that

$$\mathcal{R}(\mathcal{P}_n, f) \to \int_a^b f(x)\, dx.$$

When f is continuous, we can combine this remark with Proposition 5.2.3 to see that, so long as we pick a sequence of partitions with mesh size going to zero, we can take *any* choice of sample points in each component of each partition, and rest assured that the resulting Riemann sums will converge to the integral of f over $[a, b]$. In other words, "don't worry—be happy". :-)
Recall that we started our formulation of area with the premise that area is additive. An analogous property holds for definite integrals.

Lemma 5.2.5. *The definite integral is additive on domains: suppose f is defined on $[a, b]$ and c is an interior point, $a < c < b$.*
Then f is integrable on the interval $[a, b]$ precisely if it is integrable on each of the subintervals $[a, c]$ and $[c, b]$, and in that case,

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx.$$

Proof. Notice that if $\mathcal{P}'$ and $\mathcal{P}''$ are partitions of the subintervals $[a, c]$ and $[c, b]$ respectively, then the union $\mathcal{P} = \mathcal{P}' \cup \mathcal{P}''$ is a partition of $[a, b]$ which includes the subdivision point c; conversely, if $\mathcal{P}$ is any partition of $[a, b]$, we can refine it if necessary to include c, and then form the partitions of the subintervals

$$\mathcal{P}' = \{p \in \mathcal{P} \mid p \leq c\}$$
$$\mathcal{P}'' = \{p \in \mathcal{P} \mid p \geq c\}.$$

In either of these situations, it is easy to check that

$$\mathcal{L}(\mathcal{P}, f) = \mathcal{L}(\mathcal{P}', f) + \mathcal{L}(\mathcal{P}'', f)$$
$$\mathcal{U}(\mathcal{P}, f) = \mathcal{U}(\mathcal{P}', f) + \mathcal{U}(\mathcal{P}'', f).$$

In particular, a sequence of partitions $\mathcal{P}_n$ for $[a, b]$ satisfies the convergence condition precisely if the corresponding sequences of partitions $\mathcal{P}'_n$ for $[a, c]$ and $\mathcal{P}''_n$ for $[c, b]$ both satisfy it, and in this case

$$\int_a^b f(x)\, dx = \lim \mathcal{L}(\mathcal{P}_n, f)$$
$$= \lim \mathcal{L}(\mathcal{P}'_n, f) + \lim \mathcal{L}(\mathcal{P}''_n, f) = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx.$$

□

We can also establish several properties that relate the integrals of different functions over the same interval. We start with a particularly easy and useful version of monotonicity.

Lemma 5.2.6 (Comparison Theorem)**.** *Suppose f is integrable on $[a, b]$ and*

$$\alpha \leq f(x) \leq \beta \text{ for all } x \in [a, b]\,.$$

Then

$$\alpha(b - a) \leq \int_a^b f(x)\, dx \leq \beta(b - a).$$

If in addition f is continuous on $[a, b]$, then the inequalities are strict unless $f(x) = \alpha$ (resp. $f(x) = \beta$) for all $x \in [a, b]$.

Proof. The basic inequality really follows from our earlier remarks concerning the "trivial partition" and its refinements on page 323.

However, we review this argument here as a warmup for our proof of the second assertion in the lemma.

For any partition $\mathcal{P}$, we have, on each component I_j,

$$\alpha \leq \inf_{I_j} f \leq \sup_{I_j} f \leq \beta$$

so summing over all components we have

$$\alpha(b-a) = \sum_{j=1}^{n} \alpha \triangle x_j \leq \mathcal{L}(\mathcal{P}, f) \leq \mathcal{U}(\mathcal{P}, f) \leq \sum_{j=1}^{n} \beta \triangle x_j = \beta(b-a).$$

Since the integral lies between the lower and upper sum, we obtain the inequality.

Furthermore, if f is continuous and $f(x_0) > \alpha$ for some $x_0 \in [a, b]$, then we can find a small interval $J \subset [a, b]$ with positive length, containing x_0, on which $f(x) - \alpha$ is bounded away from zero, say

$$f(x) > \alpha + \varepsilon > \alpha \text{ for } x \in J.$$

Then for any partition $\mathcal{P}$ which includes the endpoints of J, we can, for each component interval I_j contained in J, modify our earlier inequality to read

$$\alpha + \varepsilon \leq \inf_{I_j} f \quad (I_j \subset J)$$

and hence summing over all components we have

$$\begin{aligned}\mathcal{L}(\mathcal{P}, f) &\geq \sum_{I_j \subset J} (\alpha + \varepsilon) \triangle x_j + \sum_{I_j \not\subset J} \alpha \triangle x_j \\ &= \alpha(b-a) + \varepsilon \|J\| > \alpha(b-a).\end{aligned}$$

This shows the left inequality in the theorem is strict when f is continuous and not equal to α everywhere; a similar argument takes care of the right-hand inequality. □

A very similar argument gives the more general statement

Proposition 5.2.7 (Monotonicity of Integrals)**.** *If f and g are integrable on $[a, b]$ and*

$$f(x) \leq g(x) \text{ for } x \in [a, b]$$

then

$$\int_a^b f(x)\, dx \leq \int_a^b g(x)\, dx.$$

If f and g are also continuous on $[a, b]$, then the inequality is strict unless $f(x) = g(x)$ for all $x \in [a, b]$.

Proof. Again, we see that for any partition $\mathcal{P}$,

$$\mathcal{L}(\mathcal{P}, f) \leq \mathcal{L}(\mathcal{P}, g)$$

and taking suprema over all partitions gives the inequality.
If f and g are also continuous and $f(x_0) < g(x_0)$ for some $x_0 \in [a, b]$, we again find an interval J of positive length containing x_0 where

$$f(x) + \varepsilon < g(x) \text{ for } x \in J.$$

Then for any partition $\mathcal{P}$ containing the endpoints of J, we have (Exercise 9)

$$\mathcal{L}(\mathcal{P}, f) + \varepsilon \|J\| \leq \mathcal{U}(\mathcal{P}, g). \tag{5.12}$$

But the supremum of $\mathcal{L}(\mathcal{P}, f)$ over *all* partitions $\mathcal{P}$ is the same as that over all partitions containing the endpoints of J, so

$$\int_a^b f(x)\,dx + \varepsilon \|J\| \leq \int_a^b g(x)\,dx.$$

□

Also, the integral is "linear" in the integrand (the function being integrated).

Proposition 5.2.8 (Linearity of Integrals)**.** *If f and g are integrable over $[a, b]$ and $r, s \in \mathbb{R}$, then $h(x) = rf(x) + sg(x)$ is integrable, and*

$$\int_a^b (rf(x) + sg(x))\; dx = r \int_a^b f(x)\,dx + s \int_a^b g(x)\,dx.$$

Proof. We will do this in two steps: first, that the integral is *homogeneous*

$$\int_a^b rf(x)\; dx = r \int_a^b f(x)\,dx$$

and second that it is *additive*

$$\int_a^b (f(x) + g(x))\; dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

Linearity then follows easily (Exercise 11).

To prove homogeneity, we note that, if $r \geq 0$, then for any partition $\mathcal{P}$, on each component interval I_j,

$$\inf_{I_j} rf = r \inf_{I_j} f$$
$$\sup_{I_j} rf = r \sup_{I_j} f$$

so

$$\mathcal{L}(\mathcal{P}, rf) = r\mathcal{L}(\mathcal{P}, f)$$
$$\mathcal{U}(\mathcal{P}, rf) = r\mathcal{U}(\mathcal{P}, f)$$

and taking suprema (*resp.* infima) over all partitions gives the result.
If $r < 0$, then in an analogous way

$$\begin{aligned} \mathcal{L}(\mathcal{P}, rf) &= r\mathcal{U}(\mathcal{P}, f) \\ \mathcal{U}(\mathcal{P}, rf) &= r\mathcal{L}(\mathcal{P}, f) \end{aligned} \tag{5.13}$$

and we get the same result (Exercise 10).
For additivity, we need to be more circumspect, since the inf (*resp.* sup) of a sum of functions need not equal the sum of their infima (*resp.* suprema). However, for all $x \in I_j$,

$$\inf_{I_j} f + \inf_{I_j} g \leq f(x) + g(x) \leq \sup_{I_j} f + \sup_{I_j} g$$

so for any Riemann sum associated to $\mathcal{P}$,

$$\mathcal{L}(\mathcal{P}, f) + \mathcal{L}(\mathcal{P}, g) \leq \mathcal{R}(\mathcal{P}, f + g) \leq \mathcal{U}(\mathcal{P}, f) + \mathcal{U}(\mathcal{P}, g).$$

Then using a sequence of partitions satisfying the convergence condition, we get the desired result. □

The linearity of the integral, together with additivity on domains, allows us to interpret more general integrals in terms of areas. Suppose the function f is integrable on $[a, b]$ and switches sign at only finitely many points: for example, suppose $a < c < b$ with $f(x) \geq 0$ when $a \leq x \leq c$ and $f(x) \leq 0$ when $c \leq x \leq b$ (clearly, then, $f(c) = 0$). Then we already know that

$$\int_a^c f(x)\, dx = A(S_+)$$

where S_+ is the region above the x-axis and below the graph of f

$$S_+ = \{(x,y) \,|\, a \leq x \leq c,\ 0 \leq y \leq f(x)\}.$$

Also, since $f(x) \leq 0$ on $[c,b]$, we have $-f(x) \geq 0$ there, so

$$-\int_c^b f(x)\,dx = \int_c^b (-f(x))\ dx$$

gives the area of the region above the x-axis and below the graph of $-f$; but this is the reflection across the x-axis of the region

$$S_- = \{(x,y) \,|\, c \leq x \leq b,\ 0 \geq y \geq f(x)\}$$

below the x-axis and *above* the graph of f. It follows that

$$\int_c^b f(x)\,dx = -A(S_-)\,.$$

We have established the simplest case of

Remark 5.2.9. *If f is integrable on $[a,b]$ and switches sign at only finitely many points, then*

$$\int_a^b f(x)\,dx = A(S_+) - A(S_-)$$

where S_+ is the region bounded by $y = f(x)$ above *the x-axis and S_- is the region bounded by $y = f(x)$* below *the x-axis.*

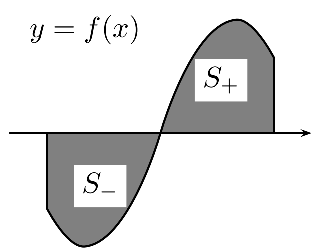


Figure 5.12: Remark 5.2.9

Another interpretation of the integral is as an *average*. Suppose f is continuous on $[a,b]$. Let us divide $[a,b]$ into N equal parts, by the partition $\mathcal{P}_N$: that is,

$$\triangle x_j = \triangle x = \frac{b-a}{N}$$

and

$$p_j = a + j\triangle x, \quad j = 0, \ldots, N.$$

Now, we "sample" the function by evaluating f at a randomly chosen point $x_j \in I_j = [p_{j-1}, p_j]$ for $j = 1, \ldots, N$, and consider the average of these numbers[10]:

$$A_N = \frac{f(x_1) + \ldots + f(x_N)}{N} = \frac{1}{N}\sum_{j=1}^{N} f(x_j)\,.$$

If we sample more and more points, what happens to these averages? Well, if we were to use the same sample points to form a Riemann sum, we would get

$$\mathcal{R}(\mathcal{P}_N, f) = \sum_{j=1}^{N} f(x_j)\,\triangle x_j = \sum_{j=1}^{N} f(x_j)\,\frac{b-a}{N} = (b-a)A_N.$$

Since f is integrable, we know that

$$\mathcal{R}(\mathcal{P}_N, f) \to \int_a^b f(x)\,dx$$

so that

$$A_N \to \frac{1}{b-a}\int_a^b f(x)\,dx.$$

Thus, we can define the **average of** f over $[a,b]$ to be

$$\mathrm{avg}_{[a,b]} f = \frac{1}{b-a}\int_a^b f(x)\,dx.$$

Note that if f is constant, say $f(x) = h$, for $a \le x \le b$, then $\mathrm{avg}_{[a,b]} f = h$. For other functions, we still think of $\mathrm{avg}_{[a,b]} f$ as the "average value" of f on $[a,b]$. The following easy result says that, at least for f continuous, the word "value" is justified in this context.

[10]Note that for different values of N, the number $f(x_j)$ for any particular j may have different meanings.

Proposition 5.2.10 (Integral Mean Value Theorem)**.** *If f is continuous on $[a, b]$, then f attains its average somewhere in (a, b); that is, for some $c \in (a, b)$,*

$$f(c) = avg_{[a,b]}f = \frac{1}{b-a}\int_a^b f(x)\,dx.$$

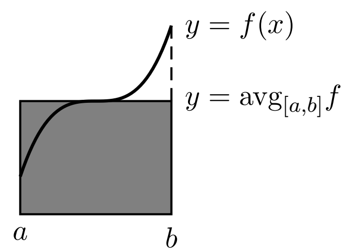


Figure 5.13: Integral Mean Value Theorem

Proof. Let

$$\alpha = \min_{[a,b]} f = f(x_{min})$$
$$\beta = \max_{[a,b]} f = f(x_{max})$$

which exist by the Extreme Value Theorem (Theorem 3.3.4). Then the Comparison Theorem for Integrals says

$$\alpha(b-a) \leq \int_a^b f(x)\,dx \leq \beta(b-a)$$

or

$$f(x_{min}) = \alpha \leq \mathrm{avg}_{[a,b]}f \leq \beta = f(x_{max})\,.$$

But then the Intermediate Value Theorem (Theorem 3.2.1) says that for some c between x_{min} and x_{max},

$$f(c) = \mathrm{avg}_{[a,b]}f.$$

We would like to say that c can be chosen *not* to be an endpoint. When $\alpha \neq \beta$, the comparison theorem (Lemma 5.2.6) says that

$$\alpha(b-a) < \int_a^b f(x)\,dx < \beta(b-a)$$

so that c is not equal to x_{min} or x_{max}, and hence is interior to $[a, b]$. □

Finally, we establish two ways in which behavior at a single point (or finitely many) need not affect the integral.

Lemma 5.2.11. *Suppose f is integrable on $[a,b]$, and g is defined on $[a,b]$, equal to f at all points of $[a,b]$ with one exception, $c \in [a,b]$. Then g is integrable, and*

$$\int_a^b f(x)\,dx = \int_a^b g(x)\,dx.$$

Proof. Let $\delta = |g(c) - f(c)|$. For any partition $\mathcal{P}$, let I_j be a component interval of $\mathcal{P}$ containing c. Then

$$\inf_{I_j} f - \delta \leq \inf_{I_j} g \leq \sup_{I_j} g \leq \sup_{I_j} f + \delta.$$

Since the upper (and lower) sums of f and g differ only in the term (or, if c is a division point, two terms) corresponding to I_j, we see that

$$|\mathcal{L}(\mathcal{P}, f) - \mathcal{L}(\mathcal{P}, g)| \leq 2\delta \|I_j\|$$
$$|\mathcal{U}(\mathcal{P}, f) - \mathcal{U}(\mathcal{P}, g)| \leq 2\delta \|I_j\|.$$

Now, pick a sequence $\mathcal{P}_n$ of partitions satisfying the convergence condition for f. Refining if necessary, we can assume that the component interval(s) of $\mathcal{P}_n$ containing c has length at most $\varepsilon_n \to 0$. Then

$$|\mathcal{L}(\mathcal{P}_n, f) - \mathcal{L}(\mathcal{P}_n, g)| \leq 2\delta\varepsilon_n$$
$$|\mathcal{U}(\mathcal{P}_n, f) - \mathcal{U}(\mathcal{P}_n, g)| \leq 2\delta\varepsilon_n$$

and it follows that

$$\mathcal{U}(\mathcal{P}_n, g) - \mathcal{L}(\mathcal{P}_n, g) \leq 4\delta\varepsilon_n \to 0,$$

so g is integrable and

$$\int_a^b f(x)\,dx = \lim \mathcal{L}(\mathcal{P}_n, f) = \lim \mathcal{L}(\mathcal{P}_n, g) = \int_a^b g(x)\,dx.$$

□

Our final result gives the most general class of functions which we will be integrating. In fact, as Riemann showed, it is possible to integrate a more general class of functions; this is the content of § 5.9. A function f defined on $[a,b]$ is **piecewise continuous** on $[a,b]$ if there are finitely many points of discontinuity for f on $[a,b]$.

Theorem 5.2.12. *If f is defined and bounded on $[a,b]$ and $c \in [a,b]$ is the only point of discontinuity on $[a,b]$, then f is integrable on $[a,b]$.*
Hence, every bounded, piecewise-continuous function is integrable.

Proof. We shall prove the result assuming $c = a$; the proof for $c = b$ is analogous, and then for $a < c < b$ it follows from additivity on domains (Lemma 5.2.5).
Since f is bounded on $[a,b]$, we can take α, β so that

$$\alpha \le f(x) \le \beta \text{ for all } x \in [a,b] \,.$$

Note that on any subinterval J of $[a,b]$ with length $\|J\|$,

$$\left(\sup_J f\right)\|J\| - \left(\inf_J f\right)\|J\| \le (\beta-\alpha)\|J\|.$$

Let a_n be a sequence of points with

$$a < a_n < a + \ell_n, \quad n = 1, 2, ...,$$

where

$$(\beta - \alpha)\ell_n < \frac{1}{2n},$$

and set

$$J_n = [a_n, b] \,.$$

For each $n = 1, 2, ...,$ f is continuous on J_n, hence integrable on J_n. Let $\mathcal{P}'_n$ be a partition of J_n with

$$\mathcal{U}(\mathcal{P}'_n, f) - \mathcal{L}(\mathcal{P}'_n, f) < \frac{1}{2n},$$

and let $\mathcal{P}_n$ be the partition of $[a,b]$ formed by adding in a: thus, the leftmost component interval of $\mathcal{P}_n$ is $[a, a_n]$ and the others are precisely the component intervals of $\mathcal{P}'_n$. Thus

$$\mathcal{L}(\mathcal{P}_n, f) = \left(\inf_{[a,a_n]} f\right)\ell_n + \mathcal{L}(\mathcal{P}'_n, f)$$

$$\mathcal{U}(\mathcal{P}_n, f) = \left(\sup_{[a,a_n]} f\right)\ell_n + \mathcal{U}(\mathcal{P}'_n, f)$$

and so

$$\begin{aligned}\mathcal{U}(\mathcal{P}_n, f) - \mathcal{L}(\mathcal{P}_n, f) &= \left(\sup_{[a,a_n]} f\right)\ell_n - \left(\inf_{[a,a_n]} f\right)\ell_n + \mathcal{U}(\mathcal{P}'_n, f) - \mathcal{L}(\mathcal{P}'_n, f)\\ &\leq (\beta - \alpha)\ell_n + \frac{1}{2n}\\ &\leq \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n} \to 0.\end{aligned}$$

Hence, f is integrable on $[a, b]$. □

We close this section with a brief summary of the history of the definitions and results in the last two sections: as we mentioned in the introduction to this chapter, seventeenth-century researchers, from Bonaventura Cavalieri (1598-1647) to Gottfried Wilhelm Leibniz (1646-1714), had in various ways regarded the integral as a (vaguely defined) sum of "infinitesimal" line segments. Augustin-Louis Cauchy (1789-1857), first in a paper in 1814 (when he was 25) and then in a book-length exposition in 1823, reformulated this idea more carefully as a limit of finite sums. He formed what we call Riemann sums but always using the value of the function at the left endpoint of a component interval as the height of the rectangle over that interval. He was dealing with (piecewise) continuous functions, and noted that the limiting value of the sums was independent of the choice of partitions, so long as the mesh size went to zero. He also had the Integral Mean Value Theorem (Proposition 5.2.10) and Theorem 5.2.12. However, thirty-one years later, in 1854, Bernhard Georg Friedrich Riemann (1826-1866) was trying to handle more general functions, and he introduced the idea of looking at an *arbitrary* value of the function in the component interval for the height of the rectangle. He recognized that with this definition functions did not need to be piecewise continuous in order to be integrable, and in fact gave an example to illustrate his point. We will explore his characterization of precisely what conditions are needed to insure integrability in § 5.9. In 1875 Jean-Gaston Darboux (1842-1917) clarified some of Riemann's ideas by introducing the ideas of the upper and lower sum; further refinements were carried out in the 1880's by Vito Volterra (1860-1940) and Giuseppe Peano (1858-1932). The latter also carried out, in 1887, a deep study of the idea of area, which was further refined and popularized in an 1893 book by Camille Marie Ennemond Jordan (1838-1922); these ideas were central to the revolutionary further extension of integration theory by Henri Léon Lebesgue (1875-1941).

Exercises for § 5.2

Answers to Exercises 1acegikmo, 3, 4aceg, 5, 12a are given in Appendix B.

Practice problems:

1. For each definite integral $\int_a^b f(x)\,dx$ below, (i) Sketch the area between the graph of the function and the x-axis; (ii) use known area formulas and the results of this section to evaluate the integral.

   (a) $\int_0^2 x\,dx$ (b) $\int_0^2 (x+1)\,dx$ (c) $\int_0^3 (x-1)\,dx$

   (d) $\int_{-1}^2 (1-x)\,dx$ (e) $\int_{-1}^2 (4-x)\,dx$ (f) $\int_{-1}^1 |x|\,dx$

   (g) $\int_{-1}^1 \sqrt{1-x^2}\,dx$ (h) $\int_0^2 \sqrt{4-x^2}\,dx$ (i) $\int_{-2}^2 3\sqrt{4-x^2}\,dx$

   (j) $\int_{-2}^2 (1+\sqrt{4-x^2})\,dx$ (k) $\int_0^2 (\sqrt{4-x^2}-x)\,dx$

   (l) $\int_0^{1/\sqrt{2}} (\sqrt{1-x^2}-x)\,dx$

   (m) $\int_0^3 \lfloor x \rfloor\,dx$, where $\lfloor x \rfloor$ is the largest integer $\leq x$.

   (n) $\int_0^3 f(x)\,dx$, where $f(x) = x - \lfloor x \rfloor$ is the fractional part of x.

   (o) $\int_0^2 f(x)\,dx$, where

   $$f(x) = \begin{cases} \sqrt{1-x^2} & \text{for } 0 \leq x \leq 1, \\ x & \text{for } 1 < x \leq 2. \end{cases}$$

2. Into how many equal parts do we need to divide $[0,1]$ to insure that any Riemann sum based on this partition will give

   $$\int_0^1 \sin\frac{\pi x^2}{2}\,dx$$

   to within three decimal places? (*Hint:* The function is strictly increasing on the given interval.)

3. Which is bigger, $\int_0^1 \sin\frac{\pi x^2}{2}\,dx$ or $\int_0^1 \sin\frac{\pi x^4}{2}\,dx$?

4. Suppose f is continuous on $[a,b]$ $(a<b)$ and
$$\int_a^b f(x)\,dx = 0.$$
Which of the following statements is necessarily true? For each, either provide a proof or a counterexample.

   (a) $f(x)=0$ for *all* $x\in[a,b]$.

   (b) $f(x)=0$ for *at least one* $x\in[a,b]$.

   (c) $\int_a^b |f(x)|\,dx = 0$.

   (d) $\left|\int_a^b f(x)\,dx\right| = 0$.

   (e) $\int_a^b (f(x)+1)\,dx = b-a$.

   (f) $\int_a^b |f(x)+1|\,dx = b-a$.

   (g) For every function g which is integrable on $[a,b]$,
$$\int_a^b [f(x)+g(x)]\,dx = \int_a^b g(x)\,dx.$$

5. What is the average value of the function $f(x)=\sqrt{1-x^2}$ over the interval $[-1,1]$? Over the interval $[0,1]$? (*Hint:* Interpret the integral as an area, and use known area formulas.)

Theory problems:

6. Show that if we interpret a definite integral whose *lower* limit of integration is greater than its *upper* limit of integration as the *negative* of the integral with the limits in the right order, that is,
$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx \quad \text{if } a>b$$
then for any three numbers a,b,c
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

7. (a) Show that for any function f, any partition $\mathcal{P} = \{p_0, ...p_n\}$, and any sample set $\{x_j \in I_j, j = 1, \ldots, n\}$,

$$|\mathcal{R}(\mathcal{P}, f, \{x_j\})| \leq \mathcal{R}(\mathcal{P}, |f|, \{x_j\}).$$

(Recall that $\mathcal{R}(\mathcal{P}, f, \{x_j\})$ is the Riemann sum for f using partition $\mathcal{P}$ and sample points $\{x_j\}$.)

(b) Use this to prove that, if f and its absolute value $|f|(x) := |f(x)|$ are both integrable on $[a, b]$, then

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx.$$

Note: In Exercise 12, we will show that whenever f is integrable, so is $|f|$. This is more involved than the above.[11]

8. Prove Proposition 5.2.3 as follows:

(a) Show that for any bounded function f on $[a, b]$ and any partition $\mathcal{P} = \{p_0, ..., p_n\}$ of $[a, b]$,

$$\mathcal{U}(\mathcal{P}, f) - \mathcal{L}(\mathcal{P}, f) = \sum_{j=1}^{n} (\sup_{I_j} f - \inf_{I_j} f) \triangle x_j.$$

(b) Given $\varepsilon > 0$, use Theorem 3.7.6 to find $\delta > 0$ such that $|x - x'| < \delta$ guarantees

$$|f(x) - f(x')| < \frac{\varepsilon}{b - a}.$$

(c) Show that for any interval $I \subset [a, b]$ of length δ or less,

$$\sup_{x \in I} f(x) - \inf_{x \in I} f(x) < \frac{\varepsilon}{b - a}.$$

(d) Then show that for any partition with mesh size $\operatorname{mesh}(\mathcal{P}) < \delta$, we have

$$\mathcal{U}(\mathcal{P}, f) - \mathcal{L}(\mathcal{P}, f) < \varepsilon.$$

9. Show that if we have functions f and g which satisfy

[11] Thanks to Jason Richards for conversations that led to this separation of the two exercises.

- $f(x) \leq g(x)$ for every $x \in [a, b]$
- $J \subset [a, b]$ is an interval of positive length and $\varepsilon > 0$ is such that

$$f(x) + \varepsilon < g(x) \text{ for } x \in J$$

then for any partition $\mathcal{P}$ containing the endpoints of J, we have

$$\mathcal{L}(\mathcal{P}, f) + \varepsilon \|J\| \leq \mathcal{U}(\mathcal{P}, g).$$

(*Hint:* First suppose that J itself is a component interval of $\mathcal{P}$.)

10. (a) Show that if $r < 0$ then for any function f defined on an interval I

$$\sup_{x \in I} r f(x) = r \inf_{x \in I} f(x)$$
$$\inf_{x \in I} r f(x) = r \sup_{x \in I} f(x) .$$

(b) Use this to prove Equation (5.13).

(c) Prove that the integral is homogeneous: for any $r \in \mathbb{R}$,

$$\int_a^b r f(x) \, , dx = r \int_a^b f(x) \, dx.$$

11. How do we combine the homogeneity and additivity of the integral to show that it is linear?

Challenge problems:

12. (a) Give an example of a function f defined on $[0, 1]$ for which $|f|$ is integrable, but f is *not* integrable.

(b) Show that for any function f and any partition $\mathcal{P}$ of $[0, 1]$,

$$\mathcal{U}(\mathcal{P}, |f|) - \mathcal{L}(\mathcal{P}, |f|) \leq \mathcal{U}(\mathcal{P}, f) - \mathcal{L}(\mathcal{P}, f).$$

(*Hint:* Use Equation (2.3) in Exercise 7, § 2.5)

(c) Use this to show that if f is integrable on $[0, 1]$, then so is $|f|$.

5.3 The Fundamental Theorem of Calculus

So far our treatment of the integral has been very abstract, and in general the integral of a function is difficult to calculate precisely from the definition, although the convergence condition allows us to find sequences of numbers which eventually approximate the integral.

In this section, we establish a stunning result, appropriately called the *Fundamental Theorem of Calculus*, which states in precise terms that integration and differentiation are mutually inverse operations. This phenomenon was observed by Barrow (and perhaps Fermat) in the early 1600's. A geometric proof was given (independently) by Newton and Leibniz in the late 1600's; this result leads to the technique of *formal integration* which we will develop in the next few sections (§ 5.4-§ 5.6).

The Fundamental Theorem of Calculus really has two parts. The first states that *every continuous function f is the derivative of some function*[12] F, and describes this function in terms of definite integrals. The second part, which follows from this description, says that *if F is continuously differentiable on $[a,b]$, then the definite integral of its derivative $f(x) = F'(x)$ is given by the formula*

$$\int_a^b f(x)\,dx = F(b) - F(a)\,.$$

The technique of formal integration essentially consists of starting with f for which we can find an antiderivative F (*i.e.*, $F'(x) = f(x)$) and then applying the formula above.

To describe the function F in the first part of the Fundamental Theorem of Calculus, we need to consider "definite" integrals whose endpoints can move.

Definition 5.3.1. *Suppose f is integrable on $[a,b]$. Then we define a new function, Φ, on $[a,b]$, by*

$$\Phi(t) = \begin{cases} 0 & \textit{if } t = a \\ \int_a^t f(x)\,dx & \textit{if } t > a. \end{cases}$$

When $f(x) \geq 0$ on $[a,b]$, we can think of $\Phi(t)$ as the area under the graph $y = f(x)$ between $x = a$ and the (moveable) right edge $x = t$ (Figure 5.14). We will often omit the explicit statement of the case $t = a$ in the definition above.

We will warm up to the Fundamental Theorem of Calculus by first proving

[12]Of course, by definition, F is continuously differentiable.

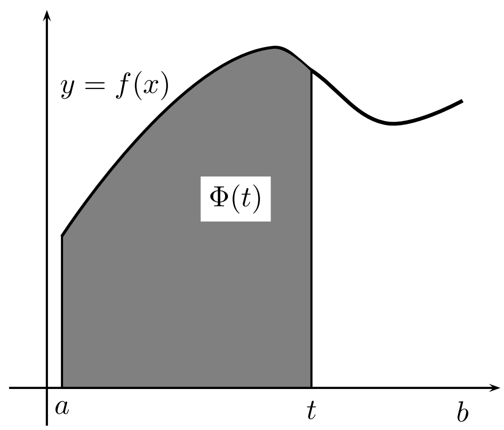


Figure 5.14: $\Phi(t) = \int_a^t f(x)\,dx$

Lemma 5.3.2. *For any function f which is (bounded and) integrable on $[a, b]$, the function*

$$\Phi(t) = \int_a^t f(x)\,dx$$

is continuous for $a \leq t \leq b$.

Proof. Pick $c \in [a, b]$; we will compute the one-sided limits of $\Phi(t)$ at $t = c$. Suppose $a \leq c < b$ and $t_i \downarrow c$; we write

$$t_i = c + s_i$$

where $s_i \downarrow 0$. By additivity on domains (Lemma 5.2.5), we have

$$\Phi(t_i) = \int_a^{t_i} f(x)\,dx = \int_a^c f(x)\,dx + \int_c^{t_i} f(x)\,dx$$
$$= \Phi(c) + \int_c^{t_i} f(x)\,dx,$$

or

$$\Phi(t_i) - \Phi(c) = \int_c^{t_i} f(x)\,dx.$$

Now, pick bounds for $f(x)$ on $[a, b]$, say

$$\alpha \leq f(x) \leq \beta \text{ for } x \in [a, b]\,.$$

Then we can apply the Comparison Theorem (Lemma 5.2.6) on $[c, t_i]$ to conclude that

$$\alpha(t_i - c) \leq \int_c^{t_i} f(x)\,dx \leq \beta(t_i - c).$$

But by construction,

$$t_i - c = s_i \to 0$$

so the Squeeze Theorem gives

$$\int_c^{t_i} f(x)\,dx \to 0,$$

and hence

$$\Phi(t_i) \to \Phi(c)$$

for any sequence $t_i \downarrow c$ in $[a, b]$, in other words

$$\lim_{t \to c+} \Phi(t) = \Phi(c) \text{ for } a \leq c < b.$$

For a sequence $t_i \uparrow c$, we start with

$$\begin{aligned}\Phi(c) = \int_a^c f(x)\,dx &= \int_a^{t_i} f(x)\,dx + \int_{t_i}^c f(x)\,dx \\ &= \Phi(t_i) + \int_{t_i}^c f(x)\,dx,\end{aligned}$$

or equivalently,

$$\Phi(c) - \Phi(t_i) = \int_{t_i}^c f(x)\,dx.$$

But an argument nearly identical to the above shows that the right side goes to zero, and so

$$\lim_{t \to c-} \Phi(t) = \Phi(c) \text{ for } a < c \leq b.$$

□

To prove *differentiability* of Φ at $t = c$, we need to estimate the difference quotient

$$\frac{\triangle\Phi}{\triangle t} = \frac{\Phi(t) - \Phi(c)}{t - c}.$$

Again we work with $t > c$. From the calculation in the preceding proof, if $t > c$,

$$\frac{\Phi(t) - \Phi(c)}{t - c} = \frac{1}{t - c}\int_c^t f(x)\,dx.$$

But this is the same as the average defined in the preceding section:

$$\frac{\triangle\Phi}{\triangle t} = \mathrm{avg}_I f$$

where I is the closed interval with endpoints t and c. If f is continuous at $x = c$, then Proposition 5.2.10 says that this equals the value of f at some point of I (that is, some point between t and c). As $t \to c$, this value must (invoking continuity of f at $x = c$) converge to $f(c)$; we have calculated that

$$\lim_{t\to c} \frac{\triangle\Phi}{\triangle t} = \lim_{t\to c} \frac{\Phi(t) - \Phi(c)}{t - c} = \lim_{t\to c} \mathrm{avg}_I f = \lim_{x\to c} f(x) = f(c)\,.$$

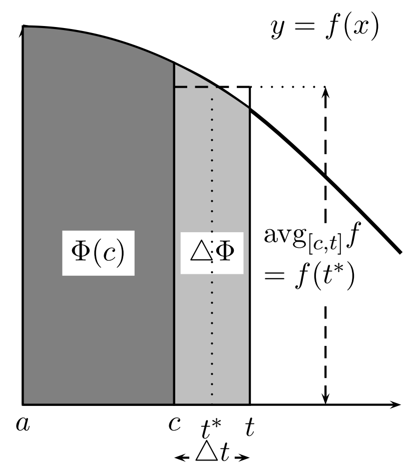


Figure 5.15: Fundamental Theorem of Calculus

This proves

Proposition 5.3.3 (Fundamental Theorem of Calculus I)**.** *If f is continuous on $[a, b]$ then*

$$\Phi(t) = \int_a^t f(x)\,dx$$

is differentiable[13] *at $t = c$, with*

$$\Phi'(c) = f(c) \text{ for all } c \in [a, b]\,.$$

[13](from the right if $c = a$ or from the left if $c = b$)

From this, the second part of the Fundamental Theorem of Calculus follows rather easily.

Proposition 5.3.4 (Fundamental Theorem of Calculus II)**.** *If f is continuous on $[a, b]$ and F is DICE on $[a, b]$ with*

$$F'(x) = f(x) \text{ for all } x \in (a, b)$$

then

$$\int_a^b f(x)\,dx = F(b) - F(a)\,.$$

Proof. We note that by Proposition 5.3.3, F has the same derivative as Φ, where

$$\Phi(t) = \int_a^t f(x)\,dx;$$

that is,

$$\Phi'(x) = F'(x) \text{ for all } x \in (a, b)\,.$$

In other words,

$$\Phi(x) - F(x)$$

is a continuous function on the closed interval $[a, b]$ whose derivative is zero everywhere in (a, b). As a consequence of the Mean Value Theorem (more precisely, of Corollary 4.9.3), it must be constant on $[a, b]$:

$$\Phi(x) - F(x) = C \quad a \leq x \leq b.$$

But we can evaluate both functions at $x = a$:

$$\Phi(a) - F(a) = 0 - F(a)$$

or

$$\Phi(x) = F(x) - F(a) \text{ for all } x \in [a, b]\,.$$

In particular, at $x = b$, we have

$$\int_a^b f(x)\,dx = \Phi(b) = F(b) - F(a)\,.$$

□

The procedure of evaluating F at two points and taking the difference comes up so often that there is a standard notation for it:

$$F(x)|_a^b = F(b) - F(a)\,.$$

We pronounce the left side above as "$F(x)$ *(evaluated) from* a *to* b".
To summarize, we pull together Proposition 5.3.3 and Proposition 5.3.4 into a full formal statement of the Fundamental Theorem of Calculus:

Theorem 5.3.5 (Fundamental Theorem of Calculus)**.** *Suppose f is continuous on $[a,b]$.*
Then

1. *The function*

$$\Phi(t) = \int_a^t f(x)\,dx$$

is differentiable for $a < t < b$, and

$$\Phi'(c) = \left.\frac{d}{dt}\right|_{t=c} \left[\int_a^t f(x)\,dx\right] = f(c) \quad a < c < b.$$

2. *If F is continuous on $[a,b]$ and satisfies*

$$F'(x) = f(x) \quad a < x < b$$

then

$$\int_a^b f(x)\,dx = F(x)|_a^b = F(b) - F(a)\,.$$

The second part of the Fundamental Theorem of Calculus lets us calculate the integrals of many functions by inverting the differentiation process. This can involve many "tricks of the trade" which we will consider in the next three sections. A function F which is continuous on $[a,b]$ and satisfies the condition

$$F'(x) = f(x) \quad a < x < b$$

is called an **antiderivative** of f on $[a,b]$. The first part of the Fundamental Theorem of Calculus says that one such antiderivative (provided f is continuous) is the function

$$\Phi(t) = \int_a^t f(x)\,dx.$$

The argument in the second part of the Fundamental Theorem of Calculus shows that every *other* antiderivative of f on $[a, b]$ differs from this one by a constant, and any antiderivative can be used to calculate the definite integral of f over $[a, b]$. Because of this connection, a list of all the antiderivatives of f is often called the **indefinite integral**[14] of f, and denoted by writing the integral without the endpoints (also called the **limits of integration**) a and b: we write

$$\int f(x)\, dx = F(x) + C$$

where F is any particular antiderivative of f. This means that *every* antiderivative of f looks like F plus some constant (denoted C): different values of this constant yield different functions on the list.

An easy and fundamental example of this is the integration of polynomials. Let us look at

$$\int x^n\, dx, \quad n = 0, 1, ...$$

where n is a nonnegative integer. We wish to find an antiderivative of x^n: since we know differentiation of a power of x leads to another power of x

$$\frac{d}{dx}[x^m] = mx^{m-1}$$

we see that to get x^n on the right, we need to start with $m = n + 1$. But the differentiation formula leads to a factor of $n + 1$ in front; to eliminate it, we need to start with its reciprocal in front:

$$\frac{d}{dx}\left[\frac{x^{n+1}}{n+1}\right] = \frac{n+1}{n+1}x^{n+1-1} = x^n.$$

Thus, one antiderivative of x^n is $x^{n+1}/(n+1)$. All others differ from it by a constant, so we have the indefinite integral formula

$$\int x^n\, dx = \frac{x^{n+1}}{n+1} + C, \quad n = 0, 1, ...$$

For example, when $n = 1$, one antiderivative of $f(x) = x$ is

$$F(x) = \frac{x^2}{2};$$

[14]The *definite* integral is a specific number, while the *indefinite* integral is an infinite list of functions.

but others are

$$F(x) = \frac{x^2}{2} + 1$$
$$F(x) = \frac{x^2}{2} - \sqrt{2}$$

and so on. So we write

$$\int x\,dx = \frac{x^2}{2} + C$$

or, equally well,

$$\int x\,dx = \frac{x^2+1}{2} + C$$

since $(x^2+1)/2$ is another antiderivative of x. Note that for a given antiderivative of x, the value of C in the first expression is different from its value in the second. The point, however, is that the collection of *all* functions generated by letting C run through *all* real (constant) values in the first expression is the same as the collection obtained in this way from the second expression.

The differentiation formula did not depend on m being an integer, so the integration formula

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C$$

also works for non-integer values of n: the only problem occurs when $n = -1$, since then the formula requires division by 0. This is no surprize: no power of x can give the antiderivative of x^{-1}: the only candidate would be x^0, but this gives a constant, whose derivative is 0: in fact, the formula

$$\int 0\,dx = C$$

is valid. But we can search through our known functions and locate a function whose derivative *is* $1/x$, namely the natural logarithm:

$$\frac{d}{dx}[\ln x] = \frac{1}{x}.$$

It follows that[15]

$$\int \frac{1}{x}\,dx = \ln x + C.$$

[15]Note that $\ln x$ is only defined for $x > 0$, so we need more to handle, for example, $\int_{-2}^{-1}(1/x)\,dx$. It turns out that the formula that works (for intervals that do not include $x = 0$) is $\int(1/x)\,dx = \ln|x| + C$. See Exercise 5 in § 4.4.

The formula for integrating x^n for n a nonnegative integer immediately gives us the integrals of polynomials, using the linearity of the derivative (or of the integral): for example, to find

$$\int (x^2 - 2x + 3)\, dx$$

we note that

$$\int x^2\, dx = \frac{x^3}{3} + C$$
$$\int (-2x)\, dx = -x^2 + C$$
$$\int 3\, dx = 3x + C.$$

That is, one antiderivative of x^2, $-2x$, and 3, respectively, is $x^3/3$, x^2, and $3x$. Clearly then, one antiderivative of the sum $x^2 - 2x + 3$ is the sum $x^3/3 - x^2 + 3x$; every other one differs from this by a constant, so

$$\int (x^2 - 2x + 3)\, dx = \frac{x^3}{3} - x^2 + 3x + C.$$

Notice that when we add the three formulas above, the sum of the three "$+C$"'s is written as "$+C$", not "$+3C$": this is because the sum of three arbitrary constants is still an arbitrary constant.

Exercises for § 5.3

Answers to Exercises 1acegi, 2, 3 are given in Appendix B.

Practice problems:

1. For each definite integral below, use the Fundamental Theorem of Calculus, together with your knowledge of derivatives, to

   (i) find an antiderivative for the integrand; and (ii) evaluate the definite integral.

   (a) $\int_0^1 4x^3\, dx$ (b) $\int_{-1}^1 4x^3\, dx$ (c) $\int_{-1}^2 (3x^2 + 1)\, dx$

   (d) $\int_1^4 \frac{x^{-1/2}}{2}\, dx$ (e) $\int_1^4 \frac{x^{-3/2}}{2}\, dx$ (f) $\int_1^4 \frac{3}{2}\sqrt{x+1}\, dx$

(g) $\displaystyle\int_1^2 \frac{1}{x}\,dx$ (h) $\displaystyle\int_0^1 e^x\,dx$ (i) $\displaystyle\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos x\,dx$

Theory problems:

2. If $f(x)$ is continuous on $[a,b]$ except for a jump discontinuity at $x = c$, what do you expect the behavior of $\Phi(x)$ to be? Consider the example

$$f(x) = \operatorname{sgn}(x) = \begin{cases} +1 & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0 \\ -1 & \text{if} \quad x < 0 \end{cases}$$

(see p. 493 for a picture). What is $\Phi(x)$ in this case?

Challenge problems:

3. Consider the function

$$\Phi(x) = \int_0^x \cos t^2\,dt.$$

(a) Find all critical points of Φ.

(b) Find all intervals on which Φ is increasing.

(c) Find all intervals on which Φ is decreasing.

(d) Determine all local maxima and local minima of Φ.

4. Consider the function

$$\Phi(x) = \int_0^{x^2} e^{-t^2}\,dt.$$

(a) Show that the function has a minimum at $(0,0)$.

(b) Show that this is the only local extremum of Φ.

(c) Show that Φ has a point of inflection at $x = \frac{1}{\sqrt{2}}$.

History notes:

Early calculations of $\int x^k\,dx$: Archimedes was the first to find an area bounded by a parabola, via an argument using triangles similar in style to his quadrature of the circle [31, pp. 233-252]. The formula for the area under the curve $y = x^k$ for k any *integer* greater than one was first formulated by Bonaventura Cavalieri (1598-1647) in 1635, based on intuitive arguments and a certain amount of inference from perceived patterns. The corresponding formula for k a positive *rational* number was formulated by John Wallis (1616-1703) in 1655, again based on inference from perceived patterns. More careful proofs were published by Pierre de Fermat (1601-1665), Gilles Personne de Roberval (1602-1675), Evangelista Torricelli (1608-1647) (a pupil of Cavalieri) and Blaise Pascal (1623-1662).

5. **$\int x^k\,dx$, k a positive integer:**

   (a) Show that if $\mathcal{P}_n$ is the division of $[0, b]$ into n equal parts then

$$\mathcal{L}(\mathcal{P}_n, x^k) = b^{k+1}\frac{0^k + 1^k + ... + (n-1)^k}{n^{k+1}}$$
$$\mathcal{U}(\mathcal{P}_n, x^k) = b^{k+1}\frac{1^k + 2^k + ... + n^k}{n^{k+1}}.$$

   (b) Fermat and Roberval, corresponding in 1636, based their analysis on the inequality

$$1^k+2^k+...+(n-1)^k < \frac{n^{k+1}}{k+1} < 1^k+2^k+...+(n-1)^k+n^k \quad (5.14)$$

   although it appears unclear whether either one had a proof of this [33, pp. 481-484]. Use Equations (5.4)-(5.6) (p. 321) to verify this inequality for $k = 1, 2, 3$.

   (c) Use induction on n (see Appendix A) to prove the inequality (5.14) in general. (*Note:* In the induction step, k is fixed; you are going from n to $n + 1$.)

   (d) Show that this inequality implies

$$\lim_{n\to\infty} \frac{1^k + 2^k + ... + n^k}{n^{k+1}} = \frac{1}{k+1}$$

   and conclude that the area under $y = x^k$, k a positive integer, up to the vertical line $x = b$, is

$$A = \frac{b^{k+1}}{k+1}.$$

   (e) Wallis, in 1655, attacked the same area by considering its ratio to the area of the circumscribed rectangle (*i.e.*, of height b^k)

[20, pp. 113-4], [51, pp. 244-6], which he considered approximated by the sum of heights of vertical "slices" through $x = p_i$ for p_i a sequence of equally spaced points in $[0, b]$. The ratio between these heights at $x = p_i$ is $p_i^k : b^k$. Show that if we pick the p_i to form an arithmetic progression of $n+1$ numbers (*i.e.*, , $p_i = i\delta$, $i = 0, \ldots, n$, where $\delta > 0$ is some constant), then the ratio of the two sums (written as a fraction) is

$$\frac{0^k + 1^k + 2^k + ... + n^k}{n^k + n^k + n^k + ... + n^k}.$$

For $k = 3$, Wallis concluded that this fraction equals $\frac{1}{4} + \frac{1}{4n}$, and hence that the limit as $n \to \infty$ is $\frac{1}{4}$. (He had previously determined that for $k = 1, 2$ the corresponding limits were $\frac{1}{2}, \frac{1}{3}$.) Verify this using Equations (5.4)-(5.6). Wallis then conjectured by analogy that for *every* positive integer k,

$$\lim_n \frac{0^k + 1^k + 2^k + ... + n^k}{n^k + n^k + n^k + ... + n^k} = \frac{1}{k+1}.$$

Show that this also gives the value obtained by Fermat and Roberval.

6. **Fermat's alternate calculation of $\int x^k\,dx$, $k > 0$ rational:** A second procedure was given by Fermat in 1658 which allowed him to handle rational values of k and also to handle "hyperbolas" (corresponding to $k < 0$), with the exception of $k = -1$ (see Exercise 11 in § 5.7). The idea was to form an *infinite* partition, whose points are in a *geometric* instead of arithmetic progression. (*cf* [33, p. 483],[47, pp. 240-241], [20, 109-111])

   Pick $r \in (0, 1)$ and consider the infinite partition $\mathcal{P}_r$ of $[0, b]$ given by

   $$p_i = r^i b, \quad i = 0, 1, 2, , ...$$

   (See Figure 5.16.) Note that the points p_i are now numbered right-to-left instead of left-to-right.

   (a) Show that when k is a positive integer, the sum of areas of

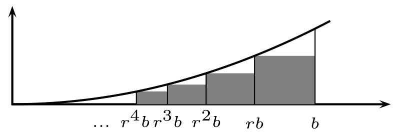


Figure 5.16: Fermat's alternate calculation of $\int_0^b x^k\,dx$, $k > 0$

inscribed rectangles for this infinite partition is given by

$$\begin{aligned} A(r) &= \sum_{i=0}^{\infty}(r^{i+1}b)^k r^i(1-r)b \\ &= \frac{b^{k+1}(1-r)}{r}\sum_{i=0}^{\infty}(r^{k+1})^i = \frac{b^{k+1}(1-r)}{r}\cdot\frac{r^{k+1}}{1-r^{k+1}} \\ &= b^{k+1}r^k\frac{1-r}{1-r^{k+1}} = \frac{b^{k+1}r^k}{1+r+r^2+...+r^k}. \end{aligned}$$

(b) Show that the fraction in this last expression goes to $\dfrac{1}{k+1}$ as $r \to 1^-$, and deduce the formula

$$A = \frac{b^{k+1}}{k+1}.$$

(c) If the exponent k in the equation $y = x^k$ is rational rather than integer, say $k = \frac{p}{q}$, then the only change in the preceding argument is to substitute $s = r^{1/q}$ and re-evaluate the fraction in the next-to-last line as

$$\frac{1-r}{1-r^{k+1}} = \frac{1-s^q}{1-s^{p+q}} = \frac{1+s+s^2+...+s^{q-1}}{1+s+s^2+...+s^{p+q-1}}$$

and noting that as $r \to 1$ also $s \to 1$, and the last fraction goes to $\dfrac{q}{p+q} = \dfrac{1}{k+1}$.

7. **Wallis' calculation of $\int x^k\,dx$, $k > 0$ rational:** ([51, pp. 246-7],[20, pp. 115-6],[33, pp. 486-7]) Wallis boldly conjectured that the formula limit

$$\lim_n \frac{0^k + 1^k + 2^k + ... + n^k}{n^k + n^k + n^k + ... + n^k} = \frac{1}{k+1}$$

generalizes to $k = \frac{p}{q}$, thus generalizing the formula for the area under $y = x^k$ to rational positive k. He was able to verify this for $p = 1$.

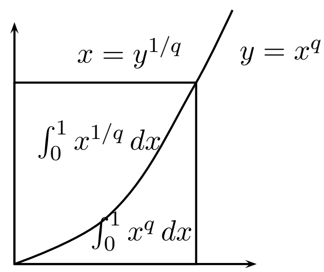


Figure 5.17: Area for fractional powers.

Consider the unit square, divided by the curve $y = x^q$, which can also be written $x = y^{1/q}$ (Figure 5.17). The area *above* this curve is congruent to the area *under* the curve $y = x^{1/q}$ $(0 \leq x \leq 1)$, so the sum of the area under $y = x^q$ and that for $y = x^{1/q}$ is the area of the square, or 1. Thus, in our notation,

$$\int_0^1 x^{1/q}\,dx = 1 - \int_0^1 x^q\,dx = 1 - \frac{1}{q+1} = \frac{q}{q+1}$$

and you should check that this is the same as the formula obtained by Fermat in case $k = \frac{1}{q}$.

8. **Leibniz's version of the Fundamental Theorem of Calculus[16]:** [51, pp. 282-284], [38, pp.132-135]

   Leibniz's formulation of the Fundamental Theorem of Calculus (see Figure 5.18[17]) is set forth in a paper published in 1693 in the *Acta Eruditorum.*

   Given the curve $AH(H)$ (dashed in Figure 5.18), Leibniz wants to find its *quadratrix*, the curve whose abcissa gives the area under the

---

[16]Thanks to my colleague Lenore Feigenbaum for helping me understand this proof.

[17]Figure 5.18 is a version of Leibniz's original figure, as transcribed in [51, p. 283] and [38, p. 133]; I have kept the labeling of points, but in the labeling of lengths have interchanged x and y. I have also replaced the triangle Leibniz used as his "characteristic triangle" with one that is more clearly related to the increments: in fact, Leibniz's proof is based on the similarity between the triangle he uses and the one we use.

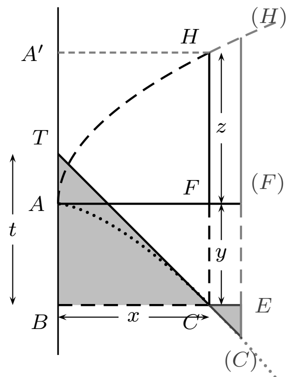


Figure 5.18: Leibniz's proof of the Fundamental Theorem of Calculus

curve. He constructs the curve $AC(C)$ (dotted in Figure 5.18) so that it starts at A and at each position C the tangent to the curve (the line CT) hits the vertical axis through A at such a point that the distance BT equals[18] the area of the rectangle[19] $AFHA'$: that is, he demands that

$$t = zx. \tag{5.15}$$

He then claims that the abcissa y gives the area of the region AFH under the curve $AH(H)$ from A to F. To this end, he forms the *characteristic triangle* $\triangle CE(C)$ for the purported quadratrix, corresponding to the increment $H(H)$ in the original curve. The relevant sides of the characteristic triangle (the small gray triangle in Figure 5.18) are $dx = CE$ and $dy = E(C)$. The triangle $\triangle CE(C)$ is similar to the triangle $\triangle CBT$ (the large gray triangle in Figure 5.18), so we have, using our labeling of lengths,

$$t : x = dy : dx. \tag{5.16}$$

Substitution of Equation (5.15) into Equation (5.16) yields

$$zx : x = dy : dx$$

or

$$dy = z\,dx$$

[18] Here, I am modernizing the proof: Leibniz picks a line segment a–which we can think of as a unit of area–and demands that the ratio of BT to BC be the same as the ratio of the area of the rectangle $AFHA'$ to a—that is, $t : a = zx : a$, or $t : x = z : 1$.

[19] I have added the point A' and the line AA' to the original figure for clarity.

from which it follows that

$$\int dy = \int z\,dx.$$

The integral on the left is (using Leibniz's notion that the integral is the sum of increments) equal to y, while the integral on the right is the desired area AFH. Thus, the constructed curve $AC(C)$ is the quadratrix of the original curve $AH(H)$.

How does this translate into our version of (the first part of) the Fundamental Theorem of Calculus?

9. **Newton's version of the Fundamental Theorem of Calculus:** [51, pp.303-4], [20, pp. 194-6]

Newton's formulation of the Fundamental Theorem of Calculus can be explained in terms of Figure 5.19. We set $be = 1$, and consider the

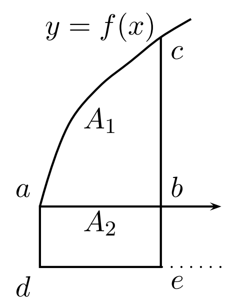


Figure 5.19: Fundamental Theorem of Calculus according to Newton

ratio between the rates of change of the two areas as the line ebc moves. Draw a diagram to show that this is given by the ratio $bc : be$, immediately giving

$$\frac{\dot{A}_1}{\dot{A}_2} = \frac{bc}{be} = \frac{f(x)}{1}.$$

This exposition follows that of Edwards [20, pp. 194-6], who refers to the fifth and seventh problems in Newton's 1666 tract, *De Analysi per Æquationes Numero Terminorum Infinitas*, his earliest exposition of his calculus, which was only published posthumously, in an English translation by John Stewart in 1745. These problems appear as numbers VII and IX in *Methodus Fluxionum et Serierum*

Infinitarum, a 1671 reworking of *De Analysi*, again published posthumously in an English translation in 1737. The English translations of these two works are reprinted in [55, vol. 1].

A different proof, much closer in spirit (and even in notation) to Leibniz's proof above, was given by Newton in an appendix to his *Opticks* of 1704. This is reproduced in [51, p. 303ff]. The standard collection of Newton's original works, with English translations where appropriate, is the monumental [56].

5.4 Secrets of Formal Integration: Manipulation Formulas

The Fundamental Theorem of Calculus in the previous section allows us to evaluate many integrals using antiderivatives. On the most naive level, this means we can read any derivative formula backwards to deduce an integral formula. However, it is also possible to adapt some of the manipulation formulas for derivatives—notably the chain rule and the product rule—into manipulation formulas for integrals. These, together with a few clever tricks, greatly expand our repertoire of integration formulas.
We owe the process of formal integration (as well as the formal differentiation of §§4.2-4.5) in large part to Leibniz. He initiated the "differential" and integral notation we use today, with the express goal of creating an efficient way to reach correct conclusions automatically, without repeating the chain of arguments that justify them every time. In this, he succeeded brilliantly, to the extent that many users of calculus blissfully ignore all questions of justification; this is what makes the subject a *calculus*, an effective set of tools for solving problems.

Formal Differentials

We begin by reviewing our basic derivative formulas, adapting each for use as an integration formula. We will at the same time write these in a form that closely resembles what appears under the integral sign, following the notation initiated by Leibniz. If F is a differentiable function, then its **formal differential** will be written

$$dF = F'\,dx.$$

This looks (formally!) like the result of multiplying dF/dx (the alternate notation for F') by its denominator. As an explanation, this is nonsense, but it does describe (formally) how things behave.[20] We saw, as a special case of the integration of polynomials, the indefinite integral formula

$$\int dx = x + C.$$

The differential notation lets us write the Fundamental Theorem of Calculus in the analogous form

$$\int F'(x)\ dx = \int dF = F(x) + C$$

(at least if F' is continuous). This formalism will allow us to handle many integration problems as ways of transforming an expression appearing under an integral sign into something recognizable as a formal differential. In the last section, we saw that the differentiation formula

$$d[x^m] = mx^{m-1}dx$$

can be manipulated to give

$$x^n dx = \frac{1}{n+1}d[x^{n+1}] \quad (n \neq -1)$$

which we interpret, in turn, as

$$\int x^n\, dx = \int \frac{1}{n+1}d[x^{n+1}] = \frac{1}{n+1}\int d[x^{n+1}] = \frac{x^{n+1}}{n+1} + C \quad (n \neq -1).$$

Similarly,

$$d\ln|x| = \frac{1}{x}dx$$

turns into

$$\int \frac{1}{x}\, dx = \int d\ln|x| = \ln|x| + C.$$

What about other formulas? The easiest to remember is

$$de^x = e^x\, dx$$

which gives

$$\int e^x\, dx = \int de^x = e^x + C.$$

[20]In fact, Leibniz thought of dF as a kind of infinitesimal difference of F; this is translated by Child in [14] as the *moment* of F. When F is a function of time, then, as Child notes, dF corresponds to what Newton calls the *fluxion* of F, and taking dF/dt to mean the "velocity" of F, the formula $dF = (dF/dt)dt$ makes perfect sense.

We also have the pair of trigonometric differential formulas

$$d\sin x = \cos x\,dx$$
$$d\cos x = -\sin x\,dx$$

which translate to the integral formulas

$$\int \cos x\,dx = \int d\sin x = \sin x + C$$
$$\int \sin x\,dx = \int -d\cos x = -\cos x + C.$$

A second pair of trigonometric differentiation formulas involves the tangent and secant:

$$d\tan x = sec^2 x\,dx$$
$$d\sec x = \sec x \tan x\,dx$$

leading to the integral formulas

$$\int \sec^2 x\,dx = \int d\tan x = \tan x + C$$
$$\int \sec x \tan x\,dx = \int d\sec x = \sec x + C.$$

Notice that these *don't* tell us how to integrate $\sec x$ or $\tan x$ on their own; this is one of the places where we need new tricks. There is also a family of formulas involving $\csc x$ and $\cot x$, which are parallel to these, but with a "minus" sign (see Exercise 3).
Our final differentiation formulas involve *inverse* trig functions:

$$d\arcsin x = \frac{dx}{\sqrt{1-x^2}}$$
$$d\arctan x = \frac{dx}{x^2+1}$$
$$d\,\mathrm{arcsec}\,x = \frac{dx}{x\sqrt{x^2-1}} \quad x > 0$$

which yield the integration formulas

$$\int \frac{dx}{\sqrt{1-x^2}}dx = \arcsin x + C$$
$$\int \frac{dx}{x^2+1}dx = \arctan x + C$$
$$\int \frac{dx}{x\sqrt{x^2-1}}dx = \mathrm{arcsec}\,x + C \quad x > 0.$$

These are particularly interesting, because they cross the line between algebra and trigonometry: in particular, the second formula shows that the *integral* of a rational function need not be a rational function.

The Chain Rule and Change of Variables

The chain rule (for real-valued functions of a real variable) can be written in the following way: suppose we can express y as a function of the variable u

$$y = f(u)$$

and in turn u can be expressed as a function of x:

$$u = u(x).$$

Then, of course, y can be expressed as a (composite) function of x, whose derivative is given by the formula

$$\frac{dy}{dx} = f'(u)\,u'(x) = \frac{dy}{du}\frac{du}{dx}.$$

Now, if y is regarded as a function of u, we write its formal differential as

$$dy = f'(u)\,du$$

but, if y is regarded as a function of x, we write

$$dy = \frac{dy}{dx}\,dx = f'(u)\,\frac{du}{dx}\,dx.$$

The second expression for dy is obtained from the first by replacing du with $(du/dx)\,dx$; this is not at all surprizing, from a formal perspective, since the very definition of du would give the equation

$$du = \frac{du}{dx}\,dx.$$

Note, however, that it points out the need to pay attention to the "differential" part of any such expression.

This formula has consequences for integrals. As a first example, consider the definite integral

$$\int_0^{\pi/3} \sin 2x\,dx.$$

If we had $\sin\theta$ instead of $\sin 2x$ under the integral, we could use the fact that $-\cos\theta$ is an antiderivative of $\sin\theta$ to evaluate the integral. If we set

$$\theta = 2x$$

then indeed

$$\sin\theta = \sin 2x.$$

However, to evaluate the integral, we also need to deal with the differential dx at the end. The rule above says

$$d\theta = d(2x) = 2\,dx$$

and we can solve this for dx

$$dx = \frac{1}{2}d\theta.$$

Finally, the limits of integration are $x = 0$ and $x = \pi/3$ (that is, we integrate over $x \in [0, \pi/3]$). In terms of θ our limits are $\theta = 2 \cdot 0 = 0$ to $\theta = 2 \cdot (\pi/3) = 2\pi/3$. Hence we can rewrite our definite integral in terms of θ

$$\int_0^{\pi/3} \sin 2x\,dx = \int_0^{2\pi/3} (\sin\theta)\left(\frac{1}{2}d\theta\right) = \frac{1}{2}\int_0^{2\pi/3} \sin\theta\ d\theta.$$

To evaluate this, we note that θ is merely a placeholder: we are looking for the area under the graph of $y = \sin\theta$ over the interval $0 \le \theta \le 2\pi/3$; and changing the name of the variable from θ to x (or vice-versa) doesn't change the area. Since $-\cos\theta$ is an antiderivative of $\sin\theta$, we have

$$\begin{aligned}\frac{1}{2}\int_0^{2\pi/3} \sin\theta\ d\theta = \frac{1}{2}(-\cos\theta) &= -\frac{1}{2}\cos\frac{2\pi}{3} + \frac{1}{2}\cos 0 \\ &= -\frac{1}{2}\left(-\frac{1}{2}\right) + \frac{1}{2}(1) = \frac{1}{4}.\end{aligned}$$

Note that if we had ignored the differential and the limits of integration, and simply evaluated $-\cos\theta$ over the original limits of integration, we would have obtained

$$\Big(-\cos\theta\Big)_0^{\pi/3} = -\cos\frac{\pi}{3} + \cos 0 = -\left(\frac{1}{2}\right) + 1 = \frac{1}{2} \quad ☹$$

which, of course, is *wrong*. The moral of this story is that we can perform substitutions in a definite integral, but we must deal not just with the integrand, but *also* with the "differential" term *and* the limits of integration.

A second example illustrates a more subtle (and more powerful) use of this device. Let us try to evaluate the indefinite integral

$$\int x \sin x^2 \, dx.$$

We would like to replace x^2 with θ inside the sine function—that is, we wish to rewrite the integral in terms of

$$\theta = x^2.$$

To do this, we need also to deal with the differential: it is easy to calculate that

$$d\theta = 2x \, dx.$$

But our *given* integral contains xdx; again, we can solve for this:

$$x \, dx = \frac{1}{2} d\theta$$

and then carry out the substitution:

$$\int (\sin x^2)(x \, dx) = \int (\sin \theta)(d\theta) = -\cos \theta + C.$$

In this form, the answer is rather useless, but we can now substitute *back* the relation $\theta = x^2$ to conclude that

$$\int (\sin x^2) x \, dx = -\cos x^2 + C.$$

Notice that if we had tried to use the same substitution on the indefinite integral

$$\int \sin x^2 \, dx$$

which looks simpler than the earlier one, we would have been in trouble, since we have only dx (not $x \, dx$) under the integral, and so solving for $d\theta = 2x \, dx$ in terms of θ would give us headaches. In fact, there is no way to evaluate this last integral in "closed form"[21]: no function that can be expressed algebraically in terms of our standard functions has derivative $\sin x^2$.

[21] It *can* be done via power series: see Exercise 36 in § 6.5.

As a final example, we consider the indefinite integral

$$\int \frac{e^x}{e^x + 1}\, dx.$$

This is harder to see as a pattern. However, in view of the integration formula

$$\int \frac{du}{u} = \ln|u| + C$$

we might try to substitute

$$u = e^x + 1.$$

Luckily for us, the corresponding differential

$$du = e^x\, dx$$

is precisely the numerator of our expression. So we can perform the substitution

$$u = e^x + 1$$

to get

$$\int \frac{e^x\, dx}{e^x + 1} = \int \frac{du}{u} = \ln|u| + C = \ln(e^x + 1) + C.$$

The effective use of substitution in integrals requires practice—it is an art that depends on recognizing patterns.

The Product Rule and Integration by Parts

The product rule for derivatives says that the derivative of a product equals a sum of two terms, each consisting of *one* of the factors *differentiated* and the *other left alone*. In terms of differentials, this can be written

$$d(uv) = u\, dv + v\, du.$$

If we put integral signs around this, we obtain

$$uv = \int u\ dv + \int v\ du$$

or, solving for the first integral, we get the **integration by parts** formula

$$\int u\ dv = uv - \int v\ du.$$

An example illustrates how this can be used. Consider the indefinite integral

$$\int x \sin x \, dx.$$

We know how to integrate (as well as differentiate) each of the factors x and $\sin x$ in the integrand, but their *product* is a different matter. Suppose that we set $u = x$ (the first factor) in the integration by parts formula. To apply the formula, we need to find v with

$$dv = \sin x \, dx.$$

But this is easily found by integration of this factor alone:

$$v = \int dv = -\cos x + C.$$

The "$+C$" is an extravagance here: we need to find *some* function v satisfying our condition (not *all* of them). So we set

$$u = x, \quad v = -\cos x.$$

Then

$$du = dx, \quad dv = \sin x \, dx$$

so

$$\begin{aligned} \int x \sin x \, dx &= \int u \, dv \\ &= uv - \int v \, du \\ &= x(-\cos x) - \int (-cosx) \, dx \\ &= -x \cos x + \int \cos x \, dx. \end{aligned}$$

This last integral is easy; we conclude that

$$\int x \sin x \, dx = -x \cos x + \sin x + C.$$

Let us review this process, using our example as a guide. The left side of the integration-by-parts formula

$$\int u \, dv = uv - \int v \, du$$

is an interpretation of our original integral: the integrand is a product of two factors (in our case, x and $\sin x$). One of these is designated "u" (in our case, $u = x$), and the rest of the integrand (including the differential dx) is designated "dv": this means that v is an *anti*derivative of the second factor (in our case, $\sin x = dv$ gives $v = -\cos x$). What appears on the right side of the formula? The first term uv ($-x \cos x$ in our case) involves no integration: the punchline is the second term, $\int v\, du$, whose integrand is the *derivative* of u times the *antiderivative* of dv. In our case, this is $\int -\cos x\, dx$, an easier integral than the original.

So to use the formula, we parse our given integrand as a product of two factors in such a way that the *derivative* of the first factor times an *antiderivative* of the *second* is easier to integrate than the original integrand. This takes practice, skill, and foresight. For example, if in our original problem

$$\int x \sin x\, dx$$

we had picked

$$u = \sin x,\ dv = x\, dx$$

then

$$du = \cos x\, dx, \quad v = \frac{x^2}{2}$$

and the integration-by-parts formula would have led to

$$\begin{aligned}\int x \sin x\, dx &= \int u\ dv = uv - \int v\, du \\ &= \frac{x^2}{2} \sin x - \int \frac{x^2}{2} \cos x\, dx.\end{aligned}$$

While this is a true statement, it is useless for purposes of evaluating the integral, since the integral on the right is *worse* than the original!

The general parsing strategy for integration by parts is guided by a few observations. We need to be able to integrate "dv" so if one part is easy to integrate and the other is not, then (since we can differentiate more easily than we can antidifferentiate) the harder part should normally go into "u". This applies, for example to $\ln x$, $\tan x$, $\sec x$, and rational functions (with nontrivial denominator). Powers of x go *up* when we integrate and *down* when we differentiate, so generally putting a (positive) power of x into "u" is more likely to simplify the integral. The functions $\sin x$, $\cos x$, and e^x are relatively neutral in this regard: the effect of integration or of differentiation does not complicate matters.

Sometimes, integration by parts needs to be used several times to get to an integral we can do: each step simplifies the problem. For example, the integral

$$\int x^2 e^x \, dx$$

is most naturally parsed as

$$u = x^2, \quad dv = e^x$$

giving

$$du = 2x \, dx, \quad v = e^x$$

so

$$\int x^2 e^x \, dx = x^2 e^x - \int 2xe^x \, dx.$$

The integral on the right is still not obvious, but we *have* reduced the power of x that appears. That was so much fun, we do it again, parsing $2xe^x \, dx$ as

$$u = 2x \quad dv = e^x \, dx$$

to get

$$du = 2 \, dx \quad v = e^x$$

and hence

$$\begin{aligned} \int 2xe^x \, dx &= 2xe^x - \int 2e^x \, dx \\ &= 2xe^x - 2\int e^x \, dx \\ &= 2xe^x - 2e^x + C. \end{aligned}$$

Substituting back into our earlier formula, we find

$$\begin{aligned} \int x^2 e^x \, dx &= x^2 e^x - [2xe^x - 2e^x] + C \\ &= (x^2 - 2x + 2)e^x + C. \end{aligned}$$

Sometimes, we need faith. For example, to evaluate

$$\int x \ln x \, dx$$

we *can't* put $u = x$, as we would like to, because we don't know how to integrate $\ln x \, dx$; this means we need

$$u = \ln x$$

so

$$x\,dx = dv$$

and the power of x will have to go *up* when we integrate:

$$v = \frac{x^2}{2}.$$

Fortunately in this case the differentiation of $\ln x$ comes to our rescue:

$$du = d\ln x = \frac{1}{x}\,dx$$

so

$$\begin{aligned}\int x\ln x\,dx = \int u\,dv = uv - \int v\,du \\ &= \frac{x^2}{2}\ln x - \int \frac{x}{2}\,dx \\ &= \frac{x^2}{2}\ln x - \frac{x^2}{4} + C.\end{aligned}$$

A second example requires, not just faith, but nerves of steel. Let us try to evaluate the integral

$$\int e^x \sin x\,dx.$$

Our observations rule out $u = e^x \sin x$ (since then $dv = dx$, so $v = x$ and the resulting integral $\int v\,du$ will involve terms like $xe^x \sin x$ and $xe^x \cos x$, an integration nightmare). But otherwise, the choice of u is unclear. Let us just try

$$u = e^x, \quad dv = \sin x\,dx.$$

Then

$$du = e^x\,dx, \quad v = -\cos x$$

and

$$\begin{aligned}\int e^x \sin x\,dx = \int u\,dv = uv - \int v\,du \\ &= -e^x \cos x - \int(-\cos x)(e^x\,dx) \\ &= -e^x \cos x + \int e^x \cos x\,dx.\end{aligned}$$

At this stage, our integration problem has changed, but not for the better: it is essentially on the same level of difficulty as our starting point.
However, we forge ahead, applying a similar parsing to our new integral

$$\int e^x \cos x \, dx$$

as

$$u = e^x, \quad dv = \cos x \, dx$$

so that

$$du = e^x \, dx, \quad v = \sin x$$

and

$$\int e^x \cos x \, dx = e^x \sin x - \int e^x \sin x \, dx.$$

This is truly discouraging: we have come back to our *first* integral!
However, a very cunning individual noticed that all is not lost. Substitute this back into our *original* integration-by-parts formula, to get

$$\begin{aligned} \int e^x \sin x \, dx &= -e^x \cos x + \int e^x \cos x \, dx \\ &= -e^x \cos x + e^x \sin x - \int e^x \sin x \, dx. \end{aligned}$$

Now step back a few paces and observe that this equation can be solved *algebraically* for our unknown integral. Moving the right-hand integral to the left side, we get

$$2 \int e^x \sin x \, dx = -e^x \cos x + e^x \sin x$$

and dividing by 2 (and adding on the constant of integration)

$$\int e^x \sin x \, dx = -\frac{1}{2} e^x \cos x + \frac{1}{2} e^x \sin x + C.$$

This process is so involved that we should check the validity of the answer, by differentiating the right side and verifying that this yields the integrand $e^x \sin x$.
This sleight-of-hand works for integrands like $e^{ax} \sin bx$ (or $e^{ax} \cos bx$) with any $a, b \neq 0$ (Exercise 2). One word of warning: the *first* time you apply integration parts to these integrals, you are free to parse either way (either $u = e^{ax}$ or $u = \sin bx$). However, the *second* time, when you are integrating

$e^{ax}\cos bx$, you must make the *parallel* choice (*i.e.*, either u is the exponential factor *both* times, or else u is the trig function *both* times). Otherwise, you really *do* go in circles.

Exercises for § 5.4

Answers to Exercises 1acegikmo, 2 are given in Appendix B.

Practice problems:

1. Evaluate each integral below:

(a) $\int (3x^2 + \sqrt{x})\,dx$

(b) $\int \frac{x^3 + 3x^2 + 1}{\sqrt{x}}\,dx$

(c) $\int_1^2 (x-1)^5\,dx$

(d) $\int \sqrt{2x-1}\,dx$

(e) $\int \frac{x^2\,dx}{x^2+1}$

(f) $\int \frac{dx}{(3+2x)^3}$

(g) $\int \sin(3x+1)\,dx$

(h) $\int \frac{\cos x\,dx}{1+\sin^2 x}$

(i) $\int \frac{x\,dx}{\sqrt{x^2+1}}$

(j) $\int_0^1 (x+1)^2(x^2+1)\,dx$

(k) $\int e^{3x-2}\,dx$

(l) $\int \ln 3x\,dx$

(m) $\int x\cos 3x\,dx$

(n) $\int \frac{dx}{\sqrt{1-4x^2}}$

(o) $\int \frac{x\,dx}{\sqrt{1-4x^2}}$

(p) $\int \frac{e^x\,dx}{e^{2x}+1}$

(q) $\int xe^{-x^2}\,dx$

(r) $\int_0^1 x^3e^{-x^2}\,dx$

(s) $\int x^2\ln x\,dx$

(t) $\int \ln x\,dx$

(u) $\int -\frac{1}{x^2}\ln\left(\frac{1}{x}\right)\,dx$

(v) $\int \frac{1}{x}(\ln x)^2\,dx$

(w) $\int \frac{1}{x\ln x}\,dx$

(x) $\int \sec^2 2x\,dx$

Theory problems:

2. Use the double-integration-by-parts trick on p. 385 to:

   (a) Evaluate $\int e^x \cos x\,dx$;

   (b) Evaluate $\int e^{2x} \sin 3x\,dx$;

   (c) Find a general formula for the indefinite integrals

   $$\int e^{ax} \sin bx\,dx \qquad \int e^{ax} \cos bx\,ddx$$

   when a and b are arbitrary nonzero constants.

3. Show that

$$\int \csc^2 x\,dx = -\cot x + C$$
$$\int \csc x \cot x\,dx = -\csc x + C.$$

History note:

4. **Leibniz's derivation of integration by parts:** [20, pp.245-7] Before he had the Fundamental Theorem of Calculus, Leibniz formulated his "transmutation theorem", which is equivalent to integration by parts. We give his derivation, referring to Figure 5.20. We wish to evaluate the area of the region $aABb$ bounded by the curve $\widehat{AB}$ with equation $y = f(x)$, the vertical lines Aa and Bb, and the horizontal line ab, given by

$$area(aABb) = \int_a^b y\,dx.$$

   (a) Let P be a point on the curve $\widehat{AB}$, with coordinates (x, y), and let T be the point where the line tangent to $\widehat{AB}$ at P crosses the y-axis; let OS be a line through the origin O perpendicular to this tangent (and meeting it at S). Denote the lengths of the line segments OS and OT, respectively, by p and z.

   (b) Construct a "characteristic triangle" $\triangle PQR$ with a horizontal leg of length dx, a vertical leg of length dy, and hypotenuse ds. Show that the triangles $\triangle PQR$ and OTS are similar.

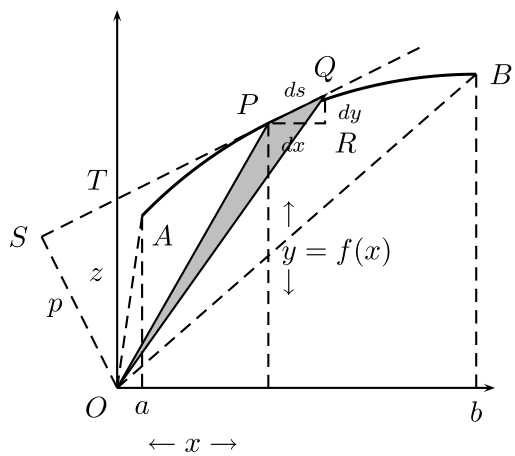


Figure 5.20: Integration by Parts after Leibniz

(c) Conclude from this that $p : z = dx : ds$, or

$$z\,dx = p\,ds.$$

(d) The triangle $\triangle OPQ$ can be regarded as having base ds and height p, so its area is given by

$$area(\triangle OPQ) = \frac{1}{2}p\,ds = \frac{1}{2}z\,dx.$$

(e) By dividing it into triangles like $\triangle OPQ$ and summing, the area of the sector $\measuredangle OAB$ is given by

$$area(\measuredangle OAB) = \frac{1}{2}\int_a^b z\,dx.$$

(f) Now, show that

$$area(aABb) + area(\triangle OaA) = area(\measuredangle OAB) + area(\triangle OBb).$$

(g) Thus,

$$\begin{aligned}\int_a^b y\,dx &= \frac{1}{2}\int_a^b z\,dx + \frac{1}{2}(Ob)\cdot(bB) - \frac{1}{2}(OA)\cdot(aA)\\ &= \frac{1}{2}\int_a^b z\,dx + \frac{1}{2}(b\cdot f(b)) - \frac{1}{2}(a\cdot f(A))\\ &= \frac{1}{2}\left((xy)\Big|_a^b + \int_a^b z\,dx \right).\end{aligned}$$

(h) Using the fact that the slope of the line TP is $\dfrac{dy}{dx}$, show that

$$z = y - x\frac{dy}{dx}.$$

(i) Substitute this into the previous equation to get the integration-by-parts formula

$$\int_a^b y\,dx = (xy)\Big|_a^b - \int_a^b x\,dy.$$

5.5 Trigonometric tricks

Sines and cosines

There are a number of integration tricks that play off trigonometric identities, like

$$\sin^2 x + \cos^2 x = 1$$

against the differentiation formulas

$$\begin{aligned} d\sin x &= \cos x\ dx\\ d\cos x &= -\sin x\ dx.\end{aligned}$$

Integrals of the form

$$\int \sin^k x \cos x\ dx \text{ or } \int \cos^k x \sin x\ dx$$

are fairly obvious candidates for the substitution trick: in the first case, the substitution $u = \sin x$ gives $du = \cos x\ dx$ and turns the integral into

$\int u^k du$. Integrals involving *both* $\sin x$ and $\cos x$ to a higher power can be simplified using trig identities. For example the integral

$$\int \cos^3 x \sin^2 x \; dx$$

can be rewritten using the identity

$$\cos^2 x = 1 - \sin^2 x$$

as

$$\begin{aligned} \int \cos^3 x \sin^2 x \; dx &= \int \cos^2 x \sin^2 x \cos x \; dx \\ &= \int (1 - \sin^2 x) \sin^2 x \cos x \; dx \end{aligned}$$

for which the substitution

$$u = \sin x, \quad du = \cos x \; dx$$

gives

$$\begin{aligned} \int (1 - \sin^2 x) \sin^2 x \cos x \; dx &= \int (1 - u^2) u^2 du \\ &= \int (u^2 - u^4) du = \frac{1}{3} u^3 - \frac{1}{5} u^5 + C \\ &= \frac{1}{3} \sin^3 x - \frac{1}{5} \sin^5 x + C. \end{aligned}$$

This idea works whenever $\cos ax$ (or $\sin ax$) appears to an *odd* power: suppose $\cos ax$ appears to an odd power. We can "peel off" one factor of $\cos ax$, attach it to dx, and let u be the other function of the pair, $\sin ax$. Then du is obtained from the peeled-off term and the rest can be written in terms of u, using the trig identity to write the remaining *even* power of $\cos ax$ in terms of $u = \sin ax$.

This trick, however, runs into trouble when *both* $\sin ax$ and $\cos ax$ appear to an *even* power. For example[22]

$$\int \sin^2 x \; dx$$

[22] $1 = cos^0 x$ is an even power of $\cos x$.

cannot be handled this way. In this situation, the **double-angle formula**

$$\cos 2\theta = \cos^2\theta - \sin^2\theta = 1 - 2\sin^2\theta$$

comes to our rescue, rewritten as

$$\sin^2\theta = \frac{1}{2}(1 - \cos 2\theta).$$

This means we can write

$$\int \sin^2 x \; dx = \frac{1}{2}\int (1 - \cos 2x) \; dx = \frac{1}{2}\int dx - \frac{1}{2}\int \cos 2x \; dx.$$

The first integral is easy, and the second is not hard (using $u = 2x$, $dx = \frac{1}{2}du$):

$$\begin{aligned}\frac{1}{2}\int dx - \frac{1}{2}\int \cos 2x \; dx &= \frac{x}{2} - \frac{1}{2}\int \cos u \frac{du}{2} \\ &= \frac{x}{2} - \frac{1}{4}\sin u + C = \frac{x}{2} - \frac{1}{4}\sin 2x + C.\end{aligned}$$

Note that the same double-angle formula also leads to

$$\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$$

which can be used in a similar way when $\cos ax$ appears with an even exponent.

Tangents and Secants

There are tricks of a similar nature which juggle the identity

$$\tan^2\theta + 1 = \sec^2\theta$$

and the differentiation formulas

$$\begin{aligned} d\tan x &= \sec^2 x \; dx \\ d\sec x &= \tan x \sec x \; dx. \end{aligned}$$

Many problems involving these functions can be treated by simply rewriting them in terms of $\sin x$ and $\cos x$. However, two annoying integrals need to be addressed directly: those of either function alone.

The integral of $\tan x$ is best treated in terms of $\sin x$ and $\cos x$:

$$\int \tan x \, dx = \int \frac{\sin x}{\cos x} dx.$$

This actually fits an earlier pattern: it involves an odd power of $\sin x$, so we make the substitution

$$u = \cos x, \quad du = -\sin x \, dx$$

to write

$$\begin{aligned}\int \frac{\sin x}{\cos x} dx = \int -\frac{du}{u} &= -\ln|u| + C = -\ln|\cos x| + C \\ &= \ln|\sec x| + C.\end{aligned}$$

(You should check that the answer is, indeed, an antiderivative of $\tan x$.) The second integral requires a more cunning plan.[23] Note that the differentiation formulas for $\tan x$ and $\sec x$ each give the other function of the pair, multiplied by $\sec x$. Thus, the derivative of their *sum* is just that sum, multiplied by $\sec x$. This means that if we write

$$\int \sec x \, dx = \int \frac{\sec x(\tan x + \sec x)}{\tan x + \sec x} \, dx$$

and note that the numerator in the second integrand is the derivative of the denominator:

$$d(\tan x + \sec x) = (\sec^2 x + \tan x \sec x) \, dx,$$

then we can substitute

$$u = \tan x + \sec x, \quad du = (\sec^2 x + \tan x \sec x) \, dx.$$

This gives

$$\int \sec x \, dx = \int \frac{du}{u} = \ln|u| + C = \ln|\tan x + \sec x| + C.$$

Integrals of the form

$$\int \tan^m x \sec^n x \, dx$$

where m and n are non-negative integers can be handled by methods similar to those for powers of $\sin x$ times powers of $\cos x$:

[23] A more natural approach, using partial fractions, is discussed in Exercise 4 of § 5.6.

1. If $\boldsymbol{n}$ **is even and positive**, we can factor out

$$\sec^2 x\,dx = d\,\tan x$$

and what is left is a power of $\tan x$ times an *even* power of $\sec x$, which can also be expressed as a polynomial in $\tan x$. Then the substitution

$$t = \tan x$$

gives us an easy integral.

For example,

$$\begin{aligned}\int \tan x \sec^2 x\,dx &= \int (\tan x)\big(d\,\tan x\big)\\ &= \frac{\tan^2 x}{2} + C.\end{aligned}$$

2. if $\boldsymbol{m}$ **is odd and** $\boldsymbol{n \geq 1}$ we can factor out

$$\tan x \sec x\,dx = d\,\sec x$$

leaving an *even* power of $\tan x$, which can be rewritten as a polynomial in $\sec x$, together with some non-negative power of $\sec x$; thus the substitution

$$s = \sec x$$

leads to an easy integral.

For example,

$$\begin{aligned}\int \tan^3 x \sec x\,dx &= \int (\tan^2 x)\big(\tan x \sec x\,dx\big)\\ &= \int (\sec^2 x - 1)\big(d\,\sec x\big)\\ &= \frac{1}{3}\sec^3 x - \sec x + C.\end{aligned}$$

3. If $\boldsymbol{m}$ **is even**, we can use the identity

$$\tan^2 x = \sec^2 - 1$$

to rewrite the integrand as a polynomial in $\sec x$. The terms involving *even* powers of $\sec x$ can be handled as in the first item above. For the *odd* powers, we can use integration by parts: parse $\int \sec^{2n+1} x\,dx$ as

$$\begin{aligned} u &= \sec^{2n-1} x, \\ dv &= \sec^2 x\,dx \\ &= d\,\tan x \end{aligned}$$

giving us

$$\begin{aligned} \int \sec^{2n+1} x\,dx &= \sec^{2n-1} x \tan x - \int \left((2n-1)\sec^{2n-2} x \frac{d}{dx}\left[\sec x\right]\right) \tan x\,dx \\ &= \sec^{2n-1} x \tan x - (2n-1) \int \sec^{2n-1} x \tan^2 x\,dx \\ &= \sec^{2n-1} x \tan x - (2n-1) \int \sec^{2n-1} x (\sec^2 x - 1)\,dx \\ &= \sec^{2n-1} x \tan x - (2n-1) \left(\int \sec^{2n+1} x\,dx - \int \sec^{2n-1} x\,dx \right) \\ &= \left(\sec^{2n-1} x \tan x + (2n-1) \int \sec^{2n-1} x\,dx \right) - (2n-1) \int \sec^{2n+1} x\,dx \end{aligned}$$

and using the usual algebraic trick, we have the reduction formula

$$\int \sec^{2n+1} x\,dx = \frac{1}{2n} \left(\sec^{2n-1} x \tan x + (2n-1) \int \sec^{2n-1} x\,dx \right).$$

In the case $n = 1$, this gives the formula

$$\int \sec^3 x\,dx = \frac{1}{2} \left(\sec x \tan x + \int \sec x\,dx \right)$$

and using our earlier calculation, this means

$$\int \sec^3 x\,dx = \frac{1}{2} \left(\sec x \tan x + \ln(\sec x + \tan x) \right) + C.$$

Applications to Algebraic Integrals

We can also use trigonometric substitutions to handle certain algebraic integrals. For example, the integral

$$\int \sqrt{1 - x^2}\,dx$$

looks hopeless, unless we note that the substitution

$$x = \sin\theta$$

gives

$$dx = \cos\theta\, d\theta$$

and

$$1 - x^2 = 1 - \sin^2\theta = \cos^2\theta,$$

so

$$\sqrt{1 - x^2} = \pm\cos\theta.$$

In fact, if $-\pi/2 \leq \theta \leq \pi/2$, the last equation has a "+". But this is the range of the inverse sine function, so if we interpret our substitution as

$$\theta = \arcsin x$$

(*i.e.*, $\theta \in [-\pi/2, \pi/2]$), we can write

$$\int \sqrt{1 - x^2}\, dx = \int (\cos\theta)(\cos\theta\, d\theta) = \int \cos^2\theta\, d\theta.$$

The last integral involves even powers of $\sin\theta$ and $\cos\theta$, so we use the double-angle identity

$$\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$$

and write

$$\int \sqrt{1 - x^2}\, dx = \int \cos^2\theta\, d\theta = \frac{1}{2}\int (1 + \cos 2\theta)\, d\theta.$$

But this integral is easy to evaluate:

$$\begin{aligned} \frac{1}{2}\int (1 + \cos 2\theta)\, d\theta &= \frac{1}{2}\int d\theta + \frac{1}{2}\int \cos 2\theta\, d\theta \\ &= \frac{\theta}{2} + \frac{1}{4}\sin 2\theta + C. \end{aligned}$$

Finally, we need to rewrite our answer using x. The first term is easy: our substitution immediately gives

$$\frac{\theta}{2} = \frac{1}{2}\arcsin x.$$

For the second, we use more trig identities:

$$\sin 2\theta = 2\sin\theta\cos\theta$$

means
$$\frac{1}{4}\sin 2\theta = \frac{1}{2}\sin\theta\cos\theta = \frac{1}{2}x\sqrt{1-x^2}.$$
So we have
$$\int \sqrt{1-x^2}\,dx = \frac{1}{2}\arcsin x + \frac{1}{2}x\sqrt{1-x^2} + C.$$
We note that the same substitution would have worked to find the formula (which we already know from differentiation)
$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C.$$
The basic observation here is that $x = \sin\theta$ ($\theta = \arcsin x$) leads to $\sqrt{1-x^2} = \cos\theta$ and $dx = \cos\theta\,d\theta$ and turns certain integrals involving $\sqrt{1-x^2}$ into integrals in $\sin\theta$ and $\cos\theta$.

Similarly, the substitution $x = \tan\theta$ ($\theta = \arctan x$) can be used to change $1+x^2$ into $\sec\theta$ and dx into $\sec^2\theta\,d\theta$. This could be used to deduce the known formula
$$\int \frac{dx}{1+x^2} = \arctan x + C,$$
as well as other integrals, by turning them into integrals involving $\tan\theta$ and $\sec\theta$.

We explore a slight variant of this trick. The integral
$$\int \frac{dx}{x^2+4}$$
is *not* quite of the type above: we need x^2+1 to take advantage of $x = \tan\theta$. However, if we factor out the 4
$$\int \frac{dx}{x^2+4} = \frac{1}{4}\int \frac{dx}{\frac{x^2}{4}+1},$$
the substitution $u = \frac{x}{2}$ gives $dx = 2\,du$ and
$$\frac{1}{4}\int \frac{dx}{\frac{x^2}{4}+1} = \frac{1}{4}\int \frac{2\,du}{u^2+1} = \frac{1}{2}\arctan u + C$$
$$= \frac{1}{2}\arctan\frac{x}{2} + C.$$
An alternate approach to this is to note that
$$x = 2\tan\theta \quad (\theta = \arctan\frac{x}{2})$$

means

$$\begin{aligned} x^2 + 4 &= 4\tan^2\theta + 4 = 4(\tan^2\theta + 1) = 4(\sec^2\theta) \\ dx &= 2\sec^2\theta\, d\theta \end{aligned}$$

so

$$\begin{aligned} \int \frac{dx}{x^2+4} &= \int \frac{2\sec^2\theta\, d\theta}{4\sec^2\theta} = \frac{1}{2}\int d\theta = \frac{\theta}{2} + C \\ &= \frac{1}{2}\arctan\frac{x}{2} + C. \end{aligned}$$

Exercises for § 5.5

Answers to Exercises 1-2acegi, 3ace, 4acegi, 5 are given in Appendix B.

Practice problems:

1. Evaluate the following integrals:

   (a) $\int \sin x \cos x\, dx$ (b) $\int_0^{\pi/2} \sin^2 x \cos x\, dx$

   (c) $\int \sin^3 x \cos^2 x\, dx$ (d) $\int_{\pi/8}^{\pi/4} \sin^2 x \cos^2 x\, dx$

   (e) $\int \sin^4 x\, dx$ (f) $\int_0^{\pi/2} \sin^3 x \cos^3 x\, dx$

   (g) $\int \frac{\sin 2x}{\cos^2 x}\, dx$ (h) $\int \frac{\cos 2x}{\cos x}\, dx$

   (i) $\int \sin x \tan x\, dx$

2. Evaluate the following integrals:

   (a) $\int \sec^4 x\, dx$ (b) $\int \sec^5 x\, dx$

   (c) $\int \tan x \sec x\, dx$ (d) $\int \tan^2 x \sec x\, dx$

   (e) $\int \tan^2 x \sec^2 x\, dx$ (f) $\int \tan^2 \sec^3 x\, dx$

   (g) $\int \tan x \sec^3 x\, dx$ (h) $\int \tan^3 x \sec^2 x\, dx$

(i) $\int \tan x \sec^4 x \, dx$

3. Evaluate the following integrals:

(a) $\int \sqrt{4-x^2}\, dx$ (b) $\int_0^{\frac{1}{2}} (1-4x^{2)^{3/2}}\, dx$

(c) $\int \frac{\sqrt{4-x^2}}{x}\, dx$ (d) $\int \frac{x\, dx}{\sqrt{x^2+1}}$

(e) $\int \frac{dx}{\sqrt{x^2+1}}$ (f) $\int_1^2 \sqrt{x^2-1}\, dx$

4. Evaluate each integral below:

(a) $\int x \cos 2x \, dx$ (b) $\int x \sin^2 x \, dx$

(c) $\int x^2 \sin^2 x \, dx$ (d) $\int x \sin^3 2x \, dx$

(e) $\int x \sec^2 2x \, dx$ (f) $\int x \arctan x \, dx$

(g) $\int \arcsin x \, dx$ (h) $\int x^2 \arcsin x \, dx$

(i) $\int \sqrt{1-x^2} \arcsin x \, dx$

Theory problems:

5. (a) Use a double integration-by-parts (as on p. 385) to evaluate

$$\int \sin 3x \cos 4x \, dx.$$

(b) Use the same trick to evaluate

$$\int \sin ax \cos bx \, dx$$

for any pair of constants $a^2 \neq b^2$.

(c) Show that

$$\cos ax \cos bx = \frac{1}{2}(\cos(a-b)x + \cos(a+b)x)$$
$$\sin ax \sin bx = \frac{1}{2}(\cos(a-b)x - \cos(a+b)x)$$

and use this to rewrite your answer to part (b).

(d) What happens when $a = \pm b$?

6. Use integration by parts to derive the reduction formula

$$\int \cos^n x\,dx = \frac{1}{n}\cos^{n-1} x \sin x + \frac{n-1}{n}\int \cos^{n-2} x\,dx.$$

5.6 Integration of Rational Functions

In this section we deal with a general method for integrating rational functions.

Basic Integrands

Note first that an integral of the form

$$\int \frac{dx}{x+a}$$

can easily be evaluated by means of the substitution

$$u = x + a, \quad du = dx$$

to give

$$\int \frac{dx}{x+a} = \int \frac{du}{u} = \ln|u| + C = \ln|x+a| + C.$$

If we replace $x + a$ by any first-degree polynomial, then factoring out the coefficient of x leads back to this. Also, any rational function whose denominator is of degree one can be written as a polynomial plus a multiple of the type of function treated above: for example

$$\frac{x^2 + 3x + 4}{x+1} = x + 2 + \frac{2}{x+1}$$

so that

$$\begin{aligned}\int \frac{(x^2 + 3x + 4)dx}{x+1} &= \int (x+2)dx + \int \frac{2}{x+1}dx \\ &= \frac{x^2}{2} + 2x + 2\ln|x+1| + C.\end{aligned}$$

What about quadratic denominators? We saw earlier that the integral

$$\int \frac{dx}{x^2+4}$$

could be transformed by the change of variables

$$x = 2\tan\theta \ (\theta = \arctan\frac{x}{2}), \quad dx = 2\sec^2\theta \ d\theta$$

into a trig integral which we can evaluate:

$$\int \frac{2\sec^2\theta \ d\theta}{4(\tan^2\theta + 1)} = \frac{1}{2}\int d\theta = \frac{\theta}{2} + C = \frac{1}{2}\arctan\frac{x}{2} + C.$$

This same kind of trick works if the denominator can be written as the square of a first-degree polynomial plus a perfect square constant: for example

$$\int \frac{dx}{(x+2)^2 + 4}$$

can be turned into the previous kind of integral via the preliminary change of variables

$$u = x + 2, \quad du = dx$$

to get

$$\int \frac{dx}{(x+2)^2 + 4} = \int \frac{du}{u^2 + 4}$$

from which we can proceed as above.
A quadratic polynomial is **irreducible** if it can be written this way. To decide whether a quadratic is irreducible or not, we can use the device of **completing the square**. For example, to decide whether

$$4x^2 - 24x + 45$$

is irreducible, we factor out the leading coefficient from the first two terms

$$4(x^2 - 6x) + 45,$$

then ask what would turn the quantity inside the parentheses into a perfect square. Recall that

$$(x + a)^2 = x^2 + 2ax + a^2$$

so to match $x^2 - 6x + \ldots$ we need $a = -3$, and hence need to add $(-3)^2 = 9$ to the terms inside the parentheses. But adding 9 *inside* the parentheses amounts to adding $4 \times 9 = 36$ *outside*; in order not to change the total value of our expression, we need to subtract this amount off at the end:

$$4(x^2 - 6x) + 45 = 4(x^2 - 6x + 9) + 45 - 36 = 4(x - 3)^2 + 9.$$

Using this we can integrate the reciprocal of our polynomial as follows: first use the calculation above to rewrite it

$$\int \frac{dx}{4x^2 - 24x + 45} = \int \frac{dx}{4(x-3)^2 + 9}$$
$$= \frac{1}{9} \int \frac{dx}{\frac{4}{9}(x-3)^2 + 1} = \frac{1}{9} \int \frac{dx}{\left(\frac{2}{3}(x-3)\right)^2 + 1}.$$

Now, set

$$\frac{2}{3}(x-3) = \tan\theta \; (\theta = \arctan\left(\frac{2}{3}(x-3)\right)), \quad dx = \frac{3}{2}\sec^2\theta \; d\theta$$

to obtain

$$\int \frac{dx}{4x^2 - 24x + 45} = \frac{1}{9} \int \frac{\frac{3}{2}\sec^2\theta \; d\theta}{\tan^2\theta + 1} = \frac{1}{6} \int d\theta$$
$$= \frac{1}{6}\theta + C = \frac{1}{6}\arctan\left(\frac{2}{3}(x-3)\right) + C.$$

A third kind of straightforward integral occurs when the numerator is the derivative of the denominator: thus

$$\int \frac{(8x-24)\,dx}{4x^2 - 24x + 45}$$

can be evaluated via the change of variables

$$u = 4x^2 - 24x + 45, \quad du = (8x - 24)\,dx$$

to yield

$$\int \frac{(8x-24)dx}{4x^2 - 24x + 45} = \int \frac{du}{u} = \ln|u| + C = \ln(4x^2 - 24x + 45) + C.$$

Since the derivative of a quadratic is always a first-degree polynomial, we can in principle combine the preceding tricks to handle a first-degree polynomial divided by an irreducible quadratic. For example,

$$\int \frac{2x+1}{4x^2 - 24x + 45}\,dx$$

can be prepared as follows: we know that the derivative of the denominator is $8x - 24$. Dividing this by 4 gives us the correct leading coefficient so that the numerator differs from this by a constant:

$$2x + 1 = \frac{1}{4}(8x - 24) + \alpha = (2x - 6) + \alpha$$

can be solved to give $\alpha = 7$, so

$$\int \frac{2x+1}{4x^2-24x+45}\,dx = \frac{1}{4}\int \frac{(8x-24)dx}{4x^2-24x+45} + 7\int \frac{dx}{4x^2-24x+45}.$$

We did both of the right-hand integrals already, so we can compute

$$\int \frac{2x+1}{4x^2-24x+45}\,dx = \frac{1}{4}\ln(4x^2-24x+45) + \frac{7}{6}\arctan\left(\frac{2}{3}(x-3)\right) + C.$$

What if the denominator is *not* irreducible? We will discover this (for a quadratic) if we attempt completion of the square and end up with a non-positive constant on the right. For example, applying the process to

$$4x^2 - 24x + 11$$

we get

$$4(x-3)^2 + 11 - 36 = 4(x-3)^2 - 25.$$

This is a *difference* of squares, which we know how to factor:

$$\begin{aligned} 4(x-3)^2 - 25 &= \left(2(x-3)^2\right)^2 - [5]^2 \\ &= (2(x-3)-5)\,(2(x-3)+5) \\ &= (2x-11)\,(2x-1)\,. \end{aligned}$$

How does this help us integrate the reciprocal? The answer to this involves a detour into algebra, to develop the *Partial Fraction Decomposition* of a rational function. Before diving into the algebra, we will look at an analogue from arithmetic, to help us fix ideas.

Partial Fractions for Rational Numbers

A rational number can be represented either as a single fraction or as a decimal expression. Now, we can think of a decimal expression as a sum of "simple" fractions: each denominator is a power of 10, and each numerator is one of the ten integers 0,1,2,3,4,5,6,7,8, or 9. Unfortunately, when the denominator of the single-fraction representation of a rational (like $\frac{1}{3}$) does not divide some power of 10, the corresponding decimal expression is an *infinite* sum. There is, however, a way to represent any rational number as a *finite* sum of fractions[24], each of which is simple in the sense that its denominator is a power of some prime number p, and its numerator is one of the integers with absolute value strictly less than p.

[24] I thank my colleague George McNinch for useful conversations concerning this topic.

Theorem 5.6.1 (Partial Fractions for Rational Numbers)**.** *Every rational number*

$$r = A/B$$

can be written as an integer plus a finite sum of fractions

$$\frac{a}{p^i}$$

where

1. *p is a prime factor of B with multiplicity $m \geq i$, and*
2. *$|a| < p$.*

We will sketch a proof of this theorem based on two initial simplifications and two decomposition principles, leaving the details to you (Exercise 7 and Exercise 8).
The initial simplifications are the obvious ones:

1. If $|A| \geq |B|$, we can divide through and rewrite A/B as an integer plus a **proper** fraction: that is, one whose numerator is strictly less than its denominator.
2. Supposing that $|A| < |B|$, we can cancel any common divisors of A and B to put the fraction in **lowest terms**—that is, A and B are relatively prime.

So to establish Theorem 5.6.1, we can restrict attention to proper fractions in lowest terms.
Our first decomposition principle is an application of the *Euclidean algorithm*; the proof is left to you (Exercise 7):

Lemma 5.6.2. *Suppose*

$$B = b_1 b_2$$

where b_1 and b_2 are relatively prime integers, and A/B is a proper fraction in lowest terms.
Then there exist nonzero integers a_1 and a_2 such that

$$\frac{A}{B} = \frac{a_1}{b_1} + \frac{a_2}{b_2}$$

where for each $i = 1, 2$

1. *a_i and b_i are relatively prime*

2. $0 < |a_i| < b_i$.

For example, with $A = 43$, $B = 120 = 8 \cdot 15$ ($b_1 = 8$, $b_2 = 15$) we can take $a_1 = 5$ and $a_2 = -4$:

$$\frac{43}{120} = \frac{5}{8} - \frac{4}{15}.$$

Of course, we can also apply Lemma 5.6.2 with $A = 4$ and $B = 15 = 3 \cdot 5$ to further decompose the second term above and write

$$\frac{43}{120} = \frac{5}{8} - \left(\frac{2}{3} - \frac{2}{5}\right) = \frac{5}{8} - \frac{2}{3} + \frac{2}{5}.$$

The second and third terms in this decomposition have the from q/p^k required by Theorem 5.6.1, but the first term, whose denominator is $8 = 2^3$, has denominator greater than $p = 2$. However, we can express 5 as a sum of powers of 2

$$\begin{aligned} 5 &= 1 + 4 \\ &= 2^0 + 2^2 \end{aligned}$$

so that

$$\begin{aligned} \frac{5}{8} &= \frac{2^0 + 2^2}{2^3} \\ &= \frac{1}{2^3} + \frac{1}{2^1} \end{aligned}$$

which has the form required by Theorem 5.6.1.

This last step illustrates the second decomposition principle:

Lemma 5.6.3. *A proper fraction A/B in lowest terms with*

$$|A| < B = p^m, \quad p \text{ prime}$$

can be decomposed as a sum of fractions of the form

$$\frac{A}{B} = \frac{q_1}{p^1} + \frac{q_2}{p^2} + \cdots + \frac{q_m}{p^m}$$

with

$$0 \le |q_i| < p, \quad i = 1, \ldots, m.$$

The proof of this lemma rests on the *base p expansion* of the numerator A; we leave it to you (Exercise 8).

In our example, then, the required decomposition is

$$\frac{43}{120} = \frac{1}{2^3} + \frac{1}{2^1} - \frac{2}{3^1} + \frac{2}{5^1}.$$

Partial Fractions for Rational Functions

The arithmetic of integers and the algebra of polynomials share many formal properties. Besides obeying the associative, commutative and distributive laws, the collection of polynomials in one variable with real coefficients (often denoted $\mathbb{R}[x]$) obeys an analogue of the prime factorization of integers. This is known as the **Fundamental Theorem of Algebra**, observed from the early seventeenth century and first proved by Carl Friedrich Gauss (1777-1855) in his doctoral dissertation[25] of 1799. It can be stated as follows:

Theorem 5.6.4 (Fundamental Theorem of Algebra)**.** *Every polynomial*

$$f(x) = a_N x^N + a_{N-1} x^{N-1} + \cdots + a_1 x + a_0$$

with real coefficients has a representation as a product of first-degree polynomials and irreducible quadratic polynomials.

A polynomial of degree one

$$\ell(x) = ax + b \quad (a \neq 0)$$

can also be written as

$$\ell(x) = a(x - r)$$

where $r = -b/a$ is its (unique) real root. A quadratic polynomial

$$q(x) = ax^2 + bx + c$$

is **irreducible** if it has no real roots, or equivalently if $b^2 < 4ac$. In this case, it has a pair of *complex* roots $\alpha \pm \beta i$ and can be written[26] in the form

$$q(x) = a\{(x - \alpha)^2 + \beta^2\}.$$

These are the "prime" polynomials, in the sense that they are divisible only by nonzero constants and (nonzero) constant multiples of themselves. A real root $x = r$ of the polynomial $f(x)$ has **multiplicity** m if

$$f(x) = (x - r)^m \tilde{f}(x)$$

[25] Gauss gave four proofs of this result over a space of 50 years, the last in 1849. *cf* [11, pp.511-2]

[26] by completing the square

where

$$\tilde{f}(r) \neq 0;$$

this is equavalent to saying that exactly m of the first-degree polynomial factors of $f(x)$ are multiples of $(x-r)$. Analogously, an irreducible quadratic $q(x)$ has multiplicity m in $f(x)$ if the factoring of $f(x)$ involves exactly m (nonzero) constant multiples of it.
A rational function

$$f(x) = \frac{A(x)}{B(x)}, \quad A(x)\,,\; B(x) \text{ polynomials}$$

is **proper** if the degree of the numerator $A(x)$ is strictly less than that of the denominator $B(x)$. The rational expression $A(x)\,/B(x)$ is in **lowest terms** if $A(x)$ and $B(x)$ have no common first-degree or irreducible divisors, or equivalently, they share no (real or complex) roots.
The analogue of Theorem 5.6.1 is

Theorem 5.6.5 (Partial Fractions for Rational Functions)**.** *Every rational function*

$$f(x) = \frac{A(x)}{B(x)}$$

(in lowest terms) can be represented as a polynomial plus a sum of rational functions of the form

$$\frac{a(x)}{p(x)^k}$$

where

1. *$p(x)$ is a first-degree or irreducible quadratic polynomial factor of $B(x)$, and*

2. *the degree of $a(x)$ is strictly less than the degree of $p(x)$.*

This says that our sum contains terms of the following three kinds:

1. If $\deg A(x) \geq \deg B(x)$, we start with a polynomial.

2. If $p(x) = (x-r)$ is a factor of $B(x)$ (*i.e.*, r is a real root of $B(x)$) with multiplicity m, then we have summands of the form

$$\frac{a_j}{(x-r)^j}$$

for $j = 1, \ldots, m$.

3. If $p(x) = \{(x-\alpha)^2 + \beta^2\}$ is an irreducible quadratic factor of $B(x)$ (*i.e.*, $\alpha \pm \beta i$ is a pair of complex roots of $B(x)$) with mutliplicity m, then we have summands of the form

$$\frac{a_j x + b_j}{p(x)^j}$$

for $j = 1, \ldots, m$.

The proof of Theorem 5.6.5 rests on two lemmas:

Lemma 5.6.6. *Suppose $A(x)/B(x)$ is proper and in lowest terms and $x = r$ is a real root of $B(x)$ with multiplicity m—that is,*

$$B(x) = (x-r)^m B_1(x), \; B_1(r) \neq 0.$$

Then

$$\frac{A(x)}{B(x)} = \frac{a}{(x-r)^m} + \frac{\hat{A}(x)}{\hat{B}(x)}$$

where

1. *$a = A(r)/B(r)$;*
2. *$\hat{A}(x)/\hat{B}(x)$ is proper and in lowest terms;*
3. *$x = r$ is a root of $\hat{B}(x)$ with multiplicity $\hat{m} < m$ (including the possibility that $\hat{m} = 0$, which means $\hat{B}(r) \neq 0$).*

Proof. Let $a \in \mathbb{R}$ be a constant (as yet unspecified), and consider

$$\begin{aligned} \frac{A(x)}{B(x)} - \frac{a}{(x-r)^m} &= \frac{A(x)}{(x-r)^m B_1(x)} - \frac{aB_1(x)}{(x-r)^m B_1(x)} \\ &= \frac{A(x) - aB_1(x)}{(x-r)^m B_1(x)}. \end{aligned}$$

Since $B_1(r) \neq 0$, we can set

$$a = \frac{A(r)}{B_1(r)}$$

in which case

$$A(r) - aB_1(r) = 0$$

and we can write

$$A(x) - aB_1(x) = (x - r)A_1(x),$$

so

$$\frac{A(x)}{B(x)} = \frac{a}{(x-r)^m B_1(x)} + \frac{A_1(x)}{(x-r)^{m-1} B_1(x)}.$$

The degree of $A_1(x)$ (*resp.* of $(x-r)^{m-1}B_1(x)$) is one less than that of $A(x)$ (*resp.* of $B(x)$), so the last fraction above is proper. If it is not in lowest terms, we cancel common factors to obtain

$$\frac{\hat{A}(x)}{\hat{B}(x)}.$$

This may lower the multiplicity of $x = r$ as a root of the denominator, but in any case the resulting rational function will satisfy conditions (2) and (3) above. □

The analogous result for irreducible quadratic factors of $B(x)$ is

Lemma 5.6.7. *Suppose $A(x)/B(x)$ is proper and in lowest terms, and that $q(x)$ is an irreducible quadratic factor of $B(x)$ with multiplicity m. Then there exist $a, b \in \mathbb{R}$ such that*

$$\frac{A(x)}{B(x)} = \frac{ax+b}{q(x)^m} + \frac{\hat{A}(x)}{\hat{B}(x)}$$

where

1. *$\hat{A}(x)/\hat{B}(x)$ is in lowest terms;*
2. *Either $q(x)$ does not divide $\hat{B}(x)$, or its multiplicity as a divisor of $\hat{B}(x)$ is strictly less than m.*

The clearest and most efficient proof of this result mimics the proof of Lemma 5.6.6, using the two complex roots of $q(x)$ in place of the real root r of $(x - r)$. We leave this to Exercise 10, since it requires some knowledge of complex arithmetic.

Proof of Theorem 5.6.5. Given $A(x)/B(x)$ in lowest terms, we proceed as follows:
First, if $\deg A(x) \geq \deg B(x)$, we divide through, obtaining

$$A(x) = P(x)\,B(x) + A_1(x)$$

where $P(x)$ is a polynomial and $\deg A_1(x) < \deg B(()\,x)$. Thus,

$$\frac{A(x)}{B(x)} = P(x) + \frac{A_1(x)}{B(x)}$$

with $A_1(x)/B(x)$ proper and in lowest terms.
In light of this, we can henceforth assume without loss of generality that $A(x)\,/B(x)$ is proper and in lowest terms.
Next, we prove

> **Claim:** *If $A(x)\,/B(x)$ is proper and in lowest terms, and $x = r$ is a real root of $B(x)$ with multiplicity m, then there exist $a_1, ..., a_m \in \mathbb{R}$ such that*
>
> $$\frac{A(x)}{B(x)} = \frac{a_1}{(x-r)^1} + \frac{a_2}{(x-r)^2} \cdots + \frac{a_m}{(x-r)^m} + \frac{\hat{A}(x)}{\hat{B}(x)}$$
>
> *where*
>
> 1. *$\hat{A}(x)/\hat{B}(x)$ is proper and in lowest terms*
> 2. *$\hat{B}(r) \neq 0$*
> 3. *$\hat{B}(x)$ divides $B(x)$ (the factors of $\hat{B}(x)$ are the factors of $B(x)$ except for $(x-r)$, and they have the same multiplicity in both).*

We prove this claim by induction (see Appendix A) on the multiplicity m of r as a root of the denominator:

Initial step: If r is a simple root ($m = 1$) of $B(x)$, then

$$B(x) = (x-r)B_1(x)$$

with $B_1(r) \neq 0$, and Lemma 5.6.6 says that

$$\frac{A(x)}{B(x)} = \frac{a}{x-r} + \frac{\hat{A}(x)}{\hat{B}(x)}$$

where the second term is proper and in lowest terms and $\hat{B}(r) \neq 0$.

Induction step: Suppose we know the claim holds for all instances where a real root has multiplicity $\hat{m} < m$, and $x = r$ is a real root of $B(x)$ with multiplicity m. Lemma 5.6.6 says that

$$\frac{A(x)}{B(x)} = \frac{a_m}{(x-r)^m} + \frac{A_1(x)}{B_1(x)}$$

where $B_1(x)$ divides $B(x)$ and either $B_1(r) \neq 0$ or r is a root of $B_1(x)$ with multiplicity $\hat{m} < m$. If $B_1(r) \neq 0$, we are done; if not, the claim for multiplicity $\hat{m}$ says that $A_1(x)/B_1(x)$ is a sum of terms of the form

$$\frac{a_j}{(x-r)^j}, \quad j < m$$

plus $\hat{A}(x)/\hat{B}(x)$, and we are done.

This completes the induction proof, and proves the claim for all multiplicities m. $\Diamond$

In a similar way, we can prove

Claim: *If $A(x)/B(x)$ is proper and in lowest terms and $q(x)$ is an irreducible quadratic factor of $B(x)$ with multiplicity m, then there exist numbers $a_j, b_j \in \mathbb{R}$, $j = 1, \ldots, m$, such that*

$$\frac{A(x)}{B(x)} = \frac{a_1x + b_1}{q(x)^1} + \frac{a_2x + b_2}{q(x)^2} \cdots + \frac{a_m + b_m}{q_x^m} + \frac{\hat{A}(x)}{\hat{B}(x)}$$

where $\hat{B}(x)$ divides $B(x)$ and does not have $q(x)$ as a factor.

This follows from Lemma 5.6.7 via an induction argument much like the preceding.

Finally, we prove the full theorem by induction on the degree of $B(x)$, or equivalently note that each step as above reduces the degree of the denominator in the undecomposed term, and eventually the process stops. $\square$

Using Partial Fractions to Integrate Rational Functions

The integration of a rational function using Theorem 5.6.5 involves two steps, assuming that the denominator has been factored into first-degree and irreducible quadratic polynomials: first we have to find the (coefficients in the) terms of the partial fraction decomposition, and second we have to integrate each term. The most natural way to perform the first step is to write down the form of the sum as predicted by the theorem (with unknown coefficients), then combine it over a common denominator (which should match the denominator of the given function) and match coefficients in the numerator. We discuss some other tricks for doing this in Exercise 9 and Exercise 6. The second step can always be performed by variants of the techniques used for the basic integrals at the beginning of this section. We illustrate with several examples.

1. To integrate
$$\int \frac{dx}{4x^2 - 24x + 11}$$
we recall the factoring of the denominator
$$4x^2 - 24x + 11 = (2x - 11)(2x - 1).$$
Since the denominator has two simple real roots, the decomposition has the form
$$\frac{1}{(2x-11)(2x-1)} = \frac{a_1}{2x-11} + \frac{a_2}{2x-1}.$$
Now, the right side can be put over a common denominator (which is the denominator on the left) by cross-multiplying:
$$\begin{aligned}\frac{1}{(2x-11)(2x-1)} &= \frac{a_1(2x-1) + a_2(2x-11)}{(2x-11)(2x-1)} \\ &= \frac{(2a_1 + 2a_2)x + (-a_1 - 11a_2)}{(2x-11)(2x-1)}.\end{aligned}$$
We want the *numerators* to match; that is, we need
$$1 = (2a_1 + 2a_2)x + (-a_1 - 11a_2)$$
identically in x: but then the coefficients of x on either side[27] must agree, as must the constant terms:
$$\begin{aligned} 0 &= 2a_1 + 2a_2 && \text{(coefficient of } x\text{)} \\ 1 &= -a_1 - 11a_2 && \text{(constant term).}\end{aligned}$$
The solution of these equations is
$$a_1 = \frac{1}{10}, \quad a_2 = -\frac{1}{10}$$
from which we conclude that
$$\frac{1}{(2x-11)(2x-1)} = \frac{1}{10}\left(\frac{1}{2x-11}\right) - \frac{1}{10}\left(\frac{1}{2x-1}\right).$$
But then
$$\begin{aligned}\int \frac{dx}{4x^2 - 24x + 11} = \int \frac{dx}{(2x-11)(2x-1)} &= \frac{1}{10}\int \frac{dx}{2x-11} - \frac{1}{10}\int \frac{dx}{2x-1} \\ &= \frac{1}{20}\ln|2x-11| - \frac{1}{20}\ln|2x-1| + C.\end{aligned}$$

[27] The coefficient of x on the left is 0.

2. To integrate

$$\int \frac{dx}{4x^2 - 24x + 36}$$

we note that

$$\begin{aligned} 4x^2 - 24x + 36 &= 4(x-3)^2 + 36 - 36 = 4(x-3)^2 \\ &= (2(x-3))^2 , \end{aligned}$$

so

$$\int \frac{dx}{4x^2 - 24x + 36} = \int \frac{dx}{(2(x-3))^2}.$$

Thus the integrand

$$\frac{1}{4x^2 - 24x + 36} = \frac{1}{(2(x-3))^2}$$

is already decomposed into its partial fraction form.

To integrate this, we use the substitution

$$u = 2(x-3), \quad du = 2dx$$

which gives

$$\begin{aligned} \int \frac{dx}{(2(x-3))^2} &= \int \frac{du/2}{u^2} = \frac{1}{2} \int u^{-2} du \\ &= -\frac{1}{2} u^{-1} + C \\ &= -\frac{1}{2\,(2(x-3))} + C. \end{aligned}$$

3. The integration of

$$\int \frac{(4x+2)\,dx}{(x+1)^2(x^2+1)}$$

is a little more involved.

(a) The factoring of the denominator leads to a decomposition of the form

$$\frac{4x+2}{(x+1)^2(x^2+1)} = \frac{a}{(x+1)^2} + \frac{b}{x+1} + \frac{cx+d}{x^2+1}.$$

Combining terms on the right, we want

$$\begin{aligned}4x+2 &= a(x^2+1)+b(x+1)(x^2+1)+(cx+d)(x+1)^2\\ &= (b+c)x^3+(a+b+2c+d)x^2+(b+c+2d)x+(a+b+d)\end{aligned}$$

identically in x. This means the coefficients need to match:

$$\begin{aligned}0 &= b+c && \text{(coefficient of } x^3\text{)}\\ 0 &= a+b+2c+d && \text{(coefficient of } x^2\text{)}\\ 4 &= b+c+2d && \text{(coefficient of } x\text{)}\\ 2 &= a+b+d && \text{(constant term).}\end{aligned}$$

The solution of these equations is

$$a=-1, \quad b=1, \quad c=-1, \quad d=2$$

so that the partial fraction decomposition is

$$\frac{4x+2}{(x+1)^2(x^2+1)} = -\frac{1}{(x+1)^2}+\frac{1}{x+1}+\frac{-x+2}{x^2+1}.$$

(b) To integrate this function, we calculate the integral of each term separately.

i. The change of variables

$$u=x+1, \quad du=dx$$

transforms the integrals of the first two terms, giving

$$\begin{aligned}-\int\frac{dx}{(x+1)^2} &= -\int\frac{du}{u^2}=u^{-1}+C = \frac{1}{x+1}+C\\ \int\frac{dx}{x+1} &= \int\frac{du}{u} = \ln|u|+C=\ln|x+1|+C.\end{aligned}$$

ii. For the last integral, we note that the derivative of x^2+1 is $2x$, so

$$-\int\frac{(x-2)\,dx}{x^2+1} = -\frac{1}{2}\int\frac{2x\,dx}{x^2+1}+2\int\frac{dx}{x^2+1}.$$

- The first of these integrals is transformed via

$$u=x^2+1, \quad du=2x\,dx$$

into

$$-\frac{1}{2}\int\frac{2x\,dx}{x^2+1} = -\frac{1}{2}\int\frac{du}{u} = -\frac{1}{2}\ln|u|+C = -\frac{1}{2}\ln\left|x^2+1\right|+C.$$

- The second can be done via $x = \tan\theta$, or simply by recognizing it as one of our basic formulas

$$2\int \frac{dx}{x^2+1} = 2\arctan x + C.$$

Combining everything, we get

$$\begin{aligned}\int \frac{4x+2}{(x+1)^2(x^2+1)}\,dx &= \frac{1}{x+1} + \ln|x+1| - \frac{1}{2}\ln(x^2+1) + 2\arctan x + C \\ &= \frac{1}{x+1} + \ln\frac{|x+1|}{\sqrt{x^2+1}} + 2\arctan x + C.\end{aligned}$$

4. Finally, consider the integral

$$\int \frac{(2x^3-4x^2+6x)\,dx}{(x^2-2x+2)^2}.$$

(a) By completing the square, we easily see that

$$x^2-2x+2 = (x-1)^2+1$$

so this is an irreducible quadratic. Since our integrand is proper (check this!), its partial fraction decomposition has the form

$$\frac{2x^3-4x^2+6x}{(x^2-2x+2)^2} = \frac{ax+b}{(x^2-2x+2)} + \frac{cx+d}{(x^2-2x+2)^2}.$$

Putting everything over a common denominator we have

$$\begin{aligned}&\frac{(ax+b)(x^2-2x+2)+cx+d}{(x^2-2x+2)^2} \\ &\quad = \frac{ax^3+(b-2a)x^2+(2a-2b)x+2b+cx+d}{(x^2-2x+2)^2} \\ &\quad\quad = \frac{ax^3+(b-2a)x^2+(2a-2b+c)x+(2b+d)}{(x^2-2x+2)^2}.\end{aligned}$$

Matching coefficients with the given numerator, we get the equations

$$\begin{aligned} 2 &= a && \text{(coefficient of } x^3) \\ -4 &= -2a + b && \text{(coefficient of } x^2) \\ 6 &= 2a - 2b + c && \text{(coefficient of } x) \\ 0 &= 2b + d && \text{(constant term)} \end{aligned}$$

with solution

$$a = 2, \quad b = 0, \quad c = 2, \quad d = 0$$

so the partial fraction decomposition is

$$\frac{2x^3 - 4x^2 + 6x}{(x^2 - 2x + 2)^2} = \frac{2x}{(x^2 - 2x + 2)} + \frac{2x}{(x^2 - 2x + 2)^2}.$$

(b) We integrate the two fractions separately.

i. To integrate the first fraction

$$\int \frac{2x\,dx}{(x^2 - 2x + 2)}$$

we note that, since the derivative of the denominator is $2x - 2$, we should rewrite our integral as

$$\int \frac{2x\,dx}{(x^2 - 2x + 2)} = \int \frac{2x - 2\,dx}{(x^2 - 2x + 2)} + \int \frac{2\,dx}{(x^2 - 2x + 2)}.$$

- The first of these integrals can be found by using the substitution

$$u = x^2 - 2x + 2$$
$$du = (2x - 2)\,dx$$

so

$$\int \frac{2x - 2\,dx}{(x^2 - 2x + 2)} = \int \frac{du}{u}$$
$$= \ln|u| + C$$
$$= \ln(x^2 - 2x + 2) + C.$$

- The second integral should be seen as

$$\int \frac{2\,dx}{(x^2 - 2x + 2)} = \int \frac{2\,dx}{(x - 1)^2 + 1};$$

the substitution

$$u = x - 1$$
$$du = dx$$

transforms this into

$$\int \frac{2\,du}{u^2+1} = 2\arctan u + C$$
$$= 2\arctan(x-1) + C$$

Thus, the integral of the first fraction is

$$\int \frac{2x\,dx}{(x^2-2x+2)} = \ln(x^2-2x+2) + 2\arctan(x-1) + C.$$

ii. To integrate the second fraction

$$\int \frac{2x\,dx}{(x^2-2x+2)^2}$$

we again note that $2x-2$ is the derivative of x^2-2x+2, so analogously we rewrite our integral as

$$\int \frac{2x\,dx}{(x^2-2x+2)^2} = \int \frac{2x-2\,dx}{(x^2-2x+2)^2} + \int \frac{2\,dx}{(x^2-2x+2)^2}.$$

- Again, the first of these integrals can be found by using the substitution

$$u = x^2-2x+2$$
$$du = (2x-2)\,dx$$

but this time it yields

$$\int \frac{2x-2\,dx}{(x^2-2x+2)^2} = \int \frac{du}{u^2}$$
$$= -\frac{1}{u} + C$$
$$= -\frac{1}{x^2-2x+2} + C.$$

- Similarly, the second integral can be handled using the substitution

$$u = x-1$$
$$du = dx$$

which this time yields

$$\int \frac{2\,dx}{(x^2-2x+2)^2} = \int \frac{2\,du}{(u^2+1)^2}.$$

This is not a familiar integral, but the further substitution

$$\begin{aligned} u &= \tan\theta \\ (\theta &= \arctan u) \\ du &= \sec^2\theta\,d\theta \\ u^2+1 &= \sec^2\theta \end{aligned}$$

leads to

$$\begin{aligned} \int \frac{2\,du}{(u^2+1)^2} &= \int \frac{2\sec^2\theta\,d\theta}{\sec^4\theta} \\ &= \int \frac{2\,d\theta}{\sec^2\theta} \\ &= \int 2\cos^2\theta\,d\theta \\ &= \int (1+\cos 2\theta)\,d\theta \\ &= \theta + \frac{1}{2}\sin 2\theta + C \\ &= \arctan u + \sin\theta\cos\theta + C \\ &= \arctan(x-1) + \frac{x-1}{x^2-2x+2} + C. \end{aligned}$$

Thus the integral of the second term in the partial fraction decomposition is

$$\begin{aligned} \int \frac{2x\,dx}{(x^2-2x+2)^2} \\ &= \left(-\frac{1}{x^2-2x+2} + \arctan(x-1)\right) \\ &\quad + \left(\frac{x-1}{x^2-2x+2}\right) + C \\ &= \arctan(x-1) + \frac{x-2}{x^2-2x+2} + C \end{aligned}$$

Pulling this all together, the full integral is

$$\begin{aligned}\int \frac{(2x^3 - 4x^2 + 6x)\,dx}{(x^2 - 2x + 2)^2} &\\ &= \left(\ln(x^2 - 2x + 2) + 2\arctan(x-1)\right) \\ &\quad + \left(\arctan(x-1) + \frac{x-2}{x^2-2x+2}\right) + C \\ &= \ln(x^2 - 2x + 2) + 3\arctan(x-1) + \frac{x-2}{x^2-2x+2} + C.\end{aligned}$$

According to [30, p. 119], this technique was discussed around 1700 in correspondence between Gottfried Wilhelm Leibniz (1646-1714) and Johann Bernoulli (1667-1748), and was exploited systematically for the integration of rational functions by Leibniz (1702), Johann Bernoulli (1702), Leonard Euler (1701-1783) (1768), and Charles Hermite (1822-1901) (1873).
Even though our motivation for studying the partial fraction decomposition was the integration of rational functions, this is a purely algebraic technique, with uses in other contexts, like the calculation of Laplace transforms in the study of differential equations.

Exercises for § 5.6

Answers to Exercises 1acegi, 2acegik, 3ace, 9f(i,iii,v) are given in Appendix B.

Practice problems:

1. Find the partial fraction decomposition of each function below:

(a) $\dfrac{x-3}{(x-1)(x-2)}$ (b) $\dfrac{2x}{x^2-1}$

(c) $\dfrac{2}{x^2-1}$ (d) $\dfrac{x^2-2}{(x-1)^2(x-2)}$

(e) $\dfrac{x+3}{(x^2+1)(x-1)}$ (f) $\dfrac{x^3-x^2-2x}{(x^2+1)(x-1)^2}$

(g) $\dfrac{2x}{(x^2+1)(x+1)}$ (h) $\dfrac{x^2-5}{(x^2+1)(x^2-1)}$

(i) $\dfrac{2x^3 - 3x^2 - 2x - 5}{(x^2+1)^2(x^2-1)}$

2. Evaluate each of the following integrals:

(a) $\displaystyle\int \frac{2dx}{x^2-1}$ (b) $\displaystyle\int \frac{(x-5)\,dx}{(x+1)(x-2)}$

(c) $\displaystyle\int \frac{(5x+7)\,dx}{x^2+2x-3}$ (d) $\displaystyle\int \frac{dx}{x^2+4x}$

(e) $\displaystyle\int \frac{(2x^2-5x-2)\,dx}{x-3}$ (f) $\displaystyle\int \frac{(2x+6)\,dx}{x^2+6x+5}$

(g) $\displaystyle\int \frac{(5x+7)\,dx}{(x+1)(x+2)^2}$ (h) $\displaystyle\int \frac{(3x^2+6x+4)\,dx}{(x+1)(x^2+2x+2)}$

(i) $\displaystyle\int \frac{x\,dx}{(x+1)(x^2+3x+2)}$ (j) $\displaystyle\int \frac{(x^2+2x+2)\,dx}{(x+1)(x^2+1)}$

(k) $\displaystyle\int \frac{(4x^2+9x+13)\,dx}{(x+1)(x^2+2x+5)}$ (l) $\displaystyle\int \frac{(x^2+2x-1)\,dx}{(x+1)^2(x^2+4x+5)}$

3. Find the partial fraction decomposition given by Theorem 5.6.1 for each fraction below:

(a) $\dfrac{8}{15}$ (b) $\dfrac{3}{10}$ (c) $\dfrac{23}{27}$

(d) $\dfrac{5}{54}$ (e) $\dfrac{9}{35}$ (f) $\dfrac{123}{140}$

Theory problems:

4. **Another way to evaluate $\int \sec x\,dx$:**

(a) Write

$$\begin{aligned}\int \sec x\,dx &= \int \frac{dx}{\cos x}\\ &= \int \frac{\cos x\,dx}{\cos^2 x}\\ &= \int \frac{\cos x\,dx}{1-\sin^2 x}.\end{aligned}$$

(b) Substitute $s = \sin x$ and apply partial fractions.

(c) Interpret your answer in terms of x.

5. **Another integration of $\sec^3 x$:** Rewrite the integral
$$\int \sec^3 x\,dx$$
in terms of $\cos x$, multiply the numerator and denominator of the integrand by $\cos x$, and use the substitution
$$u = \sin x$$
to evaluate the integral by partial fractions.

6. Once we have set up the form of the partial fraction decomposition, an alternative to combining the fractions over a common denominator is to pick some values for x and substitute them into the expression, resulting in an equation (for each choice of value) for the unknown coefficients.

   (a) How many choices of value do we need? Are there any values which should be avoided?

   (b) Use this method to find the partial fraction decompositions of the functions given in Exercise 1.

7. The *Euclidean algorithm* (Exercise 11) can be interpreted as saying that if two integers b_1 and b_2 are relatively prime, then 1 is a linear combination of them: that is, there exist integers A_1 and A_2 such that
$$A_1 b_1 + A_2 b_2 = 1.$$
Use this to prove Lemma 5.6.2, as follows:

   (a) Show that, given A/B proper and in lowest terms, where $B = b_1 b_2$ (with b_i relatively prime), we can find a_i, $i = 1, 2$, so that
$$\frac{A}{B} = \frac{a_1}{b_1} + \frac{a_2}{b_2}.$$

   (b) Show that if for $i = 1$ or $i = 2$ a_i and b_i are not relatively prime, then A/B is not in lowest terms.

   (c) Show that the coefficients a_1 and a_2 are not unique. (*Hint:* Given one pair, add b_1 to the first numerator and subtract b_2 from the second to get another pair.)

   (d) Show that, among all possible choices for a_1 and a_2, if we pick the one for which a_1 has the least positive value, then $a_1 < b_1$.

(e) Assuming A/B positive, show that if $0 < a_1 < b_1$ then $|a_2| < b_2$.

8. Suppose p is a prime.

(a) **Show** that every integer $N > 0$ has a unique **base p expansion**

$$N = a_0 + a_1 p + a_2 p^2 + \cdots + a_k p^k$$

where $0 \leq a_j < p$ for $j = 1, \ldots, k$.
(*Hint:* Pick k so that $p^k \leq N < p^{k+1}$, then find $a_k = \max\{a \,|\, ap^k \leq N\}$.)

(b) Use this to prove Lemma 5.6.3.

Challenge problems:

9. **The Heaviside "Cover Up" Method:** Another alternative to solving equations for the coefficients in a partial fraction decomposition is the following method, attributed to Oliver Heaviside (1850-1925), who along with Maxwell developed electromagnetic theory and contributed to the development of vector calculus and other operational methods in applied mathematics.

Suppose

$$f(x) = \frac{A(x)}{B(x)}$$

is proper and in lowest terms.

(a) **Show** that if

$$B(x) = (x - r_1)(x - r_2) \cdots (x - r_k)$$

(that is, the denominator has distinct real roots $x = r_i$, $i = 1, \ldots, k$, and no irreducible quadratic divisors), then the coefficient a_i in the term $a_i/(x - r_i)$ for the partial fraction decomposition of $f(x)$ can be found by "covering up" the factor $(x - r_i)$ in the expression

$$f(x) = \frac{A(x)}{(x - r_1)(x - r_2) \cdots (x - r_k)}$$

getting

$$(x - r_i)f(x) = \frac{A(x)}{\hat{B}(x)}$$

where

$$B(x) = (x - r_i)\hat{B}(x)$$

—that is,

$$\hat{B}(x) = (x - r_1)(x - r_2)\cdots(x - r_{i-1})\cdot \langle (x - r_i) \textit{ missing } \rangle (x - r_{i+1})(x - r_{i+2})\cdots(x - r_k)$$

—and substituting $x = r_i$ in what remains. In other words, show that

$$a_i = \frac{A(r_i)}{\hat{B}(r_i)}.$$

Hint: First show that $f(x)$ can be written as a sum of proper fractions in the form

$$f(x) = \frac{a_i}{x - r_i} + \frac{\hat{A}(x)}{\hat{B}(x)}.$$

(b) **Show** that, even if there are other multiple or complex roots for the denominator $B(x)$, this procedure still finds the coefficient in $a_i/(x - r_i)$ whenever $x = r_i$ is a *simple* (multiplicity 1) root of $B(x)$. (*Hint:* Same as above.)

(c) **Show** that if $x = r$ is a real root of $B(x)$ with multiplicity $m > 1$, then "covering up" the factor $(x - r)^m$ and evaluating at $x = r$ gives the coefficient in the term $a_m/(x - r)^m$.

(d) **Show** $x = r$ is a root of multiplicity m for $B(x)$ precisely if all derivatives of $B(x)$ to order $m - 1$ are zero at $x = r$ and the m^{th} derivative is nonzero at $x = r$.

(e) **Show** that when $x = r$ is a real root of multiplicity m for $B(x)$, then the coefficient in $a_j/(x - r)^j$ is found by "covering up" the $(x - r)^m$ factor, *differentiating* $m - j$ *times*, then evaluating at $x = r$ and dividing by $(m - j)!$. (*Hint:* Write out the partial fraction decomposition, multiply by $(x - r)^m$, differentiate, and examine the terms for their values at $x = r$.)

(f) Use the "cover up" method to find the partial fraction decomposition of each function below:

i. $\dfrac{x-3}{(x-1)(x-2)}$ ii

$\dfrac{2x}{x^2-1}$

iii. $\dfrac{x^2+1}{(x-1)(x-2)(x-3)}$ iv

$\dfrac{x^2+2}{(x-1)^2(x-2)}$

v. $\dfrac{x^3+1}{x^3(x-1)}$ vi

$\dfrac{x^4+2}{x(x-1)^4}$

10. *(This problem uses some elementary complex analysis.)* If $z = \alpha + \beta i$ is a complex number, its **conjugate** is the number $\bar{z} = \alpha - \beta i$. Some basic properties of conjugates are (see § 6.6):

- $z + \bar{z}$ is real;
- $z\bar{z} = |z|^2 = \alpha^2 + \beta^2$;
- $\overline{z^k} = \bar{z}^k$. In particular, for any polynomial $P(x)$ with real coefficients, $P(\bar{z}) = \overline{P(z)}$.
- As a corollary, if $x = z$ is a root of $P(x)$ then so is $x = \bar{z}$ (with the same multiplicity).

Using this, prove Lemma 5.6.7 as follows.

(a) Note that for any complex number $\zeta = \alpha + \beta i$,

$$(x-\zeta)(x-\bar{\zeta}) = (x-\alpha)^2 + \beta^2.$$

(b) By mimicking the proof of Lemma 5.6.6, show that if $A(x)/B(x)$ (with real coefficients) is proper and in lowest terms, and

$$B(x) = \{(x-\alpha)^2 + \beta^2\}^m \tilde{B}(x)$$

where $\tilde{B}(\zeta) \neq 0$, then

$$\frac{A(x)}{B(x)} = \frac{a_1}{(x-\zeta)^m} + \frac{a_2}{(x-\bar{\zeta})^m} + \frac{\hat{A}(x)}{\hat{B}(x)}$$

where a_1 and a_2 are the (possibly complex) numbers

$$a_1 = \frac{A(\zeta)}{(\zeta-\bar{\zeta})^m \tilde{B}(\zeta)}$$
$$a_2 = \frac{A(\bar{\zeta})}{(\bar{\zeta}-\zeta)^m \tilde{B}(\bar{\zeta})};$$

and $x = \zeta$ has multiplicity $\hat{m} < m$ as a root of $\hat{B}(x)$.

(c) Show that
$$a_2 = \bar{a_1};$$

(d) Show that for any (complex) number a,
$$\frac{a}{(x-\zeta)^m} + \frac{\bar{a}}{(x-\bar{\zeta})^m} = \frac{a'x+b}{\{(x-\alpha)^2+\beta^2\}^m}$$
where a' and b are real.

History note:

11. **Euclidean Algorithm:** Proposition 1 in Book VII of the *Elements*[32, p. 296] reads:

> *Two unequal numbers being set out, and the less being continually subtracted in turn from the greater, if the number which is left never measures the one before it until an unit is left, the original numbers will be prime to one another.*

The modern interpretation of this and the proposition which follows (dealing with the case when both are divisible by another integer) is: For any two positive integers b_1 and b_2 with greatest common divisor c, there exist integers a_1 and a_2 such that
$$a_1 b_1 + a_2 b_2 = c.$$
Furthermore, if $c = 1$ these numbers can be picked so that
$$|a_1| < b_2$$
and
$$|a_2| < b_1.$$
Prove this as follows:

(a) Suppose

$$0 < b_2 < b_1;$$

let q_2 be the maximum integer for which

$$q_2 b_2 \leq b_1,$$

and set

$$b_3 = b_1 - q_2 b_2.$$

Show that

$$1 \leq q_2 \leq b_1$$

and

$$0 \leq b_3 < b_2,$$

with $q_2 < b_1$ unless $b_2 = 1$.
If $b_3 = 0$, set $c = b_2$ and **show** that c is the greatest common divisor of b_1 and b_2.

(b) If $b_3 > 0$, we can repeat the procedure, with indices shifted by one: let q_3 be the maximum integer with

$$q_3 b_3 \leq b_2,$$

and set

$$b_4 = b_2 - q_3 b_3.$$

Then, as before,

$$1 \leq q_3 \leq b_2$$

and

$$0 \leq b_4 < b_3.$$

We can continue iterating this procedure: if $0 < b_{j+1} < b_j$ then we can define q_{j+1} to be the maximum integer with

$$q_{j+1}b_{j+1} \leq b_j,$$

and set

$$b_{j+2} = b_j - q_{j+1}b_{j+1}.$$

Again,

$$1 \leq q_{j+1} \leq b_j$$

and

$$0 \leq b_{j+2} < b_{j+1}.$$

Show that after finitely many steps, we arrive at $b_{j+2} = 0$.

(c) When this happens, set $c = b_{j+1}$. **Show** that c divides each of $b_1, b_2, ..., b_j$. (*Hint:* First, show that c divides b_j and b_{j+1}; then show from the definition of b_{j+1} that a number which divides b_j and b_{j+1} also divides b_{j-1}, and proceed inductively.)

(d) **Show** that if an integer divides b_1 and b_2, then it divides each of $b_1, b_2, \ldots, b_{j+1}$. Use this to show that c as defined above is the *greatest* common divisor of b_1 and b_2. In particular, if $c = 1$, then b_1 and b_2 are relatively prime.

(e) **Show** that there exist integers a_1 and a_2 such that

$$a_1b_1 + a_2b_2 = c.$$

(*Hint:* Use $c = b_{j+1}$ and the relations $b_{i+1} = b_{i-1} - q_ib_i$ for $i = j, \ldots, 1$ to express c successively as a combination of the form $a_ib_i + a_{i-1}b_{i-1}$ for $i = j, \ldots, 2$.)

(f) **Show** that any pair of numbers a_1, a_2 satisfying this condition cannot have the same sign.

(g) **Show** that a_1, a_2 and a_1', a_2' are two pairs of numbers satisfying the condition above with $c = 1$ if, and only if, for some integer k

$$a_1' - a_1 = kb_2$$

and

$$a_2' - a_2 = -kb_1.$$

(h) Use this to show that when $c = 1$ (*i.e.*, b_1 and b_2 are relatively prime), a_1 and a_2 can be picked so that

$$|a_1| < b_2$$

and

$$|a_2| < b_1.$$

(*Hint:* Suppose $a_1 > 0 > a_2$ and $a_1 > b_2$. Then $a_1' = a_1 - b_2 > 0$, and so $a_2' = a_2 + b_1 < 0$, implying that $|a_i'| < |a_i|$ for $i = 1, 2$. In particular, if we pick the lowest positive value for a_1 then $0 < a_1 < b_2$, and then by symmetric argument, $0 > a_2 > -b_1$.)

5.7 Improper Integrals

The definite integral

$$\int_a^b f(x)\,dx$$

as formulated in § 5.1 can only make sense when both the *integrand* f and the *domain of integration* $[a, b]$ are *bounded*. It would at times be convenient to relax either of these boundedness restrictions. This leads to two kinds of **improper integrals**: those over an unbounded domain, and those with an unbounded integrand.

Unbounded Domains

The first kind of integral is illustrated by the area of the region under the curve $y = x^{-2}$ to the right of the line $x = 1$ (Figure 5.21). For any (finite) $b > 1$, the area between $x = 1$ and $x = b$ is given by the integral

$$\int_1^b x^{-2}\,dx = -x^{-1}\Big|_1^b = 1 - \frac{1}{b}.$$

As we move the cutoff point $x = b$ further right, the area up to $x = b$ converges to

$$\lim_{b\to\infty}\left(1 - \frac{1}{b}\right) = 1.$$

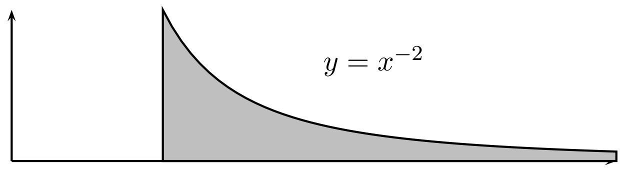


Figure 5.21: Area under $y = x^{-2}$

We would naturally say that the total area under $y = x^{-2}$ to the right of $x = 1$—in other words, the integral of $f(x) = x^{-2}$ over the unbounded interval $[1, \infty)$—is

$$\int_1^\infty x^{-2} dx = \lim_{b\to\infty} \int_1^b x^{-2} dx = \lim_{b\to\infty} \left(1 - \frac{1}{b}\right) = 1.$$

This naturally extends to integrals of functions which are allowed to change sign. We already saw in Lemma 5.3.2 that the function

$$\Phi(t) = \int_a^t f(x)\ dx$$

is continuous at every finite value of t, in other words if f is integrable on $[a, b]$ for $a < b < \infty$, then

$$\Phi(b) = \lim_{t\to b^-} \Phi(t)\,.$$

We extend this to $b = \infty$, as follows:

Definition 5.7.1. *Fix $a \in \mathbb{R}$ and suppose f is integrable on $[a, b]$ whenever $a < b < \infty$. The **integral over** $[a, \infty)$ of f is defined as the limit*

$$\int_a^\infty f(x)\ dx = \lim_{b\to\infty} \int_a^b f(x)\ dx$$

*if this limit exists. We than say the integral **converges**; otherwise it **diverges**.*

We saw that the integral of x^{-2} over $[1, \infty)$ converges. By contrast, the integral

$$\int_1^\infty x^{-1} dx = \lim_{b\to\infty} \int_1^b x^{-1} dx = \lim_{b\to\infty} (\ln x\Big|_1^b)$$
$$= \lim_{b\to\infty} \ln b = \infty$$

diverges.

Integrals over half-finite intervals are closely related to infinite series. If the integral $\int_a^\infty f(x)\ dx$ converges, then for every strictly increasing unbounded sequence

$$a = a_0 < a_1 < a_2 < \dots \quad (a_n \uparrow \infty)$$

we can write

$$\int_a^\infty f(x)\,dx = \lim_{n\to\infty} \int_a^{a_n} f(x)\ dx = \lim_{n\to\infty} \sum_{j=1}^{n} \int_{a_{j-1}}^{a_j} f(x)\ dx$$
$$= \sum_{j=1}^{n} c_j \quad \text{where } c_j = \int_{a_{j-1}}^{a_j} f(x)\ dx.$$

In general, the converse does not hold: for example

$$f(x) = \sin x, \quad a_j = 2\pi j, j = 0, \dots$$

yields

$$c_j = \int_{2\pi(j-1)}^{2\pi j} \sin x\ dx = -\cos x|_{2\pi(j-1)}^{2\pi j} = 0$$

so the series converges even though the integral

$$\int_0^\infty \sin x\ dx = \lim_{b\to\infty} (-\cos x)|_0^b = \lim_{b\to\infty} (1 - \cos b)$$

diverges.

The situation is a bit better if the integrand is always non-negative. Fix $a \in \mathbb{R}$. If $f(x) \geq 0$ for all $x \geq a$ (and f is integrable on $[a, b]$ for $a < b < \infty$), then the function

$$\Phi(t) = \int_a^t f(x)\ dx$$

is non-decreasing (Exercise 4). In particular, as $t \to \infty$, either $\Phi(t) \to \infty$ or $\Phi(t)$ is bounded, and converges. So *for non-negative integrands*, we can denote convergence of the integral by

$$\int_a^\infty f(x)\ dx < \infty,$$

just as we did for positive series. Furthermore, in this case we can test convergence using any strictly increasing unbounded sequence $a_n \uparrow \infty$.

Remark 5.7.2. *Fix $a \in \mathbb{R}$, and assume f is non-negative and integrable on $[a, b]$ for all b with $a < b < \infty$. Let*

$$a = a_0 < a_1 < a_2 < \ldots \quad a_n \uparrow \infty$$

be any unbounded strictly increasing sequence, and form the series

$$\sum_{j=1}^{\infty} c_j$$

where

$$c_j = \int_{a_{j-1}}^{a_j} f(x)\ dx.$$

Then the improper integral

$$\int_a^{\infty} f(x)\ dx$$

and the series

$$\sum_{j=1}^{\infty} c_j$$

either both diverge, or both converge to the same limit.

This follows from the fact that, if $\Phi(t)$ is non-decreasing on (a, ∞) and $t_n \uparrow \infty$, then

$$\lim \Phi(t_n) = \lim_{t \to \infty} \Phi(t)$$

(Exercise 5).
Using this observation we easily obtain the analogue of the comparison test for positive series.

Proposition 5.7.3 (Comparison Test for Integrals)**.** *Fix $a \in \mathbb{R}$. Suppose f and g are integrable on $[a, b]$ whenever $a < b < \infty$, and*

$$0 \le f(x) \le g(x) \quad \textit{for all } x \ge a.$$

Then

$$\int_a^{\infty} f(x)\ dx \le \int_a^{\infty} g(x)\ dx.$$

That is,

- *if $\int_a^{\infty} g(x)\ dx$ converges, so does $\int_a^{\infty} f(x)\ dx$.*

- *If* $\int_a^\infty f(x)\ dx$ *diverges, so does* $\int_a^\infty g(x)\ dx$.

The proof is, again, an easy application of Remark 5.7.2 and the montonicity of integrals (Proposition 5.2.7); see Exercise 6.
A useful and important connection between improper integrals and series is the following.

Proposition 5.7.4 (Integral Test for Series)**.** *Suppose f is non-negative and decreasing on $[1, \infty)$.*
Then the integral

$$\int_1^\infty f(x)\ dx$$

and the series

$$\sum_{n=1}^\infty f(n)$$

either both converge or both diverge.

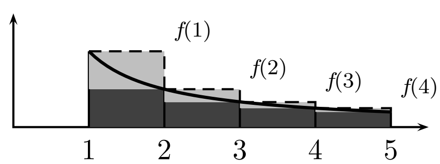


Figure 5.22: Proposition 5.7.4 (Integral Test for Series)

Proof. Let a_n be the sequence $a_n = n$, so in terms of Remark 5.7.2,

$$c_j = \int_{a_{j-1}}^{a_j} f(x)\ dx = \int_{j-1}^{j} f(x)\ dx \quad j = 2, \dots .$$

Since f is decreasing, we have

$$f(j-1) \geq f(x) \geq f(j) \qquad \text{for } j-1 \leq x \leq j$$

and hence (since the length of the interval $[j-1, j]$ is 1)

$$f(j-1) \geq c_j \geq f(j)\,,$$

which gives us

$$\sum_{j=2}^\infty f(j-1) \geq \sum_{j=2}^\infty c_j \geq \sum_{j=2}^\infty f(j)\,.$$

Now, the two outside series

$$\begin{aligned}\sum_{j=2}^{\infty} f(j-1) &= f(1) + f(2) + f(3) + \dots \\ \sum_{j=2}^{\infty} f(j) &= \qquad\quad f(2) + f(3) + \dots\end{aligned}$$

differ only in the additional term for the first series. Hence both converge, or both diverge. The inequality above then forces $\sum c_j$ to follow suit (*i.e.*, converge or diverge) and, by Remark 5.7.2, this determines convergence or divergence of the integral. □

An important consequence of the integral test is the following, which among other things puts the behavior of the harmonic series (p. 61) into a wider context, and in particular gives an alternative proof of Exercise 34 in § 2.3.

Corollary 5.7.5 (p-series Test)**.** *The series*

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges if $p > 1$ *and diverges if* $p \leq 1$.

Proof. For $p > 0$, the function

$$f(x) = x^{-p}$$

satisfies the hypotheses of the integral test, and our series is precisely

$$\sum_{n=1}^{\infty} f(n).$$

But for $p \neq 1$,

$$\begin{aligned}\int_1^{\infty} f(x)\ dx &= \lim_{b\to\infty} \int_1^b x^{-p} dx = \lim_{b\to\infty} \left.\frac{x^{1-p}}{1-p}\right|_1^b \\ &= \lim_{b\to\infty} \left(\frac{b^{1-p}}{1-p} - \frac{1}{1-p}\right)\end{aligned}$$

which converges precisely if $1 - p < 0$, or

$$1 < p.$$

The special case $p = 1$ also leads to divergence:

$$\int_1^\infty x^{-1} dx = \lim_{b\to\infty} \ln x|_1^b = \lim_{b\to\infty} \ln b = \infty.$$

□

We mention without proof (see Exercise 7) the following integral analogue of the absolute convergence test for series:

Lemma 5.7.6 (Absolute Convergence for Integrals)**.** *Fix $a \in \mathbb{R}$ and suppose f is integrable on $[a, b]$ whenever $a < b < \infty$.*
If

$$\int_a^\infty |f(x)|\ dx < \infty$$

(note that this integral has non-negative integrand), then also

$$\int_a^\infty f(x)\ dx$$

converges, and satisfies

$$\left|\int_a^\infty f(x)\ dx\right| \leq \int_a^\infty |f(x)|\ dx.$$

(See Figure 5.23.)

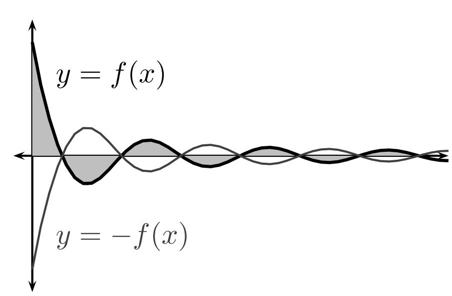


Figure 5.23: Lemma 5.7.6 (Absolute Convergence for Integrals)

Integration of f over an interval of the form $(-\infty, b]$ follows the analogous pattern: fixing b, if f is integrable on $[a, b]$ whenever $-\infty < a < b$, then we define

$$\int_{-\infty}^{b} f(x)\ dx = \lim_{a \to -\infty} \int_{a}^{b} f(x)\ dx;$$

the integral **converges** or **diverges** according to whether or not the limit exists.

For integration over the whole real line, we need to be a little more careful.

Definition 5.7.7. *Suppose f is integrable on every interval $[a, b]$, $-\infty < a < b < \infty$. We say that the integral*

$$\int_{-\infty}^{\infty} f(x)\ dx$$

converges *if for some finite $c \in \mathbb{R}$, the half-infinite integrals*

$$\int_{-\infty}^{c} f(x)\ dx, \quad \int_{c}^{\infty} f(x)\ dx$$

both *converge, and in this case we define*

$$\int_{-\infty}^{\infty} f(x)\ dx = \int_{-\infty}^{c} f(x)\ dx + \int_{c}^{\infty} f(x)\ dx.$$

If either (or both) of the half-infinite integrals diverges, then the full integral also diverges.

We note that while this definition seems to depend on a specific choice of the "cut point" $c \in \mathbb{R}$, it does not: a different choice $c' \in \mathbb{R}$ will give the same conclusion about convergence, and in case of convergence, the formula using c' will give the same value for $\int_{-\infty}^{\infty} f(x)\ dx$ (Exercise 8).

We caution the reader that this definition is *different* from taking a single limit of the form

$$\lim_{a \to \infty} \int_{-a}^{a} f(x)\ dx.$$

For example, the limit above is 0 for every odd function, like $f(x) = \sin x$ or $f(x) = x$, for both of which $\int_{-\infty}^{0} f(x)\ dx$ and $\int_{0}^{\infty} f(x)\ dx$ diverge.

Unbounded Functions

The second type of improper integral involves a function which blows up at a single point (or at a finite number of points). For example, we might ask about the area of the (unbounded) region bordered by the axes, the line $x = 1$, and the graph of $y = x^{-p}$ for some $p > 0$: this leads to the "integral"

$$\int_0^1 x^{-p}\, dx$$

which looks OK, until we note that the integrand is not bounded near $x = 0$. A more subtle example is

$$\int_0^2 \frac{dx}{(x-1)^2}$$

which blows up at $x = 1$.
We will, for simplicity of exposition, assume that the integrand is continuous where it doesn't blow up. First, we deal with blowups at the endpoints.

Definition 5.7.8. *Suppose f is continuous on $[a, b)$, but blows up at $x = b$*

$$\lim_{x\to b^-} |f(x)| = \infty.$$

We define the (improper) integral of f over $[a, b]$ as the one-sided limit

$$\int_a^b f(x)\, dx := \lim_{t\to b^-} \int_a^t f(x)\, dx;$$

the integral ***converges*** *if the limit exists and* ***diverges*** *otherwise. Similarly, if f is continuous on $[a, b)$ and blows up at $x = a$, then*

$$\int_a^b f(x)\, dx := \lim_{t\to a^+} \int_t^b f(x)\, dx.$$

Recall that $f(x) := x^{-p}$ has a convergent integral on $(1, \infty]$ if $p > 1$. Now (for $p > 0$), $f(x)$ blows up at $x = 0$, so we might investigate the integral on $[0, 1]$. A simple calculation for $p \neq 1$ gives

$$\int_0^1 x^{-p} dx = \lim_{t\to 0^+} \int_t^1 x^{-p} dx = \lim_{t\to 0^+} \left(\frac{1}{1-p} - \frac{t^{1-p}}{1-p} \right),$$

which converges precisely if $1 - p > 0$, or

$$p < 1.$$

So the situation is opposite to that at ∞. For example, $p = \frac{1}{2}$ gives

$$\int_0^1 \frac{dx}{\sqrt{x}} = \lim_{t \to 0^+} 2\sqrt{x}\Big|_t^1 = \lim_{t \to 0^+} (2 - 2\sqrt{t}) = 2$$

which converges, but $p = 2$ gives

$$\int_0^1 \frac{1}{x^2} dx = \lim_{t \to 0^+} -\frac{1}{x}\Big|_t^1 = \lim_{t \to 0^+} \left(\frac{1}{t} - 1\right)$$

which diverges.
If the blowup occurs at an interior point of an interval, we say the integral "across" the blowup conveges precisely if the integrals on either side *both* converge in the earlier sense:

Definition 5.7.9. *Suppose f is continuous at every point of $[a,b]$ except $x = c$, where $a < c < b$ and $f(x)$ blows up at $x = c$.*
Then the integral $\int_a^b f(x)\,dx$ converges precisely if BOTH *of the integrals*

$$\int_a^c f(x)\,dx = \lim_{t \to c^-} \int_a^t f(x)\,dx$$
$$\int_c^b f(x)\,dx = \lim_{t \to c^+} \int_t^b f(x)\,dx$$

converge, and then we define

$$\int_a^b f(x)\,dx := \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

For example, to evaluate

$$\int_0^2 \frac{dx}{(x-1)^2}$$

note that the integrand blows up at $x = 1$, so we need to consider the two one-sided integrals

$$\int_0^1 \frac{dx}{(x-1)^2} := \lim_{t \to 1^-} \int_0^t \frac{dx}{(x-1)^2} = \lim_{t \to 1^-} \left[-\frac{1}{x-1}\right]_0^t = \lim_{t \to 1^-} \left(1 - \frac{1}{t-1}\right)$$
$$\int_1^2 \frac{dx}{(x-1)^2} := \lim_{t \to 1^+} \int t^2 \frac{dx}{(x-1)^2} = \lim_{t \to 1^+} \left[-\frac{1}{x-1}\right]_0^t = \lim_{t \to 1^+} \left(\frac{1}{t-1} - 1\right).$$

Both limits diverge, so the integral diverges.

If we had ignored the blowup at $x = 1$ and just calculated

$$\left[-\frac{1}{x-1}\right]_0^2 = -\frac{1}{2-1} - \left(-\frac{1}{0-1}\right) = -1 - 1 = -2$$

we would have obtained the absurd result that the graph of a function which is positive wherever it is defined on $[-1, 1]$ (and goes to $+\infty$ when it is not!) has a negative integral.

Exercises for § 5.7

Answers to Exercises 1acegikmo, 2ace, 3acegi are given in Appendix B.

Practice problems:

1. Each integral below is improper. For each, (i) find any points at which the function blows up; (ii) determine whether the integral converges or diverges; and (iii) if it converges, evaluate it.

   (a) $\int_0^\infty \cos^2 3x\,dx$ (b) $\int_{-\infty}^0 \frac{dx}{(x-1)^4}$ (c) $\int_0^2 \frac{dx}{(x-1)^4}$

   (d) $\int_2^\infty \frac{dx}{(x-1)^4}$ (e) $\int_{-\infty}^\infty \frac{x\,dx}{x^2+1}$ (f) $\int_{-\infty}^\infty \frac{dx}{x^2+1}$

   (g) $\int_{-\infty}^\infty \frac{dx}{\sqrt{x^2+1}}$ (h) $\int_{-\infty}^\infty \frac{x\,dx}{\sqrt{x^2+1}}$ (i) $\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}}$

   (j) $\int_1^\infty \frac{dx}{\sqrt{x^2-1}}$ (k) $\int_0^1 \frac{dx}{x^2-1}$ (l) $\int_1^\infty \frac{dx}{x^2-1}$

   (m) $\int_{-1}^1 \frac{dx}{x^2}$ (n) $\int_{-1}^1 \frac{dx}{\sqrt[3]{x}}$ (o) $\int_{-\infty}^\infty xe^{-x^2}\,dx$

2. For each improper integral below, decide whether it converges or diverges, without trying to evaluate it.

   (a) $\int_1^\infty \frac{1}{x^2}\sin x\,dx$ (b) $\int_0^1 x\sin\frac{1}{x}\,dx$

   (c) $\int_1^\infty x^{-3/2}\sin^2 x\,dx$ (d) $\int_2^\infty \frac{dx}{(x^2-1)^{3/2}}$

   (e) $\int_2^\infty \frac{x\,dx}{(x^2-1)^{3/2}}$ (f) $\int_0^2 \frac{dx}{(x^2-1)^{4/3}}$

3. Use the Integral Test for Series (Proposition 5.7.4) to decide which series below converge, and which diverge.

(a) $\sum_{n=1}^{\infty} \frac{1}{2n^2-1}$ (b) $\sum_{n=1}^{\infty} \frac{1}{\sqrt{2n^2-1}}$ (c) $\sum_{n=2}^{\infty} \frac{1}{n^2-n}$

(d) $\sum_{n=2}^{\infty} \frac{\ln n}{n}$ (e) $\sum_{n=2}^{\infty} \frac{\ln n}{n^2}$ (f) $\sum_{n=3}^{\infty} \frac{n^2}{e^n}$

(g) $\sum_{n=2}^{\infty} \frac{1}{n \ln n}$ (h) $\sum_{n=1}^{\infty} \frac{1}{(n+1)(n+2)}$ (i) $\sum_{n=1}^{\infty} \frac{1}{n^2-2n+5}$

Theory problems:

4. Justify the statement on p. 430 that if $f(x) \geq 0$ for all $x \geq a$ and f is integrable over $[a, b]$ for all $a < b < \infty$, then

$$\Phi(t) := \int_a^t f(x)\,dx$$

is a non-decreasing function.

5. Justify the statement on p. 431 that if $\Phi(t)$ is non-decreasing on (a, ∞) and $t_n \uparrow \infty$, then

$$\lim \Phi(t_n) = \lim_{t\to\infty} \Phi(t)\,.$$

6. Use Remark 5.7.2 and Proposition 5.2.7 to prove the Comparison Test for Integrals (Proposition 5.7.3).

7. Use the inequality

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\;dx$$

(Exercise 7 in § 5.2) and Exercise 9 in § 2.5 (on Cauchy sequences) to prove Lemma 5.7.6.

8. Explain and justify the statement on p. 435 that if f is integrable on every interval $[a, b]$, and if the half-infinite integrals $\int_{-\infty}^{c} f(x)\;dx$ and $\int_{c}^{\infty} f(x)\;dx$ both converge, then so do the half-infinite integrals

obtained by replacing c with c', and then the value for $\int_{-\infty}^{\infty} f(x)\ dx$ obtained using c will be the same as that using c'. (*Hint:* Consider three integrals, of which one is finite and the other two half-infinite.)

Challenge problems:

9. Consider the integral $\int_0^1 \frac{1}{x^2} \sin x\, dx$.
   (a) Show that the integrand blows up at $x = 0$. (*Hint:* $\lim_{x\to 0} \frac{\sin x}{x} = 1$.)
   (b) Show that for x small, $\frac{1}{x^2} \sin x > \frac{1}{2x}$ and so the integral diverges.

10. Consider the integral $\int_0^1 \sin \frac{1}{x}\, dx$.
   (a) For each integer $n = 1, 2, ...,$ define $a_n := \frac{1}{n\pi}$. Show that $\sin \frac{1}{x}$ is *positive* on (a_{n+1}, a_n) if n is *even* and *negative* if n is *odd.*
   (b) Let
   $$I_0 := \int_{a_1}^1 \sin \frac{1}{x}\, dx$$
   and for $n \in \{1, 2, \ldots, \}$
   $$I_n := \int_{a_{n+1}}^{a_n} \sin \frac{1}{x}\, dx.$$
   Show that for $n = 1, 2, ...$
   $$|I_n| < a_n - a_{n+1},$$
   and I_n is *positive* for n *even*, and *negative* for n *odd.*
   (c) Show that $a_n - a_{n+1}$ monotonically decreases to zero.
   (d) Show that if $a_{n+1} \le b \le a_n$ $(n \ge 1)$ then $\int_b^1 \sin \frac{1}{x}\, dx$ is between $\sum_{k=0}^{n-1} I_k$ and $\sum_{k=0}^{n} I_k$.

(e) Use the Alternating Series Test to show that the integral converges.

History note:

11. **Fermat's calculation of $\int x^{-k} dx$:** As noted in Exercise 6 in § 5.3, Fermat used an infinite partition whose lengths are in geometric progression to calculate the area under the "hyperbola" $x^k y = 1$ (or $y = x^{-k}$) for positive $k \neq 1$. Note that this time we are looking at the area between the x-axis and the curve $y = x^{-k}$ to the *right* of the vertical line $x = a$.

([33], [51, pp.220-2]) For the hyperbola, Fermat started with an infinite partition of the interval $[a, \infty)$

$$\mathcal{P}_r = \{a, ra, r^2a, ...\}$$

where $r = \frac{m}{n} > 1$, so the sequence is increasing to ∞.

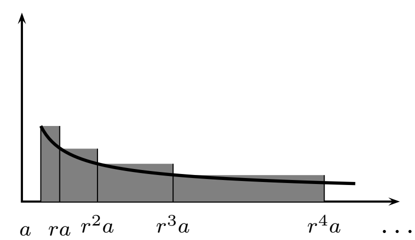


Figure 5.24: Fermat's calculation of $\int_a^\infty x^{-k}\, dx$

(a) Show that the circumscribed rectangle over $I_1 = [a, ra]$ has area

$$R_1 = (r - 1)\frac{1}{a^{k-1}}.$$

(b) Show that the subsequent circumscribed rectangles have area

$$R_j = \left(\frac{n}{m}\right)^{(j-1)(k-1)} R_1 = (r^{-(k-1)})^{j-1} R_1.$$

(c) Thus, the total circumscribed area is given by the geometric series

$$R_1 \left[1 + (r^{-(k-1)}) + (r^{-(k-1)})^2 + (r^{-(k-1)})^3 + ...\right]$$

and thus (since $r^{-1} = \frac{n}{m} < 1$) the total circumscribed area can be expressed as

$$\frac{R_1}{1-\left(\frac{1}{r}\right)^{k-1}} = \frac{(r-1)}{a^{k-1}}\frac{r^{k-1}}{r^{k-1}-1}$$
$$= \frac{1}{\frac{1}{r}+\left(\frac{1}{r}\right)^2+...+\left(\frac{1}{r}\right)^{k-2}}\left(\frac{r}{a}\right)^{k-1}.$$

(d) Then taking $r = \frac{m}{n}$ to 1 (or, in Fermat's words, letting the area of the first rectangle go to nothing), we get that the limit of the upper areas is

$$A = \frac{1}{k-1}\frac{1}{a^{k-1}}.$$

(e) Does any of this depend on k being an integer?

5.8 Geometric Applications of Riemann Sums

In this section we formulate two geometric problems—areas in the plane and volume of revolutes—in terms of the definite integral. These calculations rest on a way of thinking in terms of Riemann sums.

Area in the Plane

We have already interpreted the integral of a function in terms of the area between the x-axis and the graph of a function. We now look at the area between two graphs.
For example, let us find the area of the region bounded *below* by the curve

$$y = x^2 - 3$$

and *above* by the line

$$y = 2x.$$

These curves cross at the points $(-1, -2)$ and $(3, 6)$, so the region whose area we wish to find (see Figure 5.25) could be defined as

$$S := \{(x, y) \mid -1 \leq x \leq 3 \text{ and } x^2 - 3 \leq y \leq 2x\}.$$

In principle, by appropriately adding and subtracting the areas

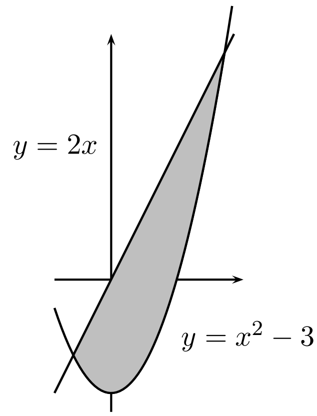


Figure 5.25: Area between $y = x^2 - 3$ and $y = 2x$

represented by various definite integrals of $x^2 - 3$ and $2x$, we could work this out (see Exercise 1). However, we will go at this more directly, following the kind of reasoning we used in defining the definite integral. We start by partitioning the interval $[-1, 3]$ of x-values that appear in S, using the partition

$$\mathcal{P} = \{-1 = p_0 < p_1 < \cdots < p_n = 3\}.$$

The vertical lines $x = p_j$, $j = 1, \ldots, n-1$ partition S into n vertical slices (Figure 5.26): the j^{th} slice S_j is the part of S lying above the component interval $I_j = [p_{j-1}, p_j]$:

$$S_j = \{(x, y) \,|\, x \in I_j \text{ and } x^2 - 3 \leq y \leq 2x\}.$$

If the slice is narrow (that is, $\|I_j\| = p_j - p_{j-1} = \triangle x_j$ is small) then the values of $x^2 - 3$ and of $2x$ don't vary too much over $p_{j-1} \leq x \leq p_j$, so we can pick $x_j \in I_j$ and approximate S_j by a rectangle whose top is at $y = 2x_j$ and whose bottom is at $y = x_j^2 - 3$. This rectangle has height

$$h_j = (2x_j) - (x_j^2 - 3) = 2x_j - x_j^2 + 3$$

and width $\triangle x_j$, so its area (which approximates that of S_j) is

$$\triangle A_j = h_j \triangle x_j = (2x_j - x_j^2 + 3)\triangle x_j.$$

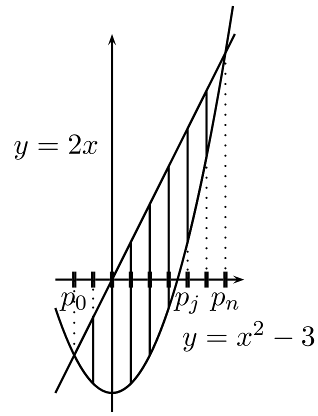


Figure 5.26: Slicing the Area

The area of S is equal to the sum of the areas of the slices S_j, $j = 1, \ldots, n$, which in turn is approximated by the sum of the areas of the rectangles:

$$A(S) = \sum_{j=1}^{n} A(S_j) \approx \sum_{j=1}^{n} \triangle A_j = \sum_{j=1}^{n} (2x_j - x_j^2 + 3)\triangle x_j.$$

If we take a sequence of partitions $\mathcal{P}_N$ with mesh size going to zero, our approximations get successively better, so we expect in turn that these approximate calculations should converge to the true value of $A(S)$. But our approximations are just Riemann sums for the the integral of the function

$$h(x) := (2x) - (x^2 - 3)$$

over the interval $[-1, 3]$, so we conclude that

$$A(S) = \lim_{N\to\infty} \mathcal{R}(\mathcal{P}_N, h) = \int_{-1}^{3} (2x - x^2 + 3)\, dx.$$

We can calculate this integral using the Fundamental Theorem of Calculus:

$$\begin{aligned}\int_{-1}^{3} (2x - x^2 + 3)\, dx &= \left(x^2 - \frac{x^3}{3} + 3x\right)\Big|_{-1}^{3} \\ &= (9 - \frac{27}{3} + 9) - (1 - \frac{1}{3} - 3) = \frac{34}{3}.\end{aligned}$$

What we have given is a "modern" elucidation of a way of thinking that was already used by the Greeks in thinking about areas and volumes, and which was used by Johannes Kepler (1571-1630) in his 1615 treatise on wine bottles; while Kepler's name is now associated primarily with his work in astronomy, he was also a major mathematician. The work in question [34] developed methods for measuring with precision the amount of wine left in a partially filled bottle, as well as the most efficient shapes for wine barrels. The method was developed systematically by Bonaventura Cavalieri (1598-1647), who had studied with Galileo. Cavalieri begins with an observation that is now called "Cavalieri's Theorem", which he stated as follows [51, p. 210] (see Figure 5.27)

> **Cavalieri's Theorem:** *If between the same parallels any two plane figures are constructed, and if in them, any straight lines being drawn equidistant from the parallels, the included portions of any one of these lines are equal, the plane figures are also equal to one another...*

(He then goes on to state the analogous property for volumes of regions in space.) This says, in our language, that if corresponding slices for two regions have the same length, then the two regions have the same area. Cavalieri then used the idea of summing up all of these (infinitely many) slices to get an area. While the procedure in this (pre-Fundamental Theorem of Calculus) form may be suspect to our eyes, it does, when translated into an algorithm for writing down appropriate integrals, give an amazingly efficient and correct procedure.

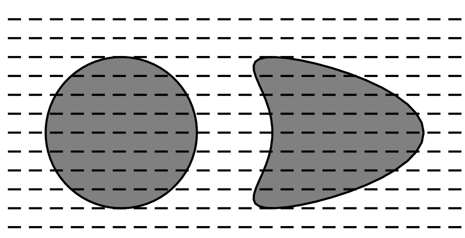


Figure 5.27: Cavalieri's Theorem

A slightly more complicated problem is the area (Figure 5.28) bounded by the graphs

$$y = x^3 - 3x, \quad y = x$$

and the vertical lines

$$x = -1, \quad x = 2.$$

The complication here is that the two graphs cross at $x = 0$, so we have

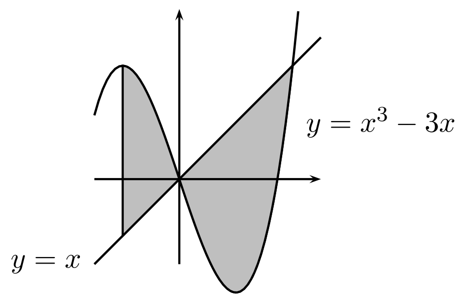


Figure 5.28: Area between $y = x^3 - 3x$ and $y = x$, $-1 \le x \le 2$

$$x \le x^3 - 3x \text{ for } -1 \le x \le 0$$
$$x \ge x^3 - 3x \text{ for } 0 \le x \le 2.$$

So when we partition the interval $[-1, 2]$, it is convenient to make sure that we include the crossover point $x = 0$, and treat this as two separate problems:

- for $-1 \le x < 0$, the slice at x_j has its *top* at (approximately) $y = x_j^3 - 3x_j$ and its *bottom* at (approximately) $y = x_j$; thus its height is approximated by

  $$h_j = (x_j^3 - 3x_j) - (x_j) = x_j^3 - 4x_j \quad -1 \le x < 0$$

  and its area is (approximately)

  $$\triangle A_j = h_j \triangle x_j = (x_j^3 - 4x_j)\triangle x_j;$$

  this means the area of the part of S to the left of $x = 0$ is approximated by a sum of the form

  $$\sum \triangle A_j = \sum (x_j^3 - 4x_j)\triangle x_j \quad -1 \le x_j < 0;$$

- for $0 < x \leq 2$, the slice at x_j has its *bottom* at (approximately) $y = x_j^3 - 3x_j$ and its *top* at (approximately) $y = x_j$; thus its height is approximated by

$$h_j = (x_j) - (x_j^3 - 3x_j) = 4x_j - x_j^3 \quad 0 < x \leq 2.$$

Thus, in this range

$$\triangle A_j = h_j \triangle x_j = (4x_j - x_j^3) \triangle x_j \quad 0 < x \leq 2$$

and the total area of the part of S to the right of the y-axis is approximated by a sum of the form

$$\sum \triangle A_j = \sum (4x_j - x_j^3) \triangle x_j \quad 0 < x_j \leq 2.$$

As we subdivide further ($\text{mesh}(\mathcal{P}_N) \to 0$), these two sums converge, respectively, to the definite integrals

$$\int_{-1}^{0} (x^3 - 4x)\, dx \text{ and } \int_{0}^{2} (4x - x^3)\, dx.$$

These are easily calculated using the Fundamental Theorem of Calculus:

$$\int_{-1}^{0} (x^3 - 4x)\, dx = \left(\frac{x^4}{4} - 2x^2 \right) \Big|_{-1}^{0} = 0 - (\frac{1}{4} - 2) = \frac{7}{4}$$
$$\int_{0}^{2} (4x - x^3)\, dx = \left(2x^2 - \frac{x^4}{4} \right) \Big|_{0}^{2} = (8 - \frac{16}{4}) - 0 = 4$$

and it follows that

$$A(S) = \int_{-1}^{0} 9x^3 - 4x)\, dx + \int_{0}^{2} (4x - x^3)\, dx = \frac{7}{4} + 4 = 5\frac{3}{4}.$$

We could have written the area in a single integral by noting that the height is in either case the distance between the two ends of each slice, which is the absolute value of their difference:

$$A(S) = \int_{-1}^{2} \left| (x^3 - 3x) - (x) \right|\, dx = \int_{-1}^{2} \left| x^3 - 4x \right|\, dx.$$

This does not, however, make the calculation any easier: we still carry it out by breaking $[-1, 2]$ into the two intervals $[-1, 0]$ and $[0, 2]$ and calculating the integrals separately, as above.

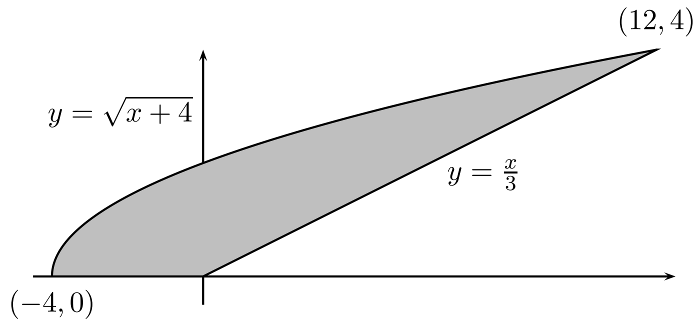


Figure 5.29: Area bounded by $y = \sqrt{x+4}$, $y = \frac{x}{3}$ and $y = 0$

As our third example, we consider the area bounded by the curves

$$y = \sqrt{x+4}, \quad y = \frac{x}{3}, \quad \text{and } y = 0.$$

This is a triangle-like region (Figure 5.29): a vertical slice to the *left* of the y-axis will have its top at $y = \sqrt{x+4}$ and its bottom on the x-axis, while a vertical slice to the *right* of the y-axis still has its *top* at $y = \sqrt{x+4}$, but now its *bottom* is at $y = x$. This would require two integrals. However, if we slice *horizontally* (Figure 5.30), our slices all have their *left* end at

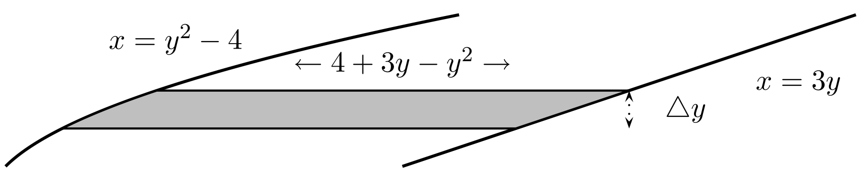


Figure 5.30: Horizontal Slice

$y = \sqrt{x+4}$, or

$$x = y^2 - 4$$

and their *right* end at

$$x = 3y.$$

So the "width" of each slice is (roughly)

$$3y - (y^2 - 4) = 4 + 3y - y^2$$

and its "height" is $\triangle y$. The area of one slice is approximately

$$\triangle A_j = (4 + 3y - y^2)\triangle y.$$

Since the values of y in this region run from $y = 0$ to $y = 4$), when we add and take an appropriate limit we get

$$\begin{aligned} A = \int_0^4 (4 + 3y - y^2) dy &= (4y + \frac{3y^2}{2} - \frac{y^3}{3})\Big|_0^4 \\ &= 16 + 24 - \frac{64}{3} = \frac{56}{3}. \end{aligned}$$

Areas in Polar Coordinates

You are probably familiar with the system of **polar coordinates**: instead of being located by its distance from the x- and y-axes, a point is located by specifying a position r on the x-axis and then specifying an angle θ (which we always give in radians) through which the x-axis is rotated counterclockwise. When $r \geq 0$, it corresponds to the distance of the point from the origin. The two systems of coordinates are related as follows (Figure 5.31): if the Cartesian coordinates of a point P are (x, y) and its

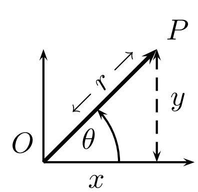


Figure 5.31: Polar Coordinates

polar coordinates are (r, θ), then

$$\begin{aligned} x &= r\cos\theta & r^2 &= x^2 + y^2 \\ y &= r\sin\theta & \tan\theta &= \frac{y}{x}. \end{aligned}$$

Polar coordinates are particularly useful in dealing with situations of rotational symmetry, but there are disadvantages as well. The most important of these is that the same geometric point can be specified by many different pairs of polar coordinates: for example, (r, θ), $(-r, \theta + \pi)$ and $(r, \theta + 2\pi)$ all give the same position; even worse, if $r = 0$, we are at the origin regardless of the value of θ.

Here we investigate the area enclosed by some curves defined by specifying r as a function of θ; these arise naturally in polar coordinates but not in Cartesian coordinates.

As a first example, we consider the curve specified by

$$r = 1 - \sin\theta.$$

If we imagine the x-axis rotating counterclockwise through one revolution, we see that the point with polar coordinates $(1 - \sin\theta, \theta)$ starts one unit out along the positive x-axis ($r = 1 - 0 = 1$) when $\theta = 0$, and as the axis rotates, the point moves closer to the origin, hitting it at $\theta = \frac{\pi}{2}$ ($r = 1 - 1 = 0$). Then the point starts to move away again, reaching distance $r = 1$ again at $\theta = \pi$, and continuing until it is $r = 1 - (-1) = 2$ units away when $\theta = \frac{\pi}{2}$; then it moves back in, reaching its initial position when $\theta = 2\pi$, and thereafter traces out the curve again and again. Figure 5.32 shows a plot of this curve, which is called a **cardioid** (guess why?)

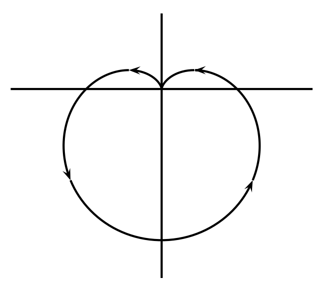


Figure 5.32: Cardioid $r = 1 - \sin\theta$

To calculate the area enclosed by this curve, we try to estimate the area $\triangle A$ swept out by the moving "stick" of radius $r = 1 - \sin\theta$ as θ increases a small amount, say from $\theta = \theta_0$ to $\theta = \theta_0 + \triangle\theta$. For $\triangle\theta$ small, r stays approximately constant, so the area is approximately that of a wedge with angle $\triangle\theta$ cut out of a circle of radius r; since $\triangle\theta$ constitutes a fraction $\frac{\triangle\theta}{2\pi}$ of a full revolution, and a full revolution encloses the full area of the disc enclosed by the circle (with area πr^2), we see that in general,

$$\triangle A = \frac{\triangle\theta}{2\pi}(\pi r^2) = \frac{r^2}{2}\triangle\theta.$$

In our case, if we cut the full "sweep" into sectors defined by inequalities of the form $\theta_{i-1} \leq \theta \leq \theta_i$, $i = 0, \ldots, N$ (where $\theta_0 = 1$ and $\theta_N = 2\pi$, we see that the i^{th} sector has area

$$\triangle A_i \approx \frac{r_i^2}{2}(\theta_i - \theta_{i-1}) \approx \frac{(1-\sin\theta_i)^2}{2}\triangle\theta_i$$

and so the total area is approximated by the sum

$$A \approx \sum_{i=1}^{N} \frac{(1-\sin\theta_i)^2}{2}\triangle\theta_i \to \int_0^{2\pi} \frac{(1-\sin\theta)^2}{2} d\theta.$$

We calculate this integral with the help of the double-angle formula:

$$\begin{aligned}
\int_0^{2\pi} \frac{(1-\sin\theta)^2}{2} d\theta &= \frac{1}{2}\int_0^{2\pi} [1 - 2\sin\theta + \sin^2\theta] d\theta \\
&= \frac{1}{2}\int_0^{2\pi} [1 - 2\sin\theta + \frac{1}{2}(1-\cos 2\theta)] d\theta \\
&= \frac{1}{2}[\theta + 2\cos\theta + \frac{\theta}{2} - \frac{1}{4}\sin 2\theta]\Big|_0^{2\pi} \\
&= \frac{3}{4}(2\pi) = \frac{3\pi}{2}.
\end{aligned}$$

As a second example, we consider the curve specified by

$$r = \sin 2\theta.$$

The point with polar coordinates $(\sin 2\theta, \theta)$ starts at the origin $(r = 0)$ when $\theta = 0$, then moves out (as the axis rotates) to a maximal distance of $r = 1$ when $\theta = \frac{\pi}{4}$, then comes back to the origin at $\theta = \frac{\pi}{2}$; subsequently, r goes *negative* in the range $\frac{\pi}{2} < \theta < \pi$, which means it traces out a closed curve (similar to the initial "loop") in the *fourth* quadrant. Next, r again becomes positive for $\pi < \theta < \frac{3\pi}{2}$, giving a "loop" in the *third* quadrant, and finally it goes negative again for $\frac{3\pi}{2} < \theta < 2\pi$, tracing out a "loop" in the *second* quadrant. After that, it traces out the same curve all over again. This curve is sometimes called a **four-petal rose** (Figure 5.33).

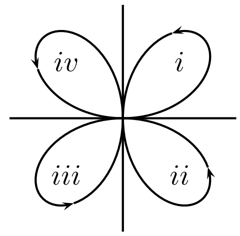


Figure 5.33: Four-petal Rose $r = \sin 2\theta$

Again, our earlier analysis says that as θ changes by $\triangle\theta_i$ from $\theta = \theta_{i-1}$ to $\theta = \theta_i$, the stick sweeps out an area roughly equal to

$$\frac{r_i^2}{2}\triangle\theta = \frac{(\sin 2\theta_i)^2}{2}\triangle\theta_i$$

and again summing and going to the limit as $N \to \infty$ we obtain the area as

$$\begin{aligned} A = \int_0^{2\pi} \frac{(\sin 2\theta)^2}{2} d\theta &= \frac{1}{4}\int_0^{2\pi} (1 - \cos 4\theta) d\theta \\ &= \frac{1}{4}(\theta - \frac{1}{4}\sin 4\theta)\Big|_0^{2\pi} \\ &= \frac{\pi}{2}. \end{aligned}$$

Volume of Revolutes

A **revolute** is the shape in 3-space obtained by rotating a curve or region in the plane about an axis. It is assumed that the curve or region lies on one side of the axis: it can have boundary points, but no interior points, on the axis.

A familiar example of a revolute is a sphere. The sphere of radius R can be obtained by rotating a semicircle of radius R about its diameter. The semicircle can be described as the upper half of the curve $x^2 + y^2 = R^2$, which is to say as the graph of

$$y = \sqrt{R^2 - x^2}, \quad -R \le x \le R;$$

to obtain the sphere, we rotate this about the x-axis (Figure 5.34). The

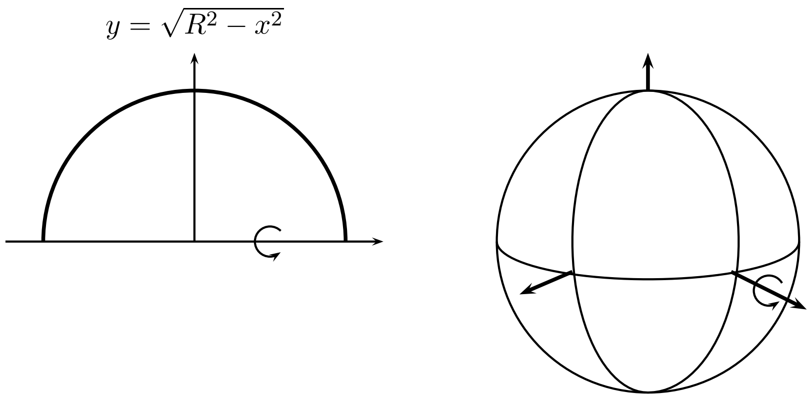


Figure 5.34: The sphere of radius R as a revolute.

interior of the sphere (the *ball* of radius R) is obtained by rotating the region

$$S = (x, y) - R \leq x \leq R,\ 0 \leq y \leq \sqrt{R^2 - x^2}$$

between the semicircle and the x-axis (*i.e.*, the *half-disc*) about the x-axis. We shall use this to derive the formula for the volume of the ball.
As in the case of planar areas, we cut our region into vertical strips S_j by means of a partition

$$\mathcal{P} = \{-R = p_0 < p_1 < \cdots < p_n = R\},$$

and approximate each strip

$$S_j = \{(x, y) \mid x \in I_j,\ 0 \leq y \leq \sqrt{R^2 - x^2}\}$$

by a rectangle going from $y = 0$ to $y = \sqrt{R^2 - x_j^2}$ for some $x_j \in I_j$.
Rotating this rectangle about the x-axis results in a coin-like solid[28] (Figure 5.35): its thickness is the length $\triangle x_j$ of I_j, while its *radius* is the height

$$h_j = \sqrt{R^2 - x^2}$$

of the rectangle. The volume of such a coin is the area of the face of the coin times its thickness, so we get as an approximation to the "slice" of the ball obtained by rotating S_j

$$\triangle V_j \approx \pi h_j^2 \triangle x_j = \pi (R^2 - x_j^2) \triangle x_j.$$

[28]Often called a "disc" in Calculus books.

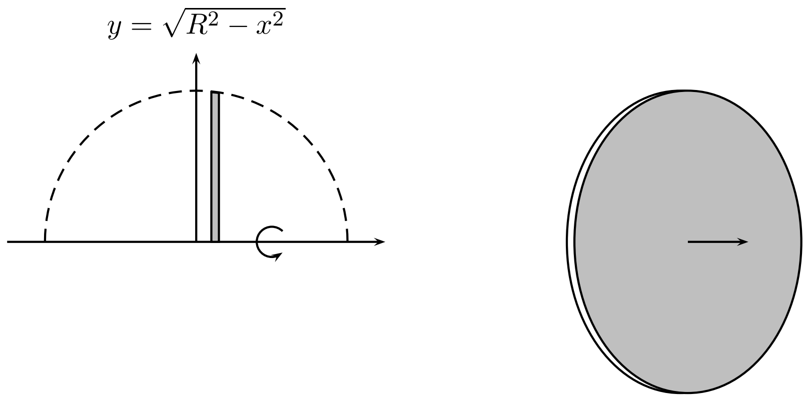


Figure 5.35: Rotating a slice.

Adding these slices from $x = -R$ to $x = R$, we obtain a Riemann sum

$$\sum_{j=1}^{n} \triangle V_j \approx \sum_{j=1}^{n} \pi(R^2 - x_j^2)\triangle x_j$$

which converges, as we take thinner slices, to

$$V = \lim_{\mathcal{P}_N} \sum_{j=1}^{n} \triangle V_j = \lim_{\mathcal{P}_N} \mathcal{R}(\mathcal{P}_N, \pi(R^2 - x^2)) = \int_{-R}^{R} \pi(R^2 - x^2)\, dx.$$

This integral is easy to compute:

$$\int_{-R}^{R} \pi(R^2 - x^2)\, dx = \pi\left(R^2 x - \frac{x^3}{3}\right)\Big|_{-R}^{R}$$
$$= \pi\left(R^3 - \frac{R^3}{3}\right) - \pi\left(-R^3 + \frac{R^3}{3}\right) = \frac{4}{3}\pi R^3$$

which is the well-known volume formula for the ball of radius R.
The idea of looking at volumes via "slices" was probably known to Democritus (*ca.* 460-370BC), who is credited by Archimedes of Syracuse (*ca.* 212-287 BC) as having first discovered that the volume of a cone is one-third that of a cylinder with the same base and height; Archimedes credits Eudoxus of Cnidos (408-355 BC) with the first proof of this fact. His own proof of this, as well as of the theorems that he was most proud

of[29], that the volume and the surface area of a sphere are each two-thirds those of a circumscribed cylinder, were based on "slicing" these figures and then calculating the volumes (and areas) of the slices via approximation by inscribed polyhedral fulcrums. He also, in *The Method*, reveals that he first discovered these facts via a mechanical argument, again involving the balancing of slices against those of a known body.
As a second example, consider the "trumpet-shaped" revolute which results from rotating the region below the graph $y = x^2$, $0 \le x \le 1$

$$S = \{(x, y) \,|\, 0 \le x \le 1,\ 0 \le y \le x^2\}$$

about the x-axis (Figure 5.36).

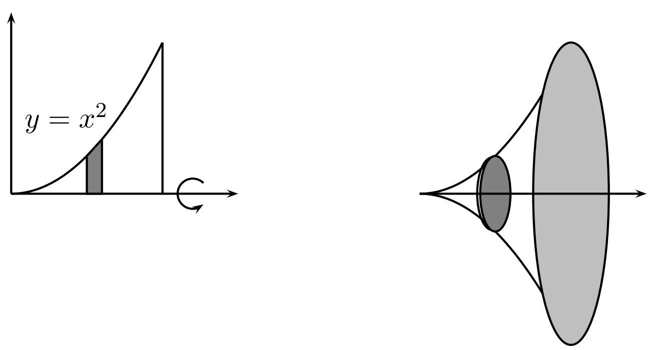


Figure 5.36: Rotating the region $0 \le y \le x^2$, $0 \le x \le 1$ about the x-axis

Again, if we slice S vertically using a partition of $[0, 1]$, we get pieces S_j which are approximated by rectangles of width $\triangle x_j$ and height $h_j = x_j^2$, so the resulting coins have volume

$$\triangle V_j = \pi h_j^2 \triangle x_j = \pi x_j^4 \triangle x_j.$$

The sum of these volumes over $0 \le x \le 1$ converges (as we slice thinner) to a definite integral

$$\sum_{j=1}^{n} \triangle V_j = \sum_{j=1}^{n} \pi x_j^4 \triangle x_j \to \int_0^1 \pi x^4 \, dx = \left.\frac{\pi x^5}{5}\right|_0^1 = \frac{\pi}{5}.$$

[29]It is reported that Archimedes asked that his tombstone be inscribed with a figure of a sphere and a circumscribed cylinder, together with the formulas giving these ratios. In 75BC, Marcus Tullius Cicero (106-43BC), the Roman *quaestor* ruling western Sicily, reported finding the tomb and restoring it with its inscription.

Suppose now that we rotate the same region S about the y-axis instead of the x-axis (Figure 5.37). If we slice *horizontally* instead of *vertically*, we

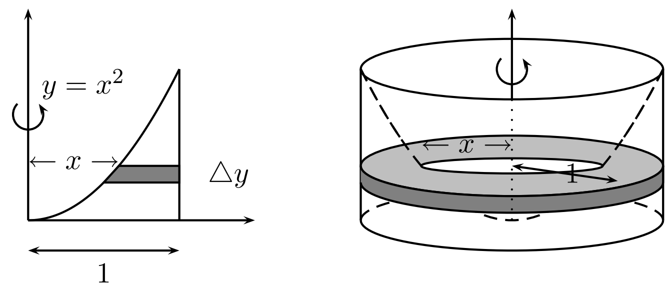


Figure 5.37: Rotating the region $0 \leq y \leq x^2$, $0 \leq x \leq 1$ about the y-axis

will get coins—but of the ancient sort, with a hole in the middle[30]. This can be formulated by writing our region using x in terms of y. The curve $y = x^2$ for $0 \leq x \leq 1$ can be rewritten as

$$x = \sqrt{y}, \quad 0 \leq y \leq 1.$$

A partition

$$\mathcal{P} = \{0 = p_0 < p_1 < ... < p_n = 1\}$$

of the interval $0 \leq y \leq 1$ slices S into *horizontal* rectangles S^j whose height is the thickness $\triangle y_j$; the j^{th} slice has its left end at $x = \sqrt{y}$ and its right end at $x = 1$. However, this slice is being rotated not about its left edge, but rather about the y-axis. If the slice were to stretch all the way to the y-axis, we would indeed have a coin whose radius is the distance from the y-axis to the right edge of S^j, at $x = 1$; the volume of such a coin would indeed be

$$\pi(1)^2 \triangle y_j.$$

However, since the *left* edge of the region is at $x = \sqrt{y_j}$, this gives the radius of the *hole* in the coin. The volume of the hole

$$\pi(\sqrt{y_j})^2 \triangle y_j$$

must be subtracted from the earlier volume to give the true volume of this

[30] These are often called "washers" in Calculus books.

slice of the revolute[31]:

$$\begin{aligned}\triangle V_j &= \pi(1)^2 \triangle y_j - \pi(\sqrt{y_j})^2 \triangle y_j \\ &= \pi(1 - y_j)\triangle y_j.\end{aligned}$$

Finally, we add these up and take the limit to get our integral formula

$$\begin{aligned}V = \lim_{\mathcal{P}_N} \sum_{j=1}^{n} \triangle V_j &= \lim_{\mathcal{P}_N} \sum_{j=1}^{n} \pi(1 - y_j)\triangle y_j \\ &= \int_0^1 \pi(1 - y)dy = \pi \left(y - \frac{y^2}{2}\right)\Big|_0^1 = \frac{\pi}{2}.\end{aligned}$$

There is another approach that can be used for this problem, which can also be useful in other settings. In the previous examples, we obtained "coins" by slicing our region *perpendicular* to the axis of rotation. If instead we slice *parallel* to the axis, we find ourselves rotating relatively long, narrow strips parallel to our axis, resulting in a family of thickened cylinders or concentric tin cans[32] approximating our solid.

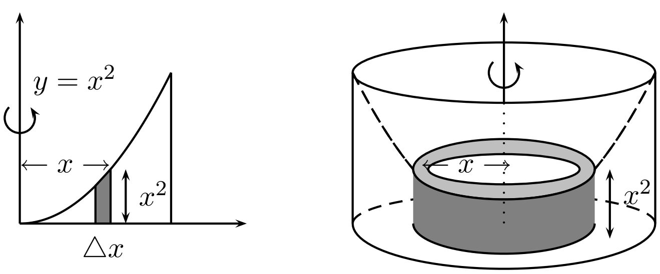


Figure 5.38: Cylinders for $0 \le y \le x^2$, $0 \le x \le 1$ rotated about the y-axis

In our case, this occurs when we slice S using a partition of the interval $0 \le x \le 1$: the j^{th} slice is a sliver S_j of height

$$h_j = x_j^2$$

and thickness $\triangle x_j$. Rotating this about the y-axis (whose distance from the sliver is x_j) gives a thin shell whose height corresponds to the height h_j

[31]Note that this involves the difference of squares, *not* the square of the difference.
[32]Often called "shells".

of the sliver, whose radius is x_j, and whose thickness is $\triangle x_j$. To measure the volume of the shell (*NB:* the amount of tin in the can, not how much soup it can hold) we imagine making a vertical cut in the side of the can and then opening it out into a flat sheet of tin, with height h_j, thickness $\triangle x_j$, and width equal to the *circumference* of a circle whose radius is x_j (Figure 5.39): thus the width is

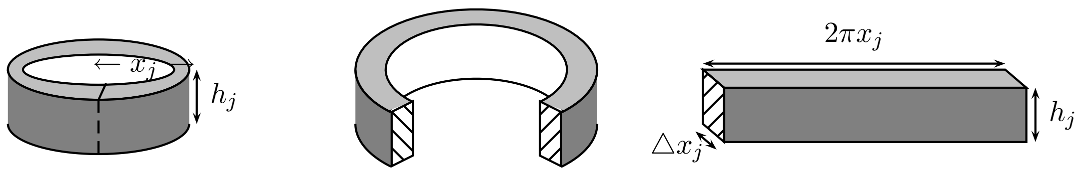


Figure 5.39: Cutting open a shell.

$$w_j = 2\pi x_j,$$

and the volume of one sheet is

$$\triangle V_j = w_j h_j \triangle x_j = 2\pi x_j (x_j)^2 \triangle x_j.$$

Adding up and taking limits, we obtain

$$\begin{aligned} V = \lim_{\mathcal{P}_N} \sum_{j=1}^{n} \triangle V_j &= \lim_{\mathcal{P}_N} \sum_{j=1}^{n} 2\pi x_j (x_j)^2 \triangle x_j \\ &= \int_0^1 2\pi x^3\, dx = 2\pi \frac{x^4}{4} \Big|_0^1 = \frac{\pi}{2} \end{aligned}$$

as before.

Exercises for § 5.8

Answers to Exercises 1, 2ace, 3a, 4, 5ace, 6 are given in Appendix B.

Practice problems:

1. Find the area bounded above by the curve $y = 2x$ and below by $y = x^2 - 3$ (Figure 5.25) as follows:

(a) Divide the area in question into three regions, one in each quadrant (Note: the region does not intersect the quadrant $x < 0$, $y > 0$).

(b) Express each of these regions in terms of (unions and/or differences of) sets of the form

$$\{(x, y) \,|\, a \leq x \leq b,\ y \text{ between the graph of } f \text{ and the } x\text{-axis}\}.$$

(c) Express the total area as sums and differences of integrals of $f(x) = 2x$ and/or $f(x) = x^2 - 3$ over different intervals.

2. For each pair of curves below, sketch the region they enclose and find its area.

   (a) $y = x^2$ and $y = 8 - x^2$.

   (b) $y = x^2$ and $x = y^2$.

   (c) $y = x^3 - 4x$ and $y = 5x$.

   (d) $y = x^3 - 2x$ and $y = 3x$.

   (e) $y = 72 - 40x - 16x^2 = 97 - (4x + 5)^2$ and $y = 16\,|x|$.

   (f) The curves $y = \cos x$ and $y = \sin x$ intersect infinitely often, cutting off a sequence of adjacent regions. Find the area of the region which crosses the y-axis.

3. Sketch the curve given by each polar equation below, and find the area it encloses.

   (a) $r = 2\cos\theta$

   (b) $r = \sin 3\theta$.

4. (a) Sketch the curves given by the polar equations $r = \sin 3\theta$ and $r = \sin 4\theta$. (*Caution:* the relation between the coefficient in front of θ and the number of "leaves" is different for the two curves.)

   (b) Find a formula for the area enclosed by the curve $r = \sin n\theta$ where n is a positive integer. (The even and odd cases are different.)

5. In each problem below, you are given a set of curves and an axis of rotation; you are to find the volume of the revolute obtained by rotating the region bounded by the given curves about the given

axis, as follows: For each, (i) sketch the region bounded by the given curves, and indicate the axis of rotation; (ii) set up an integral expressing the volume obtained by slicing *perpendicular* to the axis of rotation; (iii) set up an integral expressing the volume obtained by slicing *parallel* to the axis of rotation; and (iv) evaluate at least one of your integrals to find the volume.

(a) $y = \sqrt[3]{x}$, $y = 1$, and the y-axis, rotated about the x-axis.

(b) $y = \sqrt[3]{x}$, $x = 1$, and the x-axis, rotated about the y-axis.

(c) $y = \sqrt{x}$ and $y = x^2$, rotated about the x-axis.

(d) $y = \sin x$, $y = \cos x$, and the y-axis, rotated about the x-axis. (The region in question should be only up to the first intersection of these two graphs to the right of the y-axis.)

(e) $y = \sin x$, $y = \cos x$, and the y-axis, rotated about the y-axis.

6. A hole of radius R is bored through the center of a sphere of radius $2R$. What is the volume that remains?

Challenge problem:

7. (a) Use vertical slices to calculate the area of a circle of radius R.

   (b) Let us reflect a bit on the logic of this: to perform the integration, you probably substituted $x = R\sin\theta$ and ended up calculating a trigonometric integral. This means you used the differentiation formulas for $\sin\theta$ and $\cos\theta$ (Proposition 4.2.8). These in turn rely on the limit formulas found in Lemma 3.4.8 and Exercise 7 in § 3.4. Go back and look at how we established these: they rely on our knowing the area of a circle of radius 1! How do we get out of this logical circle?

5.9 Riemann's Characterization of Integrable Functions (Optional)

Bernhard Georg Friedrich Riemann (1826-1866) was born in the province of Hannover in Germany. The bulk of his professional life was spent at the University of Göttingen, the major university in the province, which during the latter half of the nineteenth and first third of the twentieth

century was arguably the center of mathematics in the world[33]. Riemann entered the university at the age of nineteen in 1846, completing his doctorate in 1851 and *Habilitation* (vaguely analogous to tenure) there in 1854; in 1855 Carl Friedrich Gauss (1777-1855), who held the Chair in Mathematics, died, and the post was taken by Johann Peter Gustav Lejeune Dirichlet (1805-1859), who had already come to know and deeply influence Riemann during two years the latter had spent in Berlin early in his studies. In 1859 Dirichlet died and Riemann took over the chair; toward the end of his life, fighting consumption, he spent a good deal of time in Italy, especially Pisa. He died in 1866 in Italy, at the age of 40. He published nine papers during his lifetime, and another seven were published shortly after his death; nonetheless he had an enormous influence on many areas of mathematics, including complex function theory, integration, number theory and geometry, as well as physics.
The process of obtaining *Habilitation* was and still is an involved one, and in particular a kind of advanced dissertation, called the *Habilitationsschrift*, was required. A very famous part of Riemann's *Habilitationsschrift*, entitled "On the Hypotheses which Lie at the Foundations of Geometry", is one of the seminal works of modern differential geometry. The other part, "On the Representability of a Function through Trigonometric Series" [53], concerned the study of what we now call "Fourier series". This study involved intense use of definite integrals, and in his work Riemann clarified and extended the theory of the definite integral from the form it had been given by Cauchy. Cauchy had defined the definite integral of a continuous function by dividing the interval into a number of (not necessarily equal) pieces and using the value of the function at the left end of each piece to determine the height of a rectangle over that piece. Using the uniform continuity of the function, as we do in § 5.1—which he didn't really distinguish from continuity as we know it—he proved that with successively finer divisions, the sums so obtained converge to a limit which is independent of the exact details of how the interval was divided (provided the mesh size goes to zero).
Riemann was trying to understand how general a function could be approximated by a trigonometric series, and so he wanted to understand precisely what was needed in a function to make sure his integral—which involved the arbitrary "sampling" we saw in § 5.2—was well-defined. It was already generally understood that functions with a finite number of jump discontinuities were integrable, but Riemann was looking further. He

[33]It is also known to physicists as the birthplace of quantum physics.

ended up with the following.
He defined the *variation* of a bounded function $f(x)$ on an interval I to be

$$D(f(x), I) := \sup_{x \in I} f(x) - \inf_{x \in I} f(x),$$

and considered, for any partition $\mathcal{P}$ of $[a, b]$ and any $\delta > 0$, the sum

$$\pounds(f(x), \delta, \mathcal{P})$$

of the lengths of those I_j for which

$$D(f(x), I_j) \geq \delta.$$

Then

Theorem 5.9.1 (Riemann-Integrable Functions)**.** *A bounded function $f(x)$ on the closed interval $[a, b]$ is integrable on this interval precisely if for every $\delta > 0$ and $\varepsilon > 0$ there exists Δ such that*

$$\pounds(f(x), \delta, \mathcal{P}) < \varepsilon \text{ whenever } mesh(\mathcal{P}) < \Delta.$$

Proof. First, it is clear that the condition is necessary, since for any partition $\mathcal{P}$

$$\mathcal{U}(\mathcal{P}, f) - \mathcal{L}(\mathcal{P}, f) \geq \delta \cdot \pounds(f(x), \delta, \mathcal{P}),$$

and this has to go to zero with the mesh size.
On the other hand, suppose the condition holds. Then for any partition $\mathcal{P}$,

$$\begin{aligned} &\mathcal{U}(\mathcal{P}, f) - \mathcal{L}(\mathcal{P}, f) \\ &\quad = \sum_{D(f(x), I_j) \geq \delta} D(f(x), I_j) \triangle x_j + \sum_{D(f(x), I_j) < \delta} D(f(x), I_j) \triangle x_j \\ &\quad < D(f(x), [a, b]) \pounds(f(x) . \delta, \mathcal{P}) + \delta(b - a) \end{aligned} \tag{5.17}$$

(verify this inequality!) and we can make the second term small by choosing δ, then make the first one small by making sure $\mathcal{P}$ has a sufficiently small mesh size. □

Let us apply this criterion to a few examples.
First, the Dirichlet "pathological" function

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational }, \\ 0 & \text{otherwise} \end{cases}$$

has $\mathcal{L}(f(x), \delta, \mathcal{P}) = b - a$ for any partition and any $\delta < 1$, so this function is clearly *not* integrable.
Now, consider the modification in which $f\left(\frac{p}{q}\right) = \frac{1}{q}$

$$f(x) = \begin{cases} \frac{1}{q} & \text{if } x = \frac{p}{q} \text{ in lowest terms,} \\ 0 & \text{otherwise.} \end{cases}$$

Now, for any integer q, there are at most

$$n(q) := q + (q-1) + ... + 2 + 1 + 1 = \frac{q(q-1)}{2} + 1$$

points with denominator q or less, so a partition $\mathcal{P}$ with mesh size less than

$$\frac{\varepsilon}{n(q)}$$

will have

$$\mathcal{L}(f(x), \delta, \mathcal{P}) < \varepsilon$$

for any $\delta \geq \frac{1}{q}$, and the function is *integrable*, even though it is discontinuous at the infinite number of rational points.
Riemann gave another example of a function $f(x)$ which is discontinuous at an infinite number of points but is integrable; we work this out in Exercises 1-5.

Exercises for § 5.9

Riemann's discontinuous integrable function: In his *Habilitationsschrift* [44], Riemann illustrated his characterization of integrable functions by constructing an integrable function that is discontinuous on a dense set of points[34].

1. Begin by defining

$$(x) := \begin{cases} x - n & \text{if } n - \frac{1}{2} < x < n + \frac{1}{2} \text{ for some integer } n, \\ 0 & \text{otherwise } (i.e., \text{ if } x = n + \frac{1}{2} \text{ for some integer } n). \end{cases}$$

[34] A subset of an interval is *dense* if every point of the interval is an accumulation point of the set.

Show that (x) is continuous at all points except odd multiples of $\frac{1}{2}$, where it has a (negative) "jump" of size 1:

$$\lim_{x \to n+\frac{1}{2}^-} = \frac{1}{2}$$
$$\lim_{x \to n+\frac{1}{2}^+} = -\frac{1}{2}.$$

2. For each $x \in \mathbb{R}$, define $f(x)$ by the infinite series

$$f(x) = \frac{(x)}{1^2} + \frac{(2x)}{2^2} + \frac{(3x)}{3^2} + ... = \sum_{k=1}^{\infty} \frac{(kx)}{k^2}.$$

Show that this series always converges absolutely (see Exercise 10 in § 5.9 or Lemma 6.2.1). (*Hint:* What are the possible values of (kx)?)

3. Show that given x, (nx) is discontinuous at x for some integer n precisely if $x = \frac{m}{2n}$, where m is odd. Show that if m and n are relatively prime, then for every *odd* multiple of n, say $k = n(2p+1)$, the term $\frac{(kx)}{k^2}$ has a "jump" of size

$$\frac{1}{k^2} = \frac{1}{n^2(2p+1)^2}.$$

Using this, it can be shown that $f(x)$ has a "jump" of size

$$\sum_{=p}^{\infty} 1 \left(\frac{1}{n^2(2p+1)^2} \right) = \frac{\pi^2}{8n^2}$$

(the summation $\sum_{p=1}^{\infty} \frac{1}{(2p+1)^2} = \frac{\pi^2}{8}$ was established earlier by Euler). It can also be shown that at all other points, $f(x)$ is continuous.

4. To check Riemann's criterion for integrability on $[a, b]$, we need some technical estimates on the variation of the function f on a subinterval.

First, given $N \in \mathbb{N}$, there are finitely many points in $[a, b]$ of the form $x = \frac{m}{2n}$ with m odd, m and n relatively prime, and $n < N$; call this number P_N. Show that any two points of this form are at least distance $\frac{1}{2N^2}$ apart, so a subinterval $I \subset [a,b]$ of length $\frac{1}{4N^2}$ contains at most one such point.

Second, show that if x and x' both belong a subinterval I which contains *no* points of this form, then for $k < N$,

$$(kx) - (kx') = k\{x - x'\}$$

so

$$\sum_{k=1}^{N-1} \{\frac{(kx)}{k^2} - \frac{(kx')}{k^2}\} = \left(\sum_{k=1}^{N-1} \frac{1}{k}\right) \{x - x'\}.$$

Also, since for *all* integers k and points $x \in [a,b]$

$$-\frac{1}{2} < (kx) < \frac{1}{2},$$

for any two points $x, x' \in [a,b]$

$$\left|(kx) - (kx')\right| < 1$$

and hence

$$\left|\sum_{k=N}^{\infty} \left(\frac{(kx)}{k^2} - \frac{(kx')}{k^2}\right)\right| < \sum_{k=N}^{\infty} \frac{1}{k^2}.$$

From this it follows that if $I \subset [a,b]$ is a subinterval of the above type, then the variation of f on I is less than

$$\left(\sum_{k=1}^{N-1} \frac{1}{k}\right) \|I\| + \sum_{k=N}^{\infty} \frac{1}{k^2}. \tag{5.18}$$

5. Finally, to check Riemann's criterion: given $\delta > 0$ and $\varepsilon > 0$, we need to find a partition of $[a,b]$ such that the sum $\pounds(f, \delta, \mathcal{P})$ of the lengths of the component intervals on which the variation exceeds δ is at most ε.

   To this end, pick $N \in \mathbb{N}$ such that

$$\frac{\pi^2}{8N^2} < \delta \tag{5.19}$$

and

$$\sum_{k=N}^{\infty} \frac{1}{k^2} < \frac{\delta}{2}; \tag{5.20}$$

then pick

$$\mu < \frac{1}{4N^2} \tag{5.21}$$

such that

$$\mu < \frac{\varepsilon}{P_N} \tag{5.22}$$

(where P_N is as in the previous item) and

$$\left(\sum_{k=1}^{N} \frac{1}{k}\right) \mu < \frac{\delta}{2}. \tag{5.23}$$

Use Equation (5.18) and the conditions above to prove that any partition with mesh size less than μ has the desired property.

The perfection of approximation methods in which one uses series depends not only on the convergence of the series, but also on the ability to estimate the error resulting from neglected terms; in this regard one could say that almost all the approximation methods used in the solution of geometric and mechanical problems are highly imperfect. The preceding theorem may serve on many occasions to give these methods the perfection they lack, and without which it is often dangerous to employ them.

Joseph Louis Lagrange (1736-1813)
Théorie des Fonctions (1797), [37, p. 69]

6 Power Series

Infinite series played a role in calculus from the beginning: the method of exhaustion used by Euclid and Archimedes involved adding successive areas until the total area of some inscribed polygon was as close as desired to the area of a curvilinear region; in particular, the geometric series was in essence understood by the Greeks, although a general formula was only given in 1593 by François Vièta (1540-1603)[30, p. 20].
In the seventeenth century, infinite series were used extensively. The evaluation of a number of numerical series occurred during the early days of calculus, for example Leibniz's series (Exercise 46 in § 6.5); another famous infinite representation (this time an infinite product) was Wallis' representation of π (Exercise 45 in § 6.5). Also in this period began the extensive use of *power series*, the "infinite-degree" analogue of polynomials. They were central to Newton's work. One of his first and most famous results was the extension of the Binomial Theorem to fractional powers; we explore the resulting *Binomial Series* in Exercise 44, § 6.5; in fact, this formula was at his request put on his gravestone at Westminster Abbey. Mercator gave a power series representation of the natural logarithm, and James Gregory gave the first few terms of a power series representation of a number of basic functions, like the arctangent. Newton and Gregory both understood the principles behind what we now call Taylor series, although their first publication was in Brook Taylor's *Methodus incrementorum*, a textbook on Newton's fluxion calculus

published in 1715. The special case of the Taylor series about $x = 0$, now known as Maclaurin series, was developed as a tool and used extensively in Colin MacLaurin's *Treatise on Fluxions* of 1742.

The use of power series exploded during the eighteenth century, particularly at the hands of Johann Bernoulli (1667-1748) and the grand masters of the game, Leonard Euler (1701-1783), one of the most prolific mathematicians of all time, and Joseph Louis Lagrange (1736-1813). Much of their work was based on a fairly intuitive manipulation of series, although Lagrange tried (ultimately unsuccessfully) to make power series the basis of a rigorous calculus.

The rigorous study of series began with Augustin-Louis Cauchy (1789-1857) and Niels Henrik Abel (1802-1829) and was carried further by Bernhard Georg Friedrich Riemann (1826-1866) and later Karl Theodor Wilhelm Weierstrass (1815-1897) in the nineteenth century.

Much of the work of Euler and Lagrange also made free use of complex numbers; since we have not engaged in a study of complex analysis, many of their arguments are not available to us here, but in the last section of this chapter we explore the ways in which the use of power series allowed these and subsequent researchers to extend the calculus to the complex domain.

6.1 Local Approximation of Functions by Polynomials

Even though *in principle* a function unambiguously specifies the output value $f(x)$ resulting from a given input x, this doesn't guarantee that *in practice* this value can be located as a point on the number line. In Chapter 2, we faced a situtation of this sort for the square root function $f(x) = \sqrt{x}$ evaluated at $x = 2$, and the procedures we came up with for locating $\sqrt{2}$ could easily be adapted to finding $\sqrt{x}$ for any particular positive number x, provided that x itself is properly located on the number line. However, it would hardly be practical to carry out such a procedure every time we needed a square root. And for other functions, like $\sin x$, we would have to come up with an entirely different evaluation procedure. Fortunately, it is possible using tools from calculus to come up with a systematic procedure for evaluating most of the functions we deal with. This is based on a somewhat different point of view: instead of the trial-and-error approach we used in Chapter 2, we try to identify a new function ϕ which on one hand is relatively easy to calculate and, on the other, yields values acceptably close to those of our desired function f. We refer to such a function ϕ as an **approximation** to the function f; the size of their difference $|f(x) - \phi(x)|$ will be referred to as the **error** of the approximation at x. Of course, we shouldn't expect to actually calculate the error (just as we can't actually calculate the error in a decimal approximation to $\sqrt{2}$), but we hope to establish an **error bound**, showing that the error stays within certain bounds (at least for certain values of x), indicating the **accuracy** of the approximation. We shall, in this section, be interested in *local approximations*, that is, to approximations that have good accuracy when x is near some particular value $x = a$.

Which functions can we *really* calculate in practice? The calculations we can carry out by hand are addition, subtraction and multiplication. (Division is far more complicated, and in most cases gives us answers that we only know approximately; consider the decimal expansion for 1/73). This means that the functions we can compute are **polynomials**

$$\phi(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n.$$

The complexity of calculation for a polynomial is roughly measured by its **degree**: the highest power n with a nonzero coefficient: $\alpha_n \neq 0$.

The "simplest" polynomials, by this measure, are those of degree zero, or constant functions. These don't provide very useful approximations for most functions. However, we note in passing that, given a function f which is *continuous* at $x = a$, the constant approximation $\phi(x) = f(a)$ has the property that the error goes to zero as $x \to a$.

Things get more interesting one degree up. In general, the graph of a polynomial of degree one is a straight line. If we want to study f near $x = a$, we naturally look at the line tangent at $x = a$ to the graph of f—that is, we look at the degree one polynomial

$$\phi(x) = \alpha_0 + \alpha_1 x$$

satisfying the conditions

$$\phi(a) = f(a)$$
$$\phi'(a) = f'(a).$$

We will refer to these conditions by saying the graphs of f and ϕ have **first-order contact** at $x = a$. These conditions specify a particular function which we denoted $T_a f$ in § 4.1; recall that using the point-slope formula, we can write this in the form

$$\phi(x) = T_a f(x) = f(a) + f'(a)(x - a).$$

Let us investigate the properties of $\phi = T_a f$ as an approximation to f near $x = a$. Since ϕ and f are both continuous at $x = a$, we immediately have that the error $|f(x) - T_a f(x)| \to 0$ as $x \to a$. But this is no improvement over constants. We might, however, investigate the **error ratio**: the ratio of the error $|f(x) - \phi(x)|$ to the distance $|x - a|$ from x to a:

$$\frac{|f(x) - \phi(x)|}{|x - a|} = \left|\frac{f(x) - \phi(x)}{x - a}\right|.$$

The error ratio can be viewed as measuring how much accuracy in the outputs can be squeezed out of a given input accuracy. We note that the quantity inside the absolute values is a candidate for L'Hôpital's rule (Proposition 4.10.1), which tells us that

$$\lim_{x\to a} \frac{f(x) - \phi(x)}{x - a} = \lim_{x\to a} \frac{f'(a) - \phi'(a)}{1} = 0.$$

In the "oh" notation of § 4.9, we can rewrite this as

$$|f(x) - \phi(x)| = \mathfrak{o}(|x - a|) \text{ as } x \to a.$$

Note that we have not really used the assumption that ϕ is a first-degree polynomial, but only its first-order contact with f at $x = a$. That is, we have shown

Lemma 6.1.1. *If f and ϕ are both differentiable at $x = a$ and have first-order contact there*

$$\begin{aligned} \phi(a) &= f(a) \\ \phi'(a) &= f'(a) \end{aligned}$$

then the error ratio for $\phi(x)$ as an approximation to $f(x)$ goes to zero as $x \to a$; that is,

$$|f(x) - \phi(x)| = \mathfrak{o}(|x - a|)$$

as $x \to a$. (See Figure 6.1.)

What does this tell us about the accuracy of $\phi(x)$ as an approximation to $f(x)$? Suppose that we know that the error ratio is less than 10^{-r} (r an integer) for x within a certain distance of a. For x to be near a, their decimal expressions must agree to a certain number of digits: if

$$|x - a| < 10^{-k}$$

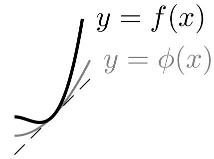


Figure 6.1: First-order contact

then the two expressions agree to the k^{th} place beyond the decimal point. But then

$$|f(x) - \phi(x)| < (10^{-r})(10^{-k}) = 10^{-(k+r)}$$

so that the outputs agree to $k + r$ places. Saying that $|f(x) - \phi(x)| = \mathfrak{o}(|x - a|)$ means that, by making k large, we can get r as large as we like: for x near a, we get many extra digits of accuracy in the output.

The preceding paragraph is really only a *qualitative* statement (just how big must k be, for example, to achieve $r = 2$?). However, in practice, this tells us that the tangent line gives a very good approximation to the original function. We consider two examples.

First, let us use this to calculate some square roots, say $\sqrt{80}$ and $\sqrt{80.9}$. The square root function $f(x) = x^{1/2}$ has derivative $f'(x) = \frac{1}{2}x^{-1/2}$; in particular, we know with absolute precision the values of these two functions at $x = 81$:

$$\begin{aligned} f(81) &= 9 \\ f'(81) &= \frac{1}{18}. \end{aligned}$$

Thus, we have

$$\phi(x) = T_{81}f(x) = 9 + \frac{1}{18}(x - 81).$$

Then

$$\begin{aligned} \phi(80) = \quad & 9 + \frac{80-81}{18} = \quad 9 - \frac{1}{18} = \quad 8\frac{17}{18} \approx \quad 8.944444444 \\ \phi(80.9) = \quad & 9 + \frac{80.9-81}{18} = \quad 9 - \frac{1}{180} = \quad 8\frac{179}{180} \approx \quad 8.994444444 \end{aligned}$$

How do these values compare to the actual square roots? One way to measure this is to square the approximation. We have

$$\begin{aligned}
[\phi(80)]^2 = & \left[8\frac{17}{18}\right]^2 = & \left[\frac{161}{18}\right]^2 = & \frac{25921}{324} \approx & 80.00308 \\
[\phi(80.9)]^2 = & \left[8\frac{179}{180}\right]^2 = & \left[\frac{1619}{180}\right]^2 = & \frac{2621161}{32400} \approx & 80.9000308
\end{aligned}$$

Another (strictly speaking, less justified) way is to compare with a calculator or computer approximation to the roots themselves. We obtained

$$\begin{aligned}
\sqrt{80} &\approx 8.94427\ldots \\
\sqrt{80.9} &\approx 8.9944427\ldots.
\end{aligned}$$

By either measure, we see that a moderate distance from x to $a = 81$ yields a strikingly good accuracy of approximation.

As another example, let us try to compute $\sin\frac{\pi}{5}$. Here, $f(x) = \sin x$ and $f'(x) = \cos x$ have precisely known values[1] at $x = \frac{\pi}{4}$, so that the tangent function is

$$\begin{aligned}
T_{\pi/4}f(x) =& \sin\frac{\pi}{4} + (\cos\frac{\pi}{4})(x - \frac{\pi}{4}) \\
=& \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}(x - \frac{\pi}{4}).
\end{aligned}$$

Hence

$$T_{\pi/4}f\left(\frac{\pi}{5}\right) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}(\frac{\pi}{5} - \frac{\pi}{4}) = \frac{1}{\sqrt{2}}(1 - \frac{\pi}{20}).$$

A decimal approximation for this quantity is

$$\frac{1}{\sqrt{2}}(1 - \frac{\pi}{20}) \approx 0.596034707$$

while a (calculator) evaluation of $\sin\pi/5$ gives

$$\sin\frac{\pi}{5} \approx 0.587785252.$$

Noting that here

$$|x - a| = \frac{\pi}{20} \approx 0.157\ldots$$

the agreement to the second digit (the difference is roughly 0.01) is quite good.

[1] That is, we can determine $1/\sqrt{2}$ to any desired accuracy.

We might ask whether any other polynomial of degree one gives a better local approximation to f at $x = a$ than the tangent line. This is answered in very strong terms as a consequence of the following result (which also has theoretical implications).

Proposition 6.1.2. *Suppose f is a function continuous at $x = a$ and ϕ is a first-degree polynomial*

$$\phi(x) = \alpha_0 + \alpha_1 x$$

such that

$$|f(x) - \phi(x)| = \mathfrak{o}(|x - a|) \text{ as } x \to a.$$

Then f is differentiable at $x = a$ and ϕ is its tangent function—that is, they have first-order contact at $x = a$:

$$\begin{aligned} f(a) &= \phi(a) \\ f'(a) &= \phi'(a). \end{aligned}$$

Proof. Consider the difference function

$$\varepsilon(x) = f(x) - \phi(x)\,,$$

which is continuous at $x = a$. Our hypothesis can be stated as

$$\lim_{x \to a} \frac{\varepsilon(x)}{x - a} = 0.$$

Now, the denominator of this expression goes to zero, hence

$$\varepsilon(a) = \lim_{x \to a} \varepsilon(x) = 0.$$

Thus we can rewrite our hypothesis in the form

$$0 = \lim_{x \to a} \frac{\varepsilon(x)}{x - a} = \lim_{x \to a} \frac{\varepsilon(x) - \varepsilon(a)}{x - a}$$

which tells us that $\varepsilon(x)$ is differentiable at $x = a$, with $\varepsilon'(a) = 0$. But

$$f(x) = \phi(x) + \varepsilon(x)$$

so

$$\begin{aligned} f(a) &= \phi(a) + \varepsilon(a) &&= \phi(a) \\ f'(a) &= \phi'(a) + \varepsilon'(a) &&= \phi'(a). \end{aligned}$$

□

The theoretical importance of this result is that it gives us another characterization of differentiability: if f has a local approximation at $x = a$ with error of order $\mathfrak{o}(|x-a|)$ by a function ϕ whose graph is a straight line, then f is differentiable at $x = a$, and ϕ gives the equation of its tangent line.

The use of the tangent function as a local approximation to f is often referred to (for obvious reasons) as the **linear approximation**[2] to f at $x = a$.

Flush with the success of linear approximation, we might ask whether we could do even better by allowing polynomials of higher degree. A natural place to start is with higher-order contact.

Definition 6.1.3. *Given n a positive integer we say two functions f and ϕ have* ***contact to order n*** *at $x = a$ if they and their derivatives to order n exist and agree at $x = a$:*

$$\begin{aligned} f(a) &= \phi(a) \\ f'(a) &= \phi'(a) \\ &\vdots \\ f^{(n)}(a) &= \phi^{(n)}(a). \end{aligned}$$

Suppose we can find a polynomial ϕ which has contact to order $n > 1$ with f at $x = a$. Then for each $k = 1, \ldots, n-1$, the k^{th} derivatives $f^{(k)}$ and $\phi^{(k)}$ have contact to order $n-k$ at $x = a$. In particular, we have first-order contact for $f^{(n-1)}(x)$ and $\phi^{(n-1)}(x)$ at $x = a$ and hence, by Lemma 6.1.1,

$$\frac{f^{(n-1)}(x) - \phi^{(n-1)}(x)}{x-a} \to 0 \text{ as } x \to a.$$

Now, the numerator of this expression is the derivative of $f^{(n-2)}(x) - \phi^{(n-2)}(x)$, which also goes to zero as $x \to a$. The denominator, likewise, is the derivative of $(x-a)^2/2$; factoring out the $\frac{1}{2}$, we see that by L'Hôpital's rule (Proposition 4.10.1)

$$\frac{f^{(n-2)}(x) - \phi^{(n-2)}(x)}{(x-a)^2} \to 0 \text{ as } x \to a.$$

Proceeding in a similar fashion, we can show for each k going from 1 to n that

$$\frac{f^{(n-k)}(x) - \phi^{(n-k)}(x)}{(x-a)^k} \to 0 \text{ as } x \to a.$$

[2]Technically, $T_a f$ is an *affine* function, not a linear one.

In particular, $k = n$ gives us

$$f(x) - \phi(x) = \mathfrak{o}(|x-a|^n).$$

This is a much higher accuracy for ϕ as a local approximation to f. The converse also holds: if

$$\frac{f(x) - \phi(x)}{(x-a)^n} \to 0 \text{ as } x \to a$$

then Proposition 6.1.2 says that

$$\frac{f'(x) - \phi'(x)}{(x-a)^{n-1}} \to 0 \text{ as } x \to a$$

(we have factored out the constant $1/n$). Again, we can work our way up, step by step, to show that for $k = 1, \ldots, n$,

$$\frac{f^{(k)}(x) - \phi^{(k)}(x)}{(x-a)^k} \to 0 \text{ as } x \to a$$

and in particular[3]

$$f^{(k)}(a) = \phi^{(k)}(a) \text{ for } k = 0, \ldots, n$$

—in other words, f and ϕ must have contact to order n. We have established the following:

Proposition 6.1.4. *Suppose f has derivatives to order n at $x = a$, and ϕ is a polynomial. Then for $k = 1, \ldots, n$,*

$$f(x) - \phi(x) = \mathfrak{o}(|x-a|^k)$$

precisely if f and ϕ have contact to order k at $x = a$.

For the careful reader we note that if, instead of assuming ϕ is a polynomial, we had simply taken it to be some function, we would need to assume that ϕ, like f, has derivatives to order n at $x = a$ (see Exercise 5). Let us now see how to find a polynomial with specified derivatives to order n at $x = a$. It is easiest to first look at the case $a = 0$. Suppose

$$p(x) = c_0 + c_1 x + \ldots + c_n x^n$$

[3] We use the convention that the "zeroth derivative" is the function itself: $f^{(0)}(x) := f(x)$

is a polynomial of degree n. We know that

$$p(0) = c_0 + c_1(0) + \ldots + c_n(0)^n = c_0$$

—that is, the value of any polynomial at 0 is its "constant term". Now, differentiation has the following effect on a polynomial:

- the constant term of $p(x)$ vanishes;
- the coefficient c_1 of x is the (new) constant term of $p'(x)$;
- every monomial $c_k x^k$ ($k = 1, \ldots, n$) in $p(x)$ becomes $kc_k x^{k-1}$ in $p'(x)$: that is, it "moves left" one place and gets multiplied by the (old) exponent of x.

Thus, we have

$$p'(x) = c_1 + 2c_2x^1 + 3c_3x^2 + \ldots + nc_nx^{n-1}.$$

But the value of *this* polynomial at $x = 0$ is *its* constant term:

$$p'(0) = c_1 + 2c_2(0) + 3c_3(0)^2 + \ldots + nc_n(0)^{n-1} = c_1.$$

Note that doing this again leads to a constant term in $p''(x)$ equal to $2c_2$ (*not* c_2). It is reasonably easy to work out that for $k = 1, \ldots, n$,

$$p^{(k)}(0) = k!c_k$$

where $k!$ (pronounced "k factorial") is the product of the integers from 1 to k:

$$1! = 1, \quad 2! = 2 \cdot 1 = 2, \quad 3! = 3 \cdot 2 \cdot 1 = 6, \ldots$$

We add the useful convention that

$$0! = 1.$$

So we have

Remark 6.1.5. *For any polynomial*

$$p(x) = c_0 + c_1x + \ldots + c_nx^n,$$

the value of the k^{th} derivative at $x = 0$ equals $k!$ times the coefficient of x^k:

$$p^{(k)}(0) = k!c_k \text{ for } k = 1, \ldots.$$

In particular, $p^{(k)}(0) = 0$ once k exceeds the degree of p—in fact, starting from $k = n + 1$, the k^{th} derivative of a polynomial of degree n is the (constant) zero function.
If we wish to evaluate the derivatives of p at another point $x = a$ ($a \neq 0$), it is easiest to proceed as follows: set

$$q(x) = p(x + a)$$

which is the same as

$$p(x) = q(x - a)$$

and note that

$$p^{(k)}(a) = q^{(k)}(0),$$

for every integer $k = 0, 1, ...$ Since q is a polynomial of the same degree as p,

$$q(x) = c_0 + c_1 x + \ldots + c_n x^n,$$

we can rewrite $p(x) = q(x - a)$ in powers of $x - a$:

$$p(x) = c_0 + c_1(x - a) + c_2(x - a)^2 \ldots + c_n(x - a)^n.$$

Then it is clear that the derivatives of p at $x = a$ are precisely the same as those of q at $x = 0$. This means that to evaluate the derivatives of a polynomial p at $x = a$, we express it in powers of $x - a$ (instead of powers of x) as above, and then

$$p^{(k)}(a) = k! c_k \text{ for } k = 1, ...$$

Turning this around we have:

Remark 6.1.6. *Suppose f has (at least) n derivatives at $x = a$ and p is a polynomial; write $p(x)$ in terms of powers of $x - a$:*

$$p(x) = c_0 + c_1(x - a) + c_2(x - a)^2 \ldots + c_d(x - a)^d$$

(where d is the degree of p). Then f and p have contact to order n at $x = a$ precisely if the coefficients of all powers of $x - a$ up to the n^{th} in the expression for $p(x)$ above satisfy

$$c_k = \frac{f^{(k)}(a)}{k!} \text{ for } k = 0, 1, \ldots, n.$$

In particular, if $f^{(n)}(a) \neq 0$, then p must have degree at least equal to n. Note that the coefficients c_k of $(x-a)^k$ for $k > n$ in p have no effect on the n^{th} order contact with f at $x = a$.
The polynomial

$$T_a^n f(x) = c_0 + c_1(x-a) + \ldots + c_n(x-a)^n$$

where

$$c_k = \frac{f^{(k)}(a)}{k!} \text{ for } k = 0, 1, \ldots, n$$

is the only polynomial of degree n or less that has contact with f to order n at $x = a$; it is called the **Taylor polynomial of degree n for f at $x = a$**. We see immediately that it is the lowest-degree polynomial approximation to f near $x = a$ for which the error is of order $\mathfrak{o}(|x-a|^n)$:

$$f(x) - T_a^n f(x) = \mathfrak{o}(|x-a|^n) \text{ as } x \to a.$$

Let us illustrate the improvement in accuracy by using $T_a^2 f(x)$ to approximate the three numbers we looked at earlier: $\sqrt{80}$, $\sqrt{80.9}$, and $\sin \pi/5$.
For $f(x) = x^{1/2}$, $a = 81$, we have

$$\begin{aligned} f(a) &= \sqrt{81} &&= 9 \\ f'(a) &= \frac{1}{2\sqrt{81}} &&= \frac{1}{18} \\ f''(a) &= -\frac{1}{4\sqrt{(81)^3}} &&= -\frac{1}{2916} \end{aligned}$$

so that

$$T_{81}^2 f(x) = c_0 + c_1(x-81) + c_2(x-81)^2$$

where

$$\begin{aligned} c_0 &= f(a) &&= 9 \\ c_1 &= f'(a) &&= \frac{1}{18} \\ c_2 &= \frac{f''(a)}{2} &&= -\frac{1}{5832}. \end{aligned}$$

Then

$$\begin{aligned} T^2_{81} f\,(80) &= 9 + \frac{1}{18}(80-81) - \frac{1}{5832}(80-81)^2 \\ &= 9 - \frac{1}{18} - \frac{1}{5832} \\ &= \frac{52163}{5832} \\ &\approx 8.944272977 \end{aligned}$$

while

$$\begin{aligned} T^2_{81} f\,(80.9) &= 9 - \frac{1}{18}(0.1) - \frac{1}{5832}(0.01) \\ &= \frac{5245559}{583200} \\ &\approx 8.99444273. \end{aligned}$$

In both cases, the numbers themselves agree with (our) calculator values for $\sqrt{80}$ and $\sqrt{80.9}$, respectively, while their squares are, respectively,

$$80.00001908 \quad \text{and} \quad 80.90000002,$$

a striking improvement over the performance of the first-order approximations.

For $f(x) = \sin x$, $a = \pi/4$, we have

$$\begin{aligned} f(a) &= \sin\frac{\pi}{4} &&= \frac{1}{\sqrt{2}} \\ f'(a) &= \cos\frac{\pi}{4} &&= \frac{1}{\sqrt{2}} \\ f''(a) &= -\sin\frac{\pi}{4} &&= -\frac{1}{\sqrt{2}} \end{aligned}$$

so

$$T^2_{\pi/4} f\,(x) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}(x - \frac{\pi}{4}) - \frac{1}{\sqrt{2}}\frac{(x-\frac{\pi}{4})^2}{2}.$$

Then

$$\begin{aligned} T^2_{\pi/4} f\,(x) &= \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}(\frac{\pi}{5} - \frac{\pi}{4}) - \frac{1}{\sqrt{2}}\frac{(\frac{\pi}{5}-\frac{\pi}{4})^2}{2} \\ &= \frac{1}{\sqrt{2}}\left(1 - \frac{\pi}{20} - \frac{\pi^2}{800}\right) \\ &\approx 0.587311127 \end{aligned}$$

which differs from the calculator value of $\sin \pi/5$ by just under 0.0005, again a marked improvement over the previous approximation.
We know from Proposition 6.1.4 that the error in using $T_a^n f$ as an approximation to f is of order $\mathfrak{o}(|x-a|^n)$ as $x \to a$. While this gives us an idea of how the accuracy improves as x gets closer to a, it doesn't yield any explicit estimate of the actual size of the error. The following result often allows us to obtain such an estimate explicitly. It was given by Joseph Louis Lagrange (1736-1813) in 1797, in [37] ([30, p. 254]).

Theorem 6.1.7 (Taylor's Theorem, Lagrange Form)**.** *Suppose f has $n+1$ continuous derivatives between $x=a$ and $x=b$. Then there exists s between a and b such that the error for the degree n Taylor polynomial as a local approximation at $x=a$ to the value of $f(b)$ satisfies*

$$f(b) - T_a^n f(b) = \frac{f^{(n+1)}(s)}{(n+1)!}(b-a)^{n+1}.$$

The expression on the right is easy to remember: it looks exactly like the *next* term we would need to write down $T_a^{n+1} f(b)$, except that the derivative is evaluated at $x=s$ instead of $x=a$.

Proof. Write $\varepsilon(x)$ for the error between $f(x)$ and $T_a^n f(x)$:

$$\varepsilon(x) = f(x) - T_a^n f(x).$$

This satisfies

$$\varepsilon^{(k)}(a) = 0 \text{ for } k = 0, 1, \ldots, n.$$

Now, consider a new function

$$g(x) = \varepsilon(x) + c_{n+1}(x-a)^{n+1}$$

where c_{n+1} is chosen so that

$$g(b) = 0$$

namely,

$$c_{n+1} = -\frac{\varepsilon(b)}{(b-a)^{n+1}}.$$

Note that the extra term has contact to order $n+1$ with 0 at $x=a$, so

$$g^{(k)}(a) = 0 \text{ for } k = 0, \ldots, n.$$

We now apply Rolle's theorem to $g(x)$ on the interval $[a,b]$: since

$$g(a) = g(b) = 0,$$

it follows that for some ξ_1 between a and b,

$$0 = g'(\xi_1) = \varepsilon'(\xi_1) + (n+1)c_{n+1}(\xi_1 - a)^n.$$

But $g'(a) = 0$ also, so for some ξ_2 between a and ξ_1,

$$g''(\xi_2) = 0.$$

Recursively, for $k = 1, \ldots, n$, we have

$$g^{(k)}(a) = 0 = g^{(k)}(\xi_k)$$

so (by Rolles's theorem) there exists ξ_{k+1} between a and ξ_k such that

$$g^{(k+1)}(\xi_{k+1}) = 0.$$

Now, let $s = \xi_{n+1}$. Since $T_a^n f\,(x)$ is a polynomial of degree n (or less), its $(n+1)^{st}$ derivative is zero. Also, the $(n+1)^{st}$ derivative of the extra term (which is a monomial of degree $n+1$) is $(n+1)!c_{n+1}$. Thus, the condition defining $\xi_{n+1} = s$ gives

$$\begin{aligned} 0 &= g^{(n+1)}(s) \\ &= \left.\frac{d^{n+1}}{dx^{n+1}}\right|_{x=s} \left(f(x) - T_a^n f\,(x) + c_{n+1}(x-a)^{n+1}\right) \\ &= f^{(n+1)}(s) - 0 + (n+1)!c_{n+1} \end{aligned}$$

and solving for c_{n+1} gives

$$c_{n+1} = -\frac{f^{(n+1)}(s)}{(n+1)!}.$$

But we defined c_{n+1} so that

$$\begin{aligned} 0 = g(b) &= \varepsilon(b) + c_{n+1}(b-a)^{n+1} \\ &= \varepsilon(b) - \frac{f^{(n+1)}(s)}{(n+1)!}(b-a)^{n+1}. \end{aligned}$$

In other words,

$$f(b) - T_a^n f\,(b) = \varepsilon(b) = \frac{f^{(n+1)}(s)}{(n+1)!}(b-a)^{n+1},$$

as required. □

Thus, instead of the *approximate* expression for $f(x)$

$$\begin{aligned} f(x) &\approx T_a^n f(x) \\ &= f(a) + f'(a)(x-a) + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n \end{aligned}$$

we can write the "exact" expression

$$\begin{aligned} f(x) &= T_a^n f(x) + \frac{f^{(n+1)}(s)}{(n+1)!}(x-a)^{n+1} \\ &= f(a) + f'(a)(x-a) + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(s)}{(n+1)!}(x-a)^{n+1} \end{aligned}$$

where of course the "s" in the last term is not known exactly: it is some number between a and x. But this is useful information. We illustrate with the two functions considered earlier.
For $f(x) = x^{1/2}$, $a = 81$ and $n = 2$, we have

$$f^{(3)}(s) = -\frac{3}{8s^{5/2}}.$$

For s between 80 and 81, $s^{5/2}$ lies between $80^{5/2}$ and $81^{5/2}$; now $81^{5/2} = 9^5 = 59049$. We don't know $80^{5/2}$, but certainly it is more than $64^{5/2} = 8^5 = 32768$. This means that $f^{(3)}(s)$ lies between

$$-\frac{3}{8 \cdot 8^5} \approx -1.14 \times 10^{-5}$$

and

$$-\frac{3}{8 \cdot 9^5} \approx --6.35 \times 10^{-6}.$$

Thus, Taylor's theorem *guarantees* that the error in our calculation of $\sqrt{80}$ via $T_{81}^2 f(80)$ lies between

$$-\frac{3}{8 \cdot 8^5}\frac{(80-81)^3}{3!} = \frac{1}{2 \cdot 8^6} \approx 2 \times 10^{-6}$$

and

$$-\frac{3}{8 \cdot 9^5}\frac{(80-81)^3}{3!} = \frac{1}{16 \cdot 9^5} \approx 10^{-6}.$$

This tells us (with certainty!) that our calculation is an *under*estimate, by at least 10^{-6}, but by no more than 2×10^{-6}.

For the calculation of $\sqrt{80.9}$, the term $(80-81)^3=-1$ is replaced by $(80.9-81)^3=-0.001$, and all the estimates are divided by 10^3.
For $f(x)=\sin x$ and $a=\pi/4$, we have

$$f^{(3)}(s)=-\cos s.$$

Since s is between $\pi/4$ and $\pi/5$, $f^{(3)}(s)$ is between $-\cos\pi/4$ and $-\cos\pi/5$. We know that

$$\cos\frac{\pi}{4}=\frac{1}{\sqrt{2}}<\frac{1}{1.4}=\frac{5}{7}$$

and $\cos\pi/5$ is less than

$$\cos\frac{\pi}{6}=\frac{\sqrt{3}}{2}<0.9.$$

Thus the error between $\sin\pi/5=f(\pi/5)$ and $T^2_{\pi/4}f\,(\pi/5)$ is between

$$-\frac{5}{7}\frac{(\frac{\pi}{5}-\frac{\pi}{4})^3}{3!}=\left(\frac{5}{7}\right)\left(\frac{\pi^3}{6\cdot 20^3}\right)=\frac{\pi^3}{7\cdot 24\cdot 400}\approx 5\times 10^{-4}$$

and

$$-0.9\frac{(\frac{\pi}{5}-\frac{\pi}{4})^3}{3!}=\left(\frac{9}{10}\right)\left(\frac{\pi^3}{6\cdot 20^3}\right)=\frac{3\pi^3}{160000}\approx 6\times 10^{-4}.$$

Note that we observed an error value in this range.
Raising the degree n in the Taylor polynomial approximation $T^n_a f$ to f at $x=a$ amounts to adding new, higher-degree terms to the existing polynomials. The addition of these terms yields improved local accuracy, so we might ask whether these approximations converge to a function

$$\mathcal{T}_a f(x)=\lim_{n\to\infty}T^n_a f\,(x)$$

which approximates f especially well. Formally, the value of this function at any point x is the sum of a series

$$\begin{aligned}\mathcal{T}_a f(x)&=f(a)+f'(a)(x-a)+f''(a)\frac{(x-a)^2}{2!}+\ldots+f^{(k)}(a)\frac{(x-a)^k}{k!}+\ldots\\&=\sum_{k=0}^{\infty}f^{(k)}(a)\frac{(x-a)^k}{k!}\end{aligned}$$

whose partial sums are the Taylor polynomials

$$T^n_a f\,(x)=\sum_{k=0}^{n}f^{(k)}(a)\frac{(x-a)^k}{k!}.$$

It is called the **Taylor series** of f around $x = a$.
The first publication of Taylor series was in the *Methodus incrementorum* (1715), a textbook published by Brooke Taylor (1685-1731) [51, pp. 328-33]. He came up with the series as part of an attempt at formulating a general theory of finite differences, following earlier ideas of Newton and Gregory (see Exercise 7). In fact, the idea was "in the air", and had been touched on by both Newton and Gregory in earlier, unpublished work.
The special case of $a = 0$ is known as **MacLaurin series**; although a special case, it was developed as a tool in the *Treatise of Fluxions* (1742) [51, pp. 338-341], published by one of Newton's best and most successful disciples, Colin Maclaurin (1698-1746). This was in part a serious attempt to answer the stinging criticisms of the logical looseness of the calculus by George Berkeley (1685-1753) in *The Analyst*. Among other things, Maclaurin gave higher-order generalizations of the Second Derivative Test for local extrema as a corollary of his series.
The evaluation of the Taylor series at any point x involves questions of convergence. We will take up these and related questions in the next two sections. As we shall see after this excursion, the Taylor series in many familiar cases actually gives back the function f in a new form.

Exercises for § 6.1

Answers to Exercises 1ace, 3 are given in Appendix B.

Practice problems:

1. In each problem below, you are given a function f and a point a. Calculate the derivatives of f to order 3 at $x = a$ and write down the Taylor polynomial of f of order 3 at $x = a$.

   (a) $f(x) = x^{-1}$, $a = 2$ (b) $f(x) = \sin x$, $a = \frac{\pi}{4}$

   (c) $f(x) = \sin x$, $a = \frac{\pi}{2}$ (d) $f(x) = x^{-1/2}$, $a = 4$

   (e) $f(x) = \ln x$, $a = 1$ (f) $f(x) = e^{-x}$, $a = 0$

2. Use Theorem 6.1.7 (Taylor's Theorem, Lagrange Form) to show that

$$\left|1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \frac{1}{7!} + \frac{1}{8!} + \frac{1}{9!} - e\right| < 10^{-6}.$$

3. Consider the function $f(x) = \cos x$.

(a) How good an approximation to $f(1) = \cos 1$ is the Taylor polynomial of order 5 at $x = 0$, $T_0^5 f(1) = 1 - \frac{1}{2} + \frac{1}{24}$?

(b) How many terms of the Taylor series for $\cos x$ at $x = 0$ do we need to approximate $\cos 1$ with an error of at most 0.0001?

Theory problems:

4. **Alternative forms of Taylor's Theorem:** Theorem 6.1.7 gives one form of the difference between f and its n^{th} degree Taylor polynomial $T_a^n f$. Here we explore some variations on this result [50, pp. 391-4].

(a) This time, instead of taking the difference between f and the Taylor polynomial based at $x = a$, we fix b and consider the difference $\delta(x)$ between $f(b)$ and the Taylor polynomial based at x and evaluated at $x = b$:

$$\begin{aligned} f(b) &= T_b^n f(x) + \delta(x) \\ &= f(x) + f'(x)(b - x) + \cdots + \frac{f^{(n)}(x)}{n!}(b - x)^n + \delta(x) . \end{aligned} \quad (6.1)$$

Now differentiate both sides of Equation (6.1) with respect to x. When we differentiate $T_b^n f$, there is a lot of cancellation, and in the end we are left with

$$0 = \frac{f^{(n+1)}(x)}{n!}(b - x)^n + \delta'(x)$$

or

$$\delta'(x) = -\frac{f^{(n+1)}(x)}{n!}(b - x)^n. \quad (6.2)$$

(b) Now, by the Fundamental Theorem of Calculus,

$$\delta(b) - \delta(a) = \int_a^b \delta'(t)\, dt$$

but by construction, since $T_b^n f(b) = f(b)$, we have $\delta(b) = 0$. On the other hand,

$$\delta(a) = f(b) - T_a^n f(b)$$

so we have

$$f(b) = T_a^n f\,(b) + [\delta(a) - \delta(b)] = T_a^n f\,(b) - \int_a^b \delta'(t)\;dt$$
$$= T_a^n f\,(b) + \int_a^b \frac{f^{(n+1)}(t)}{n!}(b-t)^n\,dt.$$

The integral on the right is called the **integral form of the remainder** for the Taylor series.

(c) If we apply the Mean Value Theorem (Proposition 4.9.2) to $\delta(x)$ on $[a,b]$, we get the existence of some $s \in (a,b)$ for which

$$\delta'(s) = \frac{\delta(b) - \delta(a)}{b-a}.$$

Using the fact that $\delta(b) = 0$ and Equation (6.2), we can conclude that *there exists some $s \in (a,b)$ for which*

$$f(b) - T_a^n f\,(b) = \frac{f^{(n+1)}(s)}{n!}(b-s)^n(b-a).$$

This statement is known as the **Cauchy form of the remainder** for the Taylor series.

(d) Examine the various proofs (above, and the proof of Theorem 6.1.7) to see precisely what we needed to assume about the function $f^{(n+1)}$: did we need to assume it was continuous in Theorem 6.1.7, or in the alternative versions above?

5. (a) Suppose f and ϕ are both differentiable at $x = a$. Show that

$$f(x) - \phi(x) = \mathfrak{o}(|x-a|)$$

precisely if

$$f'(a) = \phi'(a)\,.$$

(b) Suppose ϕ is differentiable at $x = a$ and f (which is not even assumed continuous) satisfies

$$f(x) - \phi(x) = \mathfrak{o}(|x-a|).$$

Show that f is differentiable at $x = a$.

(c) Use induction[4] on n to show that if ϕ has derivatives to order n at $x = a$ while f has derivatives to order $n - 1$ and $f(x) - \phi(x) = \mathfrak{o}(|x - a|^n)$, then f has derivatives to order n at $x = a$, and has contact to order n with ϕ.

Challenge problem:

6. (a) Show that if f is a polynomial of degree d, then

$$T_a^n f\,(x) = f(x) \text{ for all } n \geq d.$$

(b) **Cauchy's example:** Consider the function

$$f(x) = \begin{cases} e^{-x^2} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

Use Exercise 8 in § 4.10 to show that f has contact to all orders at $x = 0$ with the constant zero function $\phi(x) = 0$ (and hence has the same Taylor series there) but is not itself constant. A slight variation of this example was pointed out by Cauchy [13] in his *Résumé des lecons sur le calcul infinitesimal* (1823), rejecting Lagrange's attempt [37] to use power series as the foundation for calculus "in spite of all the respect that such a great authority must command." [26, p. 54]

History note:

7. **Taylor's derivation of his series:** Taylor, in his *Methodus incrementorum* (1715), derived his series on the basis of an interpolation formula given by Newton in [41, Book III, Lemma V] but also discovered earlier by James Gregory (1638-1675). Our exposition follows [20, pp. 284-9].

(a) Suppose we are given $n + 1$ values $y_0, y_1, ..., y_n$ and $n + 1$ equally spaced points $x_0 < x_1 < \cdots < x_n$, say $x_{i+1} - x_i = \triangle x$. We wish to find a polynomial

$$p(x) = A_0 + A_1 x + \cdots + A_n x^n$$

[4] See Appendix A for more information.

such that

$$p(x_i) = y_i \quad i = 0, \ldots, n.$$

The solution of this is given by the **Newton-Gregory Interpolation Formula** which we will state without proof. We define the k^{th} order *difference quotients* $\triangle^k y_j$ for $j = 0, \ldots, n-k$, $k = 0, \ldots, n$ inductively, via

$$\triangle^1 y_j := y_{j+1} - y_j$$
$$\triangle^{k+1} y_j := \triangle^k y_{j+1} - \triangle^k y_j.$$

Then the required polynomial is given by

$$p(x_0 + s\triangle x) = y_0 + \triangle^1 y_0 s + \frac{\triangle^2 y_0}{2!} s(s-1) + \cdots + \frac{\triangle^n y_0}{n!} s(s-1)\cdots(s-n).$$

(b) Now, Taylor wants to do the following. Set $x = x_0 + n\triangle x$, use the interpolation formula for $y_j = f(x_0 + j\triangle x)$, and then take a limit as $n \to \infty$. The interpolation formula for a given n gives

$$y = y_0 + \left(\triangle^1 y_0\right)(n) + \left(\triangle^2 y_0\right)\left(\frac{n(n-1)}{2!}\right) + \cdots + \left(\triangle^k y_0\right)\left(\frac{n(n-1)\ldots n-k+1}{k!}\right) + \cdots \cdots + \left(\triangle^{n-1} y_0\right)\left(\frac{n(n-1)\ldots 1}{(n-1)!}\right).$$

Substitute

$$n - k = \frac{x - x_k}{\triangle x}, \quad k = 0, \ldots, n-1$$

to get

$$y = y_0 + \left(\frac{\triangle^1 y_0}{\triangle x}\right)(x - x_0) + \left(\frac{\triangle^2 y_0}{(\triangle x)^2}\right)\left(\frac{(x-x_0)(x-x_1)}{2!}\right) + \cdots + \left(\frac{\triangle^k y_0}{(\triangle x)^k}\right)\left(\frac{(x-x_0)(x-x_1)\cdots(x-x_{k-1})}{k!}\right) + \cdots \cdots + \frac{(x-x_0)(x-x_1)\cdots(x-x_{n-1})}{n!}.$$

Thinking infinitesimally (more accurately, in terms of fluxions), Taylor makes the approximations

$$\frac{\triangle^k y_0}{(\triangle x)^k} \cong \frac{\dot{y_0}}{\dot{x_0}} \cong f^{(n)}(x_0)$$

and noting that as $n \to \infty$ all the $x_k \to x_0$, this gives the formula

$$f(x) = y = f(x_0) + f'(x_0)\,(x - x_0) + \frac{f^{(2)}(x_0)}{2!}(x - x_0)^2 + \cdots$$

which is the Taylor series for f at $x = x_0$.

6.2 Convergence of Series

In this section we carefully study the convergence of series, in preparation for the study of power series. Recall that a series

$$\sum_{k=0}^{\infty} a_k \tag{6.3}$$

is interpreted as a sequence of partial sums

$$S_N := \sum_{k=0}^{N} a_k, \quad N = 0, ...$$

and the series converges precisely if this sequence converges, to

$$\sum_{k=0}^{\infty} a_k = \lim S_N.$$

In practice, it can be very difficult to even determine whether a given series converges, never mind to find its limit. However, we have already seen some results that help with this. The first of these is a negative result: the series *diverges* unless the terms go to zero:

Divergence Test for Series (Proposition 2.3.10) If the series (6.3) converges, then

$$\lim a_k = 0.$$

Of course, this condition does not guarantee convergence, as shown by the harmonic series (p. 61). The *positive* results (*i.e.*, those that allow us to conclude that a series *converges*) which we have already seen are:

Comparison Test (Corollary 2.3.6) If $0 \leq a_k \leq b_k$ for all k, and $\sum_{k=0}^{\infty} b_k$ converges, then so does $\sum_{k=0}^{\infty} a_k$.

Geometric Series Test (Proposition 2.4.6) If $a_k \neq 0$ for all k and $\frac{a_{k+1}}{a_k} = r$ independent of k (*i.e.*, $a_k = a_0 r^k$ for $k = 1, 2, ...$), then $\sum_{k=0}^{\infty} a_k$

- *diverges* if $|r| \geq 1$
- *converges* if $|r| < 1$, to

$$\sum_{k=0}^{\infty} a_0 r^k = \frac{a_0}{1-r}.$$

Alternating Series Test (Proposition 2.4.8) successive terms have opposite sign, then $\sum_{k=0}^{\infty} a_k$ converges.

Integral Test (Proposition 5.7.4): If f is a positive, strictly decreasing function defined on $[1, \infty)$, then $\int_1^{\infty} f(x)\,dx$ and $\sum_{k=1}^{\infty} f(k)$ either both converge or both diverge.

***p*-series Test** (Corollary 5.7.5)

- *diverges* if $p \leq 1$
- *converges* if $p > 1$.

Note that a number of these tests are limited to series whose terms are all positive; such series are easier to think about, because the sequence of partial sums is automatically increasing in this case. The following result was mentioned in Exercise 10 in § 2.5; here we give a complete proof because of the importance of this result to what follows.

Lemma 6.2.1 (Absolute Convergence Test)**.** *Given any series*

$$\sum_{k=0}^{\infty} a_k,$$

if the series obtained by replacing each term with its absolute value (a positive *series) converges*

$$\sum_{k=0}^{\infty} |a_k| < \infty,$$

then the original series also converges.

Proof. We have assumed that

$$\sum_{k=0}^{\infty} |a_k| = L < \infty.$$

Note that the "tail" of this series (the series obtained by dropping the first few terms) also converges:

$$\begin{aligned}\sum_{k=n}^{\infty} |a_k| &= \lim_{N\to\infty} \sum_{k=n}^{N} |a_k| \\ &= \lim_{N\to\infty} \left(\sum_{k=0}^{N} |a_k| - \sum_{k=0}^{n-1} |a_k| \right) \\ &= L - \sum_{k=0}^{n-1} |a_k| .\end{aligned}$$

Denote this "tail" by τ_n:

$$\tau_n = \sum_{k=n}^{\infty} |a_k| = L - \sum_{k=0}^{n-1} |a_k| .$$

Our hypothesis (that the series of absolute values converges) says that the tails go to zero:

$$\lim \tau_n = 0.$$

Note, furthermore, that the tails are decreasing: in fact, for $n < m$,

$$\tau_n - \tau_m = \sum_{k=n}^{m-1} |a_k| \geq 0.$$

Now, we need to find a limit for the sequence of partial sums of the original sequence:

$$s_n = \sum_{k=0}^{n} a_k.$$

Comparing two such partial sums, say s_n and s_m with $n < m$, we have

$$s_m - s_n = a_{n+1} + \ldots + a_m = \sum_{k=n+1}^{m} a_k.$$

Taking absolute values, using the triangle inequality and the equation above, we have

$$|s_m - s_n| = \left|\sum_{k=n+1}^{m} a_k\right| \leq \sum_{k=n+1}^{m} |a_k| = \tau_{n+1} - \tau_{m+1}.$$

This means, whenever $n < m$,

$$\tau_{m+1} - \tau_{n+1} \leq s_n - s_m \leq \tau_{n+1} - \tau_{m+1}.$$

Now, consider the two sequences

$$\begin{aligned} \alpha_l &= s_l - \tau_{l+1} \\ \beta_l &= s_l + \tau_{l+1}. \end{aligned}$$

From the two inequalities above we obtain

$$\alpha_n \leq \alpha_m \leq \beta_m \leq \beta_n \text{ for } m > n.$$

This says each of these sequences is monotone and bounded, so convergent. Furthermore, since

$$\beta_n - \alpha_n = 2\tau_{n+1} \to 0$$

they have the same limit.
But clearly, since $\tau_{n+1} \geq 0$,

$$\alpha_n \leq s_n \leq \beta_n$$

so the s_n's converge by the Squeeze Theorem (Theorem 2.4.7). □

We say that a series

$$\sum_{k=0}^{\infty} a_k$$

converges absolutely if the series of absolute values converges:

$$\sum_{k=0}^{\infty} |a_k| < \infty.$$

The result above says that every absolutely convergent series converges. Note that *the converse is false*: there exist series which converge, but don't converge absolutely. For example, the **alternating harmonic series** given by

$$a_k = \frac{(-1)^k}{k+1}$$

that is

$$\sum_{k=0}^{\infty} a_k = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

converges by the alternating series test (Proposition 2.4.8), but the corresponding series of absolute values is the harmonic series

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \dots$$

which we found diverges (Exercise 34 in § 2.3).
As an application of this test, consider the series

$$\sum_{k=1}^{\infty} \frac{\operatorname{sgn}(\sin k)}{k^2}$$

where the **sign function** sgn is defined by

$$\operatorname{sgn}(x) = \begin{cases} +1 & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0 \\ -1 & \text{if} \quad x < 0. \end{cases}$$

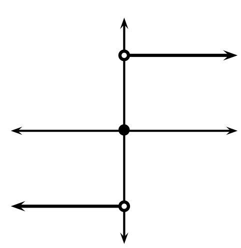


Figure 6.2: Definition of sgn(x).

The actual sequence of signs in our series is irregular: for $k = 1, 2, 3$ we have $0 < k < \pi$, so the first three terms are positive; the next three are negative and for a while, the signs alternate in runs of three, but between

$14\pi \approx 43.9$ and $15\pi \approx 47.1$ there are *four* integer values. The alternating series test cannot be applied here, but of course

$$|a_k| = \frac{1}{k^2}$$

tells us immediately that the series converges absolutely.
The comparison test lets us determine convergence for many series. Here is a class of examples.[5]

Remark 6.2.2. *Suppose*

$$p(x) = c_0 + c_1 x + \ldots + c_n x^n$$

is a polynomial of degree n *with non-negative coefficients:*

$$c_k \geq 0 \quad k = 0, 1, \ldots n-1, \quad c_n > 0.$$

Then the series

$$\sum_{k=1}^{\infty} \frac{1}{p(k)} = \frac{1}{p(1)} + \frac{1}{p(2)} + \ldots$$

converges *if* $n \geq 2$ *and* diverges *if* $n = 0$ *or* 1.

Proof. For $n \geq 2$, the non-negativity of the coefficients gives us, for $k = 1, 2, ...$

$$p(k) = c_0 + c_1 k + \ldots + c_n k^n \geq c_n k^n.$$

Also, since $n \geq 2$, for $k = 1, 2, ...,$

$$c_n k^n \geq c_n k^2.$$

Thus,

$$\frac{1}{p(k)} \leq \frac{1}{c_n k^2}$$

and we can use comparison with the convergent series

$$\sum_{k=1}^{\infty} \frac{1}{c_n k^2} = \frac{1}{c_n} \sum_{k=1}^{\infty} \frac{1}{k^2}.$$

For $n = 0$, we have

$$\lim \frac{1}{p(k)} = \lim \frac{1}{c_0} = \frac{1}{c_0} \neq 0$$

[5] A more general result, which includes this one as a special case, is Remark 6.2.4, which follows from the Limit Comparison Test, Proposition 6.2.3.

so the series diverges by the divergence test.
Finally, for $n = 1$, we are looking at

$$\sum_{k=1}^{\infty} \frac{1}{p(k)} = \sum_{k=1}^{\infty} \frac{1}{c_0 + c_1 k}.$$

If $c_0 = 0$, this is just the harmonic series multiplied by $1/c_1$, and divergence is easy to establish. However, if $c_0 > 0$, then any *direct* comparison with the harmonic series gives the wrong inequality:

$$\frac{1}{c_1 k + c_0} < \frac{1}{c_1 k} \quad k = 1, 2, \ldots \text{ if } c_0 > 0$$

and the *divergence* of the *higher* series tells us nothing about the *lower* one. However, let us note that for any integer $m \geq 1$ the series

$$\sum_{k=1}^{\infty} \frac{1}{k + m}$$

consists of the harmonic series missing its first m terms. This is also divergent (Exercise 3). Now, factoring out c_1 and letting

$$\alpha = \frac{c_o}{c_1},$$

we have for $k = 1, \ldots$

$$\frac{1}{c_1 k + c_0} = \frac{1}{c_1} \frac{1}{k + \alpha}.$$

Note that, for any integer $m > \alpha$, we have

$$\frac{1}{c_1 k + c_0} > \frac{1}{c_1} \frac{1}{k + m}$$

and *this* comparison goes the *right* way: the *lower* series *diverges*, so our series diverges, also. □

A combination of absolute convergence and comparison lets us handle series of the following form:

$$\sum_{k=1}^{\infty} \frac{1}{k^p} \sin k \quad p \geq 2.$$

First, we have

$$|a_k| = \frac{1}{k^p} |\sin k|$$

and, since $|\sin\theta| \leq 1$ for any θ,

$$|a_k| \leq \frac{1}{k^p}.$$

Hence, by the comparison test (and remark 6.2.2)

$$\sum_{k=1}^{\infty} |a_k| \leq \sum_{k=1}^{\infty} \frac{1}{k^p} < \infty.$$

But this means our original series converges absolutely, and so converges by lemma 6.2.1.
The kind of cleverness we needed to prove Remark 6.2.2 via the comparison test can be avoided by means of the following, which also gives many other convergence results.

Proposition 6.2.3 (Limit Comparison Test)**.** *Suppose*

$$\sum_{k=0}^{\infty} a_k \text{ and } \sum_{k=0}^{\infty} b_k$$

are positive series ($a_k > 0, b_k > 0$ for $k = 0, 1, 2, ...$) and consider the sequence of termwise ratios

$$c_k = \frac{a_k}{b_k} > 0.$$

Then

1. *If the sequence $\{c_k\}$ is bounded above and $\sum_{k=0}^{\infty} b_k$ converges, then $\sum_{k=0}^{\infty} a_k$ also converges.*

2. *If the sequence $\{c_k\}$ is bounded away from 0 and $\sum_{k=0}^{\infty} b_k$ diverges, then $\sum_{k=0}^{\infty} a_k$ also diverges.*

In particular, if $c_k \to \gamma$ with $0 < \gamma < \infty$, then the two series either both converge or both diverge.

Proof. If $c_k < C < \infty$ for all k, we have

$$a_k < Cb_k \quad k = 0, 1, ...$$

so $\sum b_k < \infty$ gives $\sum Cb_k < \infty$ and hence $\sum a_k < \infty$ by comparison.

If $c_k > 1/C > 0$ for all k, then

$$a_k > \frac{b_k}{C}$$

and so $\sum b_k = \infty$ gives $\sum a_k = \infty$ by comparison.
The final statement follows from the fact that a sequence of positive numbers c_k with positive (finite) limit has bounds of the form

$$\frac{1}{C} < c_k < C$$

for some real, positive C (Exercise 5), and so the appropriate earlier statement applies. □

This theorem gives a number of easy corollaries. One takes care of series defined by a rational function.

Remark 6.2.4. *Suppose*

$$r(x) = \frac{\alpha_0 + \alpha_1 x + \ldots + \alpha_p x^p}{\beta_0 + \beta_1 x + \ldots + \beta_q x^q}$$

is a rational function ($\alpha_p \neq 0 \neq \beta_q$) defined for all integers $k \geq m$, with numerator of degree p and denominator of degree q.
Then

$$\sum_{k=m}^{\infty} r(k) = r(m) + r(m+1) + \ldots$$

converges if and only if

$$q \geq p + 2.$$

Proof. First, assume $p < q$ since otherwise $\lim r(k) \neq 0$ and the divergence test applies. For $p < q$, consider

$$a_k = r(k)\,, \quad b_k = \frac{1}{k^{q-p}}$$

in the limit comparison test. We have

$$c_k = \frac{a_k}{b_k} = \frac{\alpha_0 k^{q-p} + \ldots + \alpha_p k^q}{\beta_0 + \ldots + \beta_q k^q} \to \frac{\alpha_p}{\beta_q} \neq 0$$

and hence $\sum a_k$ and $\sum b_k$ either both converge or both diverge. But $\sum b_k$ converges if and only if $q \geq p + 2$. □

For example,

$$\sum_{k=2}^{\infty} \frac{k+1}{2k^3-1}$$

converges, while

$$\sum_{k=0}^{\infty} \frac{k}{k^2+100k+1000}$$

diverges.
Another application is to the series

$$\sum_{k=1}^{\infty} \frac{1}{k!} = 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots = 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \ldots.$$

Limit comparison using $a_k = 1/k!$ and $b_k = 1/k^2$ gives

$$c_k = \frac{a_k}{b_k} = \frac{1/k!}{1/k^2} = \frac{k^2}{k!} \to 0.$$

In particular, the ratio is bounded above, and since $\sum b_k = \sum 1/k^2$ converges, so does $\sum a_k = \sum 1/k!$.
A very powerful tool is obtained by applying the limit comparison test with a geometric series.

Theorem 6.2.5 (Ratio Test)**.** *Suppose* $\sum_{k=0}^{\infty} a_k$ *has* $a_k \neq 0$ *for* $k = 0, 1, \ldots$*, and assume the limit*

$$\rho = \lim \left| \frac{a_{k+1}}{a_k} \right|$$

exists.

1. *If* $\rho > 1$*, the series diverges.*

2. *If* $\rho < 1$*, the series converges absolutely.*

3. *The information that* $\rho = 1$ *does not suffice to decide convergence or divergence of the series.*

Proof. If $\rho > 1$, then for some N, $k \geq N$ guarantees that

$$\left| \frac{a_{k+1}}{a_k} \right| > 1$$

and in particular

$$|a_k| \geq |a_N| \text{ for } k \geq N.$$

But then $\lim a_k \neq 0$, and the series diverges.
If $\rho < 1$, let $r = (\rho + 1)/2$, so $\rho < r < 1$. For some N, $k \geq N$ guarantees

$$\left| \frac{a_{k+1}}{a_k} \right| < r$$

so (for $k \geq N$)

$$|a_k| < r^{k-N} |a_N|.$$

Replacing $|a_N|$ with some constant Cr^N, we get

$$|a_k| < Cr^k$$

and since $\sum Cr^k < \infty$, we have that $\sum |a_k| < \infty$, and hence the series converges absolutely.
The observation that both of the series $\sum 1/k$ and $\sum 1/k^2$ have $\rho = 1$ establishes the last point.

□

As a few illustrations of the use of the ratio test, first consider

$$\sum_{k=1}^{\infty} \frac{k}{2^k}.$$

Here,

$$\begin{aligned} \rho = \lim \frac{a_{k+1}}{a_k} &= \lim \frac{(k+1)/2^{k+1}}{k/2^k} \\ &= \left(\lim \frac{k+1}{k} \right) \left(\lim \frac{2^k}{2^{k+1}} \right) \\ &= (1) \left(\frac{1}{2} \right) = \frac{1}{2} < 1 \end{aligned}$$

so the series converges.
Similarly, for

$$\sum_{k=1}^{\infty} \frac{k+1}{1+2^k}$$

we have

$$\rho = \lim \frac{a_{k+1}}{a_k} = \lim \frac{(k+2)/(1+2^{k+1})}{(k+1)/(1+2^k)}$$
$$= \left(\lim \frac{k+2}{k+1}\right)\left(\lim \frac{1+2^k}{1+2^{k+1}}\right)$$
$$= (1)\left(\frac{1}{2}\right) = \frac{1}{2} < 1$$

and again the series converges.
Finally, we look at the more subtle problem

$$\sum_{k=0}^{\infty} \frac{k^k}{k!}.$$

Here,

$$\rho = \lim \frac{a_{k+1}}{a_k} = \lim \frac{(k+1)^{k+1}/(k+1)!}{k^k/k!}$$
$$= \lim \frac{(k+1)^k}{k^k} = \lim \left(1 + \frac{1}{k}\right)^k.$$

We find this last limit using logarithms:

$$\ln \rho = \lim \ln \left(1 + \frac{1}{k}\right)^k = \lim \left[k \ln \left(1 + \frac{1}{k}\right)\right].$$

This leads to the indeterminate form

$$\infty \ln 1 = \infty \cdot 0$$

so we rewrite the problem to yield 0/0 and apply L'Hôpital's rule:

$$\ln \rho = \lim \frac{\ln\left(\frac{k+1}{k}\right)}{1/k} = \lim \frac{\frac{1}{k+1} - \frac{1}{k}}{-1/k^2} = \lim \frac{k^2}{k^2+k} = 1.$$

It follows immediately that

$$\rho = e^1 = e > 1$$

so the series *diverges.*
A modification of our argument for the ratio test also gives the following test, which is sometimes useful when the ratio test is hard to use.

Proposition 6.2.6 (Root Test). *Suppose*

$$\sum_{k=0}^{\infty} a_k$$

is a positive series ($a_k > 0$) for which the limit

$$\sigma = \lim \sqrt[k]{a_k}$$

exists.

1. *If $\sigma < 1$, then the series converges.*
2. *If $\sigma > 1$, then the series diverges.*
3. *If $\sigma = 1$, this gives us insufficient information to determine convergence or divergence of the series.*

Proof. Let $r = (\sigma + 1)/2$.
If $\sigma > 1$, then $\sigma > r > 1$, and for some N, $k \geq N$ guarantees

$$\sqrt[k]{a_k} > r$$

so

$$a_k > r^k \to \infty$$

and the series diverges.
If $\sigma < 1$, then $\sigma < r < 1$ and for some N, $k \geq N$ guarantees

$$\sqrt[k]{a_k} < r$$

or

$$a_k < r^k \quad (k \geq N).$$

But then, using $b_k = r^k$ in the limit comparison test,

$$c_k = \frac{a_k}{r^k}$$

is bounded, so convergence of $\sum r^k$ guarantees convergence of $\sum a_k$. □

We note in passing that the limits ρ in the ratio test and σ in the root test don't actually need to exist: all we need is the eventual estimates that the values of these limits would give us.

To formulate a more general test, we need the notion of the limit superior of a sequence. Suppose

$$\{r_k\}$$

is a sequence of nonnegative numbers. We define its **limit superior**, denoted

$$\limsup r_k$$

as follows. If the sequence is unbounded, set

$$\limsup r_k = \infty.$$

If the sequence is bounded, then for every integer N, the set $\{r_k \,|\, k \geq N\}$ is also bounded, and the numbers

$$u_N = \sup_{k \geq N} r_k$$

form a non-increasing sequence

$$u_N \geq u_{N+1} \geq \ldots \geq 0$$

(why?). Hence, they converge (by the completeness axiom) to a number,

$$\limsup r_k = \lim_{N \to \infty} \{\sup_{k \geq N} r_k\} \geq 0.$$

For example, the sequence

$$r_k = 1 + (-1)^k \left(1 - \frac{1}{k}\right) \quad k = 1, ...$$

whose first few terms are

$$1, \frac{3}{2}, \frac{1}{3}, \frac{7}{4}, \frac{1}{5}, \frac{11}{6}, \ldots$$

has

$$\limsup r_k = 2.$$

Note that if the sequence $\{r_k\}$ itself converges, then

$$\limsup r_k = \lim r_k.$$

The generalized test can then be stated as follows:

Proposition 6.2.7 (Generalized Root Test)**.** *For any series*

$$\sum_{k=0}^{\infty} a_k$$

set

$$\sigma = \limsup \sqrt[k]{|a_k|}.$$

Then:

1. *If $\sigma < 1$, the series converges absolutely.*
2. *If $\sigma > 1$, the series diverges.*
3. *If $\sigma = 1$, this test does not decide convergence or divergence of the series.*

The proof, which involves a careful reworkng of the proof of Proposition 6.2.6, is outlined in exercise Exercise 6.
Historically, it appears hard to trace the origins of the various convergence tests we have given for numerical series further back than their statement in Cauchy's *Cours d'analyse* of 1821. The recently published history of series before Cauchy by Giovanni Ferraro [24] traces a likely explanation of this: before Cauchy, series were treated in purely formal terms, and in practice their convergence was ignored or assumed. According to Ferraro [24, pp. 355-7] as well as Lützen [40, pp. 167-9], Cauchy proved the Divergence Theorem, the Comparison Test (in special cases), the Ratio Test[6] and Root Test, and showed necessity (but waved his arms at sufficiency) for convergence of the Cauchy condition on the partial sums. These appeared in the *Cours d'analyse*; Cauchy stated the Integral Test later, in 1827 ([24, p. 355]). However, the Alternating Series Test was set forth by Leibniz in letters to Johann Bernoulli in late 1713 and early 1714[7] [24, pp. 32-4].

Exercises for § 6.2

Answers to Exercises 1acegikmo, 6c are given in Appendix B.

Practice problems:

[6] According to Ferraro, it had been used previously by Carl Friedrich Gauss (1777-1855); Boyer [10, p. 500] ascribes it to Edward Waring (1734-1793).

[7] Hairer [30, p. 188] dates Leibniz's formulation of the Alternating Series Test as 1682.

1. For each series below, decide whether it converges or diverges. You don't need to evaluate the sum if it converges, but you should *explicitly* apply one of the convergence tests of this section to justify your conclusion in each case.

   (a) $\sum_{n=1}^{\infty} \frac{n^4}{2^n}$ (b) $\sum_{n=1}^{\infty} \frac{n^2}{3^n}$ (c) $\sum_{n=1}^{\infty} \frac{2^n}{n!}$

   (d) $\sum_{n=1}^{\infty} \frac{1}{n^2} \cos^2 n$ (e) $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n^2+1}}$ (f) $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n^3+1}}$

   (g) $\sum_{n=0}^{\infty} \frac{1}{(2^n+1)^p}$ (answer will depend on p) (h) $\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$ (i) $\sum_{n=1}^{\infty} \frac{\ln n}{n}$

   (j) $\sum_{n=1}^{\infty} \frac{\ln n}{n^2}$ (k) $\sum_{n=2}^{\infty} \frac{1}{\ln n}$ (l) $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n!}}$

   (m) $\sum_{n=1}^{\infty} \frac{1}{n^n}$ (n) $\sum_{n=1}^{\infty} \frac{1}{2^n+n}$ (o) $\sum_{n=4}^{\infty} \frac{1}{n!-2^n}$

2. Consider the series

$$\sum_{n=0}^{\infty} 2^{(-1)^n-n} = 2 + \frac{1}{4} + \frac{1}{2} + \frac{1}{16} + \frac{1}{8} + \frac{1}{64} \cdots$$

   (a) Show that the Ratio Test (Theorem 6.2.5) gives no information about convergence of this series (even a "generalized" version, involving lim sup and lim inf).

   (b) Use the Root Test (Proposition 6.2.6) to show that the series converges.

   (c) Is there an easier test which you could have applied here?

3. (a) Show that the series

$$\sum_{k=1}^{\infty} \frac{1}{k+m}$$

   diverges for each $m \geq 0$.

   (b) For $m < 0$, show that the series

$$\sum_{k=1}^{\infty} \frac{1}{|k+m|+1}$$

diverges.

Theory problems:

4. Show that if $a_n > 0$ for all n and $\sum a_n$ is convergent, then $\sum a_n^p$ is also convergent for all $p > 1$.

5. Suppose c_k are all positive and
$$\lim c_k = L > 0.$$
Show that for some $C > 0$
$$\frac{1}{C} < c_k < C \text{ for all } k = 1, 2, ...$$

6. Given a series
$$\sum_{k=0}^{\infty} a_k$$
with
$$\sigma = \limsup \sqrt[k]{|a_k|}$$
   (a) Show that if $\sigma < 1$, then there exists K such that the sum $\sum_{k=K}^{\infty} a_k$ converges absolutely by comparison with a geometric series. (*Hint:* Find K such that for $k \geq K$, $|a_k| < r^k$, where $r = \frac{1}{2}(\sigma + 1) < 1$) .
   (b) Show that if $\sigma > 1$, then there is a subsequence of $|a_k|$ which diverges to infinity.
   (c) Give two examples where $\sigma = 1$, for one of which the sum converges and for the other it diverges. (*Hint:* p-series) .

6.3 Unconditional Convergence

In this section, we consider the effect of order on infinite sums. If the terms of a *finite* sum are rearranged in a different order, the total sum will remain unchanged.[8] Unfortunately, this is no longer guaranteed for *infinite*

[8]This is the **commutative law** for addition.

sums. We shall see that it is fundamentally tied to the absolute convergence test (Lemma 6.2.1). The arguments in this section are rather intricate, but the two basic results, Proposition 6.3.1 and Theorem 6.3.2, are worth noting. We will make use of Proposition 6.3.1 later on.
First, we illustrate the difficulty by establishing that rearrangement can change an infinite sum. Our example is the **alternating harmonic series**

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

which converges by the alternating series test (Proposition 2.4.8). In fact, for this series,

$$\frac{1}{2} < \sum_{k=0}^{\infty} a_k < 1$$

(Exercise 2). We will rearrange this series so that it sums to 0: that is, we will construct a new sequence $\{b_k\}_{k=0}^{\infty}$ consisting of precisely the same terms but in a different order, for which

$$\sum_{k=0}^{\infty} b_k = 0.$$

The basis of the construction will be that $\sum_{k=0}^{\infty} a_k$ converges, but not absolutely. This construction proves a special case of the following general result, which can be proved by a careful modification of what we will do here. This result (and the method of proof) was first enunciated ([30, p. 192]) by Riemann in 1854:

Proposition 6.3.1. *Suppose $\sum_{k=0}^{\infty} a_k$ is a convergent series for which*

$$\sum_{k=0}^{\infty} |a_k| = \infty$$

(that is, the original series is not absolutely convergent).
Then, given any number $L \in \mathbb{R}$ there exists a rearrangement $\{b_k\}$ of $\{a_k\}$ with

$$\sum_{k=0}^{\infty} b_k = L.$$

Proof of Special Case. We will work with

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1}$$

(and $L = 0$). Our starting point is a pair of observations.
First, since $\sum_{k=0}^{\infty} a_k$ converges, the Divergence Test forces

$$\lim a_k = 0.$$

Second, since the convergence is not absolute, it is possible to show (Exercise 4) that the sum of the *positive* terms alone, as well as the sum of the *negative* terms alone, is infinite (*i.e.*, each diverges):

$$1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \ldots = \infty$$
$$-\frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \frac{1}{8} - \ldots = -\infty.$$

We construct our new series $\sum_{k=0}^{\infty} b_k$ by watching the sign of the partial sums. (See Figure 6.3.) We start with the first term

$$b_0 = 1 = a_0$$

noting that the first partial sum is positive:

$$s_0 = \sum_{k=0}^{0} b_k = 1 > 0.$$

Because of this, we take the next term from the negative terms of our original series. The first negative term is

$$b_1 = -\frac{1}{2} = a_1$$

and the corresponding partial sum

$$s_1 = \sum_{k=0}^{1} b_k = b_0 + b_1 = 1 - \frac{1}{2} = \frac{1}{2}$$

is still positive. So far, our series are the same. But now, since $s_1 > 0$, we take b_2 to be the first unused *negative* term, or

$$b_2 = -\frac{1}{4} = a_3.$$

The new partial sum

$$s_2 = \sum_{k=0}^{2} b_k = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

is still positive, so our next term is the next negative a_k,

$$b_3 = -\frac{1}{6} = a_5$$

giving

$$s_3 = \sum_{k=0}^{3} b_k = \frac{1}{4} - \frac{1}{6} = \frac{1}{12}.$$

Since this is still positive, we continue to use up our negative a_k's:

$$b_4 = -\frac{1}{8} = a_7$$

obtaining the partial sum

$$s_4 = \sum_{k=0}^{4} b_k = \frac{1}{12} - \frac{1}{8} = -\frac{1}{24}.$$

Since this partial sum is *negative*, we take our *next* b_k from the *positive* a_k's: the first unused positive term is

$$b_5 = \frac{1}{3} = a_2$$

giving the partial sum

$$s_5 = \sum_{k=0}^{5} b_k = -\frac{1}{24} + \frac{1}{3} = \frac{7}{24}.$$

This is positive, so we go back to taking our terms from the negative a_k's

$$b_6 = -\frac{1}{10} = a_9.$$

The general scheme, then, is:

> *at each stage, the next term b_{k+1} of the rearrangement is the first unused term of the original series whose sign is* opposite *to that of the current partial sum*
>
> $$s_k = \sum_{i=0}^{k} b_i.$$

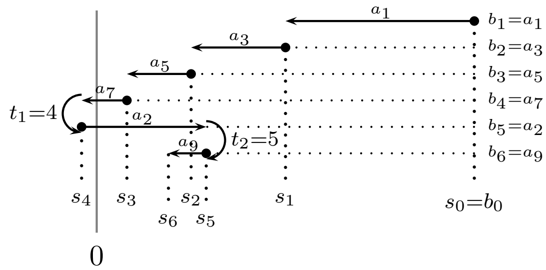


Figure 6.3: Conditional convergence

In this way, we see that the partial sums of our rearrangement are always moving toward zero, in the sense that if s_k is above (*resp.* below) 0, then s_{k+1} is below (*resp.* above) s_k. As soon as s_k overshoots 0—that is, the sign changes between s_{k-1} and s_k—we start picking our terms from the "other" list of unused original terms. Because the positive (*resp.* negative) terms alone have an infinite sum, we are guaranteed that our sums will continue to (eventually) overshoot zero repeatedly (Exercise 5a), so that our new terms b_k will eventually hit every a_k (Exercise 5b).
To see that our new series sums to zero

$$\sum_{k=0}^{\infty} b_k = 0$$

we look at the partial sums that overshoot zero. Let's call an index k a *turnaround index* if the k^{th} partial sum (of the b's) s_k has opposite sign from its predecessor s_{k-1}. We denote the successive turnaround indices by t_i, $i = 1, 2, ...$ The first few turnaround indices are

$$t_1 = 4, \quad t_2 = 5, \quad t_3 = 9, \ldots$$

We also include in our list

$$t_0 = 0.$$

By construction, the "turnaround sums" s_{t_i}, $i = 0, 1, ...$ alternate sign, and if $t_i < k < t_{i+1}$ then s_k lies between s_{t_i} and zero. Since $s_0 = 1$ is positive, we have

$$\text{for } t_i \leq k < t_{i+1}, \quad \operatorname{sgn}(b_k) = \operatorname{sgn}(s_{t_i}) = (-1)^{i+1}.$$

Note that the b_k's with index in this range are chosen to be *successive* a's with the same sign, and so the *next* b with the same sign as b_{t_i} will be

$b_{t_{i+1}+1}$. In particular, for $i = 0, 1, ...,$

$$\text{if } b_{t_i} = a_{j_1} \text{ and } b_{t_{i+2}} = a_{j_2}, \text{ then } j_1 < j_2$$

(and they have the same parity). Since

$$a_1 < a_3 < \ldots < 0 < \ldots < a_2 < a_0,$$

we conclude that

$$b_{t_1} < b_{t_3} < \ldots < 0 < \ldots < b_{t_2} < b_{t_0}$$

and, since $\lim a_j = 0$,

$$\lim b_{t_i} = 0.$$

Now, note that if $k = t_i$, then s_{k-1} and s_k lie on opposite sides of zero, so that

$$|s_{t_i}| = |s_k| \leq |s_k - s_{k-1}| = |b_k| = |b_{t_i}|,$$

and since s_{t_i} and b_{t_i} have the same sign,

$$s_{t_i} \text{ lies between zero and } b_{t_i} \text{ for } i = 0, 1, ...$$

From all of this, it follows that for $t_i \leq k \leq t_{i+1}$, s_k lies between b_{t_i} and $b_{t_{i+1}}$.

Thus, we can define two sequences that "squeeze" the s_k's, as follows: given k with $t_i \leq k < t_{i+1}$, set

$$\begin{aligned}
\alpha_k &= \min\{s_{t_i}, s_{t_{i+1}}\} \\
&= \begin{cases} s_{t_i} & \text{if } i \text{ is odd} \\ s_{t_{i+1}} & \text{if } i \text{ is even} \end{cases} \\
\beta_k &= \max\{s_{t_i}, s_{t_{i+1}}\} \\
&= \begin{cases} s_{t_{i+1}} & \text{if } i \text{ is odd} \\ s_{t_i} & \text{if } i \text{ is even.} \end{cases}
\end{aligned}$$

Then the following are easily checked:

1. $\alpha_k \leq s_k \leq \beta_k$ for $k = 0, 1, ...$

2. $\lim \alpha_k = \lim \beta_k = 0.$

Hence, the squeeze Theorem 2.4.7 gives

$$\sum_{k=0}^{\infty} b_k = \lim s_k = 0.$$

□ □

We say that a series $\sum a_k$ **converges unconditionally** if every rearrangement $\sum b_k$ has the same limit; it **converges conditionally** if different sums can be obtained via rearrangement.[9] In fact, it is possible to also show that a conditionally convergent series has rearrangements that diverge.

The previous result says that absolute convergence is *necessary* for unconditional convergence. In fact, it is also *sufficient*. This was first noted ([30, p. 192]) by Dirichlet in 1837. The general result is

Theorem 6.3.2 (Unconditional Convergence)**.** *A series*

$$\sum_{k=0}^{\infty} a_k$$

converges unconditionally precisely if it converges absolutely—that is,

$$\sum_{k=0}^{\infty} |a_k| < \infty.$$

Proof. Proposition 6.3.1 says that an unconditionally convergent series must be absolutely convergent. It remains to show that every absolutely convergent series converges unconditionally. To establish this, we revisit some ideas from the proof that absolute convergence implies convergence (Lemma 6.2.1). Let

$$A_N = \sum_{k=0}^{N} a_k$$

denote the N^{th} partial sum of the series, and

$$\tau_m = \sum_{k=m+1}^{\infty} |a_k|$$

[9] Sometimes, the term "conditionally convergent" is used when we don't *know* whether or not the convergence is unconditional. In this section, we stick to the stricter terminology.

the "tail" of its series of absolute values. Since the series is absolutely convergent, we have

$$\tau_m \to 0.$$

Suppose now that we have a rearrangement $\{b_j\}_{j=0}^{\infty}$ of $\{a_k\}_{k=0}^{\infty}$ via the correspondence

$$b_j = a_{k_j} \quad j = 0, 1, ...$$

and denote the N^{th} partial sum of the rearranged series by

$$B_N = \sum_{j=0}^{N} b_j.$$

We wish to study the difference between the N^{th} partial sum A_N of the original series and the corresponding partial sum B_N of the rearranged series. Note that if some a_k with $k \leq N$ corresponds to some b_j with $j \leq N$, (*i.e.*, $j \leq N$ and $k_j \leq N$), these two terms will cancel when we take the difference $A_N - B_N$ of the partial sums. By the triangle inequality, $|A_N - B_N|$ is bounded above by the sum of absolute values of those terms which do *not* cancel in this way.
But given m, there exists $n(m) \geq m$ such that each of the first $m + 1$ terms of the original series appears among the first $n(m) + 1$ terms of the rearranged series:

$$\{a_0, \ldots, a_m\} \subset \{b_0, \ldots, b_{n(m)}\}.$$

This means that, if $N \geq n(m)$, then each term a_k with $k \leq m$ is cancelled by some term b_j with $j \leq n(m)$ in $A_N - B_N$.
Thus, using the triangle inequality, we can bound $|A_N - B_N|$ above by a sum of terms of the form $|a_k|$ (or $|-b_j| = |a_k|$, with $k = k_j$) where each $k > m$.
But the sum of *all* such terms is the "tail" of the series of absolute values of the original series:

$$|A_N - B_N| \leq [\text{sum of terms } |a_k| \text{ with } k > m] \leq \tau_m.$$

Since $\tau_m \to 0$, we can make this as small as we like by picking N sufficiently high, and so

$$\sum_{k=0}^{\infty} a_k = \lim A_N = \lim B_N = \sum_{j=0}^{\infty} b_j$$

as required.

□

We note in passing the following important special case of Theorem 6.3.2:

Remark 6.3.3. *A positive series either diverges (to infinity) or converges unconditionally.*

Exercises for § 6.3

Answers to Exercises 1acegi are given in Appendix B.

Practice problems:

1. For each series below, decide whether it converges unconditionally or conditionally, or diverges. You should justify your conclusions.

   (a) $\sum_{n=1}^{\infty} \frac{(-1)^n n}{2^n}$ (b) $\sum_{n=1}^{\infty} \frac{(-1)^n (1-2n)}{n}$

   (c) $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ (d) $\sum_{n=1}^{\infty} \left(\frac{1}{3n} - \frac{1}{2n} \right)$

   (e) $\sum_{n=2}^{\infty} \frac{(-1)^n}{n \ln n}$ (f) $\sum_{n=0}^{\infty} \frac{(-1)^{n+1} 3^n}{n+2^n}$

   (g) $\sum_{n=0}^{\infty} \frac{(-1)^n 2^n}{3^n - n}$ (h) $\sum_{n=1}^{\infty} \frac{(-1)^n 3^n}{n!}$

   (i) $\sum_{n=1}^{\infty} n^{-3/2} \sin n$

Theory problems:

2. Show that the alternating harmonic series

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

   satisfies

$$\frac{1}{2} < \sum_{k=0}^{\infty} a_k < 1.$$

   (*Hint:* Look closely at the proof of Proposition 2.4.8 in § 2.4.)

3. Show that a sum of absolutely convergent series is itself absolutely convergent: if $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ both converge absolutely, then so does $\sum_{n=0}^{\infty}(a_n + b_n)$.

4. Show that if a series is convergent but not absolutely convergent, then the sum of the *positive* terms alone *diverges* to infinity, and the sum of the *negative* terms alone diverges to $-\infty$. (*Hint:* Proof by contradiction.)

Challenge problem:

5. (a) Show that the scheme on p. 508 in the proof of Proposition 6.3.1 will continue to overshoot zero. (*Hint:* If not, say if all partial sums after some point are *positive*, then the sum of all the subsequent *negative* terms will be bounded, a contradiction to Exercise 4.)
   (b) Show that the scheme on p. 508 in the proof of Proposition 6.3.1 will eventually include all the terms a_k.
   (c) How can the proof of the special case $L = 0$ of Proposition 6.3.1 be adapted to handle an arbitrary value of L?

6.4 Convergence of Power Series

At the end of § 6.1, we considered "taking the degree to infinity" in the Taylor approximations for $f(x)$ at $x = a$. Formally, this leads to an expression

$$\sum_{k=0}^{\infty} c_k(x-a)^k = c_0 + c_1(x-a) + c_2(x-a)^2 + \ldots .$$

Such an expression is called a **power series about** $x = a$. The **coefficients** c_k, $k = 0, 1, \ldots$ and the **basepoint** a are given numbers, while x is a real variable. If we attempt to evaluate this expression at a particular point x, we obtain an infinite series in $\mathbb{R}$. Thus, we have a function

$$p(x) = \sum_{k=0}^{\infty} c_k(x-a)^k$$

whose domain is the set of x-values for which this infinite series converges, the **domain of convergence** of the power series.
Note that regardless of the values of the coefficients, *the series always converges at the basepoint*; in fact, all the terms after the first are zero, so the series sums to its constant term:

$$p(a) = \sum_{k=0}^{\infty} c_k(a-a)^k = c_0 + c_1(0) + c_2(0)^2 + \ldots = c_0 + 0 + 0 + \ldots = c_0.$$

However, at any point $x \neq a$ other than the basepoint, convergence needs to be tested explicitly.
We explore several examples before formulating some general results.
First, the power series about $x = 0$

$$\sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \ldots$$

is a geometric series with ratio $r = x$, so Proposition 2.4.6 tells us that the series diverges if $|x| \geq 1$, while if $|x| < 1$ it converges to

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \quad |x| < 1.$$

That is, the domain of convergence of this series is the open interval $(-1, 1)$, and on this interval the series yields the function $1/(1-x)$. Note that, even though the latter formula makes sense at points *outside* the interval $(-1, 1)$ (*e.g.*, at $x = 2$, $1/(1-x) = -1$), the *series* does not converge there.
Note that if we use the same coefficients

$$c_k = 1, \quad k = 0, 1, ...$$

but move the basepoint to another position a, the resulting power series

$$\sum_{k=0}^{\infty} (x-a)^k = 1 + (x-a) + (x-a)^2 + \ldots$$

is again geometric, but this time with ratio $r = x - a$. We then find

$$\sum_{k=0}^{\infty} (x-a)^k = \frac{1}{1-(x-a)} = \frac{1}{(1+a)-x} \quad |x-a| < 1$$

with domain of convergence the open interval

$$\{x \mid |x-a| < 1\} = (a-1, a+1)\,.$$

As a second example, consider the series about $x = 0$

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots.$$

This is no longer a geometric series: the ratio between successive terms is

$$\frac{a_{k+1}(x)}{a_k(x)} = \frac{x^{k+1}/(k+1)!}{x^k/k!} = \frac{k!}{(k+1)!}x = \frac{x}{k+1}.$$

However, for any particular value of x, this ratio converges (as $k \to \infty$) to zero. Thus, using the ratio test, we have (at the point x)

$$\rho(x) = \lim \frac{|a_{k+1}(x)|}{|a_k(x)|} = \lim \frac{|x|}{k+1} = 0 < 1$$

so the series converges everywhere: the domain of convergence is the "open interval" $(-\infty, \infty)$.
By contrast, consider the series about $x = 0$

$$\sum_{k=0}^{\infty} k!x^k = 1 + x + 2x^2 + 6x^3 + \ldots.$$

This time, the ratio between successive terms is

$$\frac{a_{k+1}(x)}{a_k(x)} = \frac{(k+1)!x^{k+1}}{k!x^k} = (k+1)x$$

which is unbounded if $x \neq 0$. It follows that the series *diverges* at every $x \neq 0$; the domain of convergence consists of the basepoint $a = 0$ alone, which we can refer to as the (degenerate) closed interval

$$\{0\} = [0, 0]\,.$$

In general, the domain of convergence of a power series is an interval centered at the basepoint a: that is, it contains a and extends the same distance R in either direction from a. R is referred to as the **radius of convergence** of the series: it is determined by the sequence of coefficients

c_k, and may be any non-negative number or $+\infty$. The following theorem[10] gives more details.

Theorem 6.4.1 (Convergence of Power Series)**.** *Consider the power series*

$$\sum_{k=0}^{\infty} c_k(x-a)^k = c_0 + c_1(x-a) + c_2(x-a)^2 + \dots.$$

The domain of convergence is an interval I described by one of the following:

1. *the series converges only at the basepoint $x = a$; $I = [a, a]$*
2. *the series converges for* all *$x \in \mathbb{R}$; $I = (-\infty, \infty)$.*
3. *there exists a positive real number $0 < R < \infty$ such that the series*
   - *diverges if $|x - a| > R$*
   - *converges if $|x - a| < R$,*

   so that I is an interval with endpoints $x = a \pm R$; no general statement can be made about the behavior of the series at the endpoints: I may be open, closed, or half-open.

We define the radius of convergence R to be $R = 0$ in the first case and $R = \infty$ in the second.
If the coefficients are eventually all nonzero and the limit

$$\rho = \lim \frac{|c_{k+1}|}{|c_k|}$$

exists, then

$$R = \frac{1}{\rho}.$$

More generally, the same formula holds if we take

$$\rho = \limsup \sqrt[k]{|c_k|}.$$

Furthermore, the series converges absolutely *if $|x - a| < R$.*
(See Figure 6.4.)

[10]This result (along with the first formula for ρ) was implicit in Cauchy's treatment of 1821, but was first stated explicitly and established by Niels Henrik Abel (1802-1829) in 1826 [1][49, pp. 286-291]. Abel contributed important results to the rigorization of calculus as well as to algebra and complex analysis before his death at the age of 26. The second formula for ρ was given by Jacques Salomon Hadamard (1856-1963) in 1892 [29]. [30, p. 249]

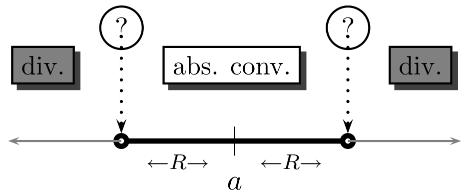


Figure 6.4: Convergence of power series

Proof. Let us first handle the easier case when all coefficients are nonzero and the ratios of successive coefficients converge:

$$\lim \frac{|c_{k+1}|}{|c_k|} = \rho \in [0, \infty) \,.$$

Let us then apply the ratio test to the power series evaluated at the non-basepoint $x \neq a$: we have

$$\rho(x) = \lim \frac{|a_{k+1}(x)|}{|a_k(x)|} = \lim \frac{\left|c_{k+1}(x-a)^{k+1}\right|}{|c_k(x-a)^k|} = \rho \cdot |x-a| \,.$$

The ratio test tells us that the series converges absolutely whenever $\rho(x) < 1$ and diverges whenever $\rho(x) > 1$. If $\rho > 0$, set $R = 1/\rho$ and note that $\rho(x) < 1$ (*resp.* > 1) precisely when $|x-a| < R$ (*resp.* $> R$). If $\rho = 0$, then $\rho(x) = 0 < 1$ for all x, which corresponds to $R = \infty$ (formally, $R = 1/\rho$ also).

For the more general case it is more efficient to use the generalized version of the root test, Proposition 6.2.7. We consider the sequence $\{\sqrt[k]{|c_k|}\}$ and set

$$\sigma = \limsup \sqrt[k]{|c_k|}.$$

Define R formally by

$$R = \frac{1}{\sigma}$$

(so $R = 0$ if $\sigma = \infty$ and $R = \infty$ if $\sigma = 0$).

Suppose $|x-a| < R$, so that

$$\sigma \cdot |x-a| < 1.$$

Then let $\beta > \sigma$ with $\beta\,|x-a| < 1$. For some N, $k \geq N$ guarantees

$$\sqrt[k]{|c_k|} < \beta$$

so that

$$\left|c_k(x-a)^k\right| < |\beta(x-a)|^k \quad (k \geq N).$$

Now, the sum of the terms on the right is geometric and converges (since the ratio is $\beta \cdot |x-a| < 1$), hence the (limit) comparison test gives absolute convergence of the power series whenever $|x-a| < R$.
If $|x-a| > R$, pick $\alpha < \sigma$ so that $\alpha \cdot |x-a| > 1$; then we can find a sequence of indices $k_i \to \infty$ for which

$$\sqrt[k_i]{|c_{k_i}|} \geq \alpha$$

and so

$$\left|c_{k_i}(x-a)^{k_i}\right| \geq |\alpha(x-a)|^{k_i} \to \infty.$$

In particular, the divergence test forces the power series to diverge when $|x-a| > R$. □

In view of this result, the domain of convergence is often called the **interval of convergence** .
In practice, the radius of convergence can often be found using the ratio test as in the preceding examples. The endpoints of the interval of convergence, however, require case-by-case study. We consider a few more examples.
The series

$$\sum_{k=1}^{\infty} \frac{x^k}{k} = x + \frac{x^2}{2} + \frac{x^3}{3} + \dots$$

has

$$\rho(x) = \lim \frac{\left|x^{k+1}/(k+1)\right|}{x^k/k} = \lim \left(\frac{k}{k+1}x\right) = |x|\,.$$

Thus, we have absolute convergence for $|x| < 1$, that is, for $-1 < x < 1$, and divergence for $|x| > 1$. As for the endpoints, we have at $x = 1$ the harmonic series

$$\sum_{k=1}^{\infty} \frac{1^k}{k} = \sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \dots$$

which diverges, while at $x = -1$ we have

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k} = -1 + \frac{1}{2} - \frac{1}{3} + \dots$$

which up to a factor of (-1) is the alternating harmonic series; this is conditionally convergent. So the interval of convergence is $[-1, 1)$.

The series
$$\sum_{k=0}^{\infty} \frac{3^k x^k}{(k+1)^2} = 1 + \frac{3x}{4} + x^2 + \frac{27x^3}{16} + \dots$$
has
$$\rho(x) = \lim \frac{\left|3^{k+1}x^{k+1}/(k+2)^2\right|}{|3^k x^k/k^2|} = \lim \left[3\,|x|\,\frac{(k+1)^2}{(k+2)^2}\right] = 3\,|x|$$
so the series converges absolutely for $|x| < \frac{1}{3}$ and diverges for $|x| > \frac{1}{3}$. At the endpoints, $x = \pm\frac{1}{3}$, we have
$$\sum_{k=0}^{\infty} \frac{3^k(\pm\frac{1}{3})^k}{(k+1)^2} = \sum_{k=0}^{\infty} \frac{(\pm 1)^k}{(k+1)^2}.$$
This is absolutely convergent at both endpoints, so the interval of convergence is $\left[-\frac{1}{3}, \frac{1}{3}\right]$.

The power series
$$1 + 2(x-1)^2 + 3(x-1)^4 + 4(x-1)^6 + \dots$$
has every other coefficient zero, so we cannot take ratios of the usual form $a_{k+1}(x)/a_k(x)$. However, we can think of this series as
$$\sum_{k=0}^{\infty} (k+1)\left[(x-1)^2\right]^k$$
and take the ratio of successive *nonzero* terms:
$$\frac{\left|(k+2)(x-1)^{2(k+1)}\right|}{|(k+1)(x-1)^{2k}|} = \frac{k+2}{k+1}(x-1)^2 \to (x-1)^2.$$
So the ratio test still applies: if $(x-1)^2 < 1$, we have convergence and if $(x-1)^2 > 1$ we have divergence. Note that these inequalities are equivalent to $|x-1| < 1$ (*resp.* $|x-1| > 1$). At the two endpoints $x = 2$ $(x-1 = 1)$ and $x = 0$ $(x-1 = -1)$ we have $(x-1)^2 = 1$, so both endpoint values of x lead to the series
$$1 + 2 + 3 + 4 + \dots$$
which clearly diverges. Our interval of convergence, then, is the open interval $(0, 2)$.

Exercises for § 6.4

Answers to Exercises 1-21 (odd only), 22 are given in Appendix B.

Practice problems:

In problems 1-21, (i) find the radius of convergence, (ii) identify the endpoints of the interval of convergence, and (iii) give the interval of convergence.

1. $\displaystyle\sum_{k=1}^{\infty} \frac{x^{k+1}}{k^2}$
2. $\displaystyle\sum_{k=1}^{\infty} \frac{x^k}{\sqrt{k}}$
3. $\displaystyle\sum_{k=0}^{\infty} \frac{x^k}{3k+1}$
4. $\displaystyle\sum_{k=0}^{\infty} \frac{3x^k}{k+1}$
5. $\displaystyle\sum_{k=1}^{\infty} \frac{x^k}{k^2+2k}$
6. $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k x^k}{\sqrt{k+1}}$
7. $\displaystyle\sum_{k=1}^{\infty} \frac{2x^k}{k!}$
8. $\displaystyle\sum_{k=1}^{\infty} \frac{(-2)^k x^k}{k!}$
9. $\displaystyle\sum_{k=1}^{\infty} \frac{x^{k-1}}{k!}$
10. $\displaystyle\sum_{k=1}^{\infty} \frac{x^k}{k^k}$
11. $\displaystyle\sum_{k=2}^{\infty} \frac{(-x)^k \ln k}{k}$
12. $\displaystyle\sum_{k=2}^{\infty} \frac{x^{2k}}{\ln k}$
13. $\displaystyle\sum_{k=0}^{\infty} \frac{x^{2k}}{2^k}$
14. $\displaystyle\sum_{k=0}^{\infty} \frac{x^{2k}}{k+1}$
15. $\displaystyle\sum_{k=0}^{\infty} \frac{3^{2k+1} x^k}{2k+1}$
16. $\displaystyle\sum_{k=1}^{\infty} \frac{kx^k}{2^k}$
17. $\displaystyle\sum_{k=1}^{\infty} \frac{x^k}{k2^k}$
18. $\displaystyle\sum_{k=1}^{\infty} \frac{(x-1)^k}{k}$
19. $\displaystyle\sum_{k=1}^{\infty} \frac{(2x+1)^k}{k^2}$
20. $\displaystyle\sum_{k=0}^{\infty} \frac{5^k (x-2)^k}{(k+1)^2}$
21. $\displaystyle\sum_{k=0}^{\infty} \frac{3^k (1-x)^k}{k+1}$

Challenge problem:

22. Consider the power series

$$\sum_{k=1}^{\infty} \frac{k^k x^k}{k!}.$$

(a) Find

$$\lim \left(\frac{(k+1)^{k+1}}{(k+1)!} \cdot \frac{k!}{k^k} \right).$$

(*Hint:* There are some fortuitous cancellations.)

(b) Find the radius of convergence of the series.

(c) Identify which numerical series you would have to test for convergence at the endpoints. (You are not expected to carry out this test.)

6.5 Handling Power Series

We saw in the last section that a power series about $x = a$

$$p(x) = \sum_{k=0}^{\infty} c_k (x-a)^k$$

defines a function on an interval I which includes the basepoint a. These functions formally resemble polynomials, albeit of "infinite degree", and in many ways they can be handled as if they were polynomials, provided we stay within the interval of convergence. In most cases, this kind of handling is justified by seeing what happens to the partial sums (which are *honest* polynomials) and then going to the limit and using Theorem 2.4.1. The distributive law tells us that we can multiply a polynomial

$$p_N(x) = c_0 + c_1 x + \ldots + c_N x^N = \sum_{k=0}^{N} c_k x^k$$

by a number or a power of x by multiplying each term by the same factor:

$$\begin{aligned}
\gamma p_N(x) &= \gamma c_0 + \gamma c_1 x + \ldots + \gamma c_N x^N &&= \sum_{k=0}^{N} \gamma c_k x^k \\
x^n p_N(x) &= c_0 x^n + c_1 x^{n+1} + \ldots + c_N x^{n+N} &&= \sum_{k=0}^{N} c_k x^{k+n} \\
& &&= \sum_{i=n}^{n+N} c_i x^i.
\end{aligned}$$

The same holds for power series:

Remark 6.5.1 (Distributive Law for Power Series)**.** *Suppose*

$$p(x) = \sum_{k=0}^{\infty} c_k (x-a)^k = c_0 + c_1(x-a) + c_2(x-a)^2 + \ldots$$

is a power series about $x = a$ with interval of convergence I.

1. *For any number γ, let*

$$b_k = \gamma c_k \quad k = 0, 1, ...$$

so that

$$\begin{aligned}\sum_{k=0}^{\infty} b_k(x-a)^k &= \gamma c_0 + \gamma c_1(x-a) + \gamma c_2(x-a)^2 + \dots \\ &= \sum_{k=0}^{\infty} \gamma c_k(x-a)^k\end{aligned}$$

is the power series that results from multiplying each term of $p(x)$ by γ.

Then $\sum_{k=0}^{\infty} b_k(x-a)^k$ converges for each $x \in I$ to $\gamma p(x)$:

$$\begin{aligned}\sum_{k=0}^{\infty} b_k(x-a)^k &= \sum_{k=0}^{\infty} \gamma c_k(x-a)^k \\ &= \gamma p(x) = \gamma \sum_{k=0}^{\infty} c_k(x-a)^k \quad x \in I\end{aligned}$$

in fact, if $\gamma \neq 0$, the interval of convergence of $\sum_{k=0}^{\infty} b_k(x-a)^k$ is precisely I.

2. *For any positive integer n, let*

$$d_k = \begin{cases} 0 & \textit{for } k < n \\ c_{k-n} & \textit{for } k \geq n \end{cases}$$

so that

$$\begin{aligned}\sum_{k=0}^{\infty} d_k(x-a)^k &= c_0(x-a)^n + c_1(x-a)^{n+1} + c_2(x-a)^2 + \dots \\ &= \sum_{k=0}^{\infty} c_k(x-a)^{k+n} = \sum_{k=0}^{\infty}(x-a)^n\big(c_k(x-a)^k\big)\end{aligned}$$

is the series that results from multiplying each term of $p(x)$ by $(x-a)^n$.

Then the interval of convergence for $\sum_{k=0}^{\infty} d_k(x-a)^k$ *is also* I*, and for each* $x \in I$*,*

$$\sum_{k=0}^{\infty} d_k(x-a)^k = \sum_{k=0}^{\infty} (x-a)^n [c_k(x-a)^k]$$
$$= (x-a)^n p(x) = (x-a)^n \sum_{k=0}^{\infty} c_k(x-a)^k.$$

Proof. These are just applications of Theorem 2.4.1 together with the distributive law applied to the partial sums.
For the first statement, we have, for $x \in I$,

$$\begin{aligned}\sum_{k=0}^{\infty} b_k(x-a)^k = \sum_{k=0}^{\infty} \gamma c_k(x-a)^k \qquad &= \lim_{N\to\infty} \sum_{k=0}^{N} \gamma c_k(x-a)^k \\ = \lim_{N\to\infty} \gamma \sum_{k=0}^{N} c_k(x-a)^k &= \gamma \lim_{N\to\infty} \sum_{k=0}^{N} c_k(x-a)^k \\ = \gamma \sum_{k=0}^{\infty} c_k(x-a)^k \qquad &= \gamma p(x)\,.\end{aligned}$$

The proof of the second statement is exactly analogous (Exercise 40). □

We also add (or subtract) polynomials termwise, and the same applies to power series (provided thay have a common basepoint).

Remark 6.5.2 (Termwise Addition of Power Series)**.** *Suppose*

$$p(x) = \sum_{k=0}^{\infty} a_k(x-a)^k$$
$$q(x) = \sum_{k=0}^{\infty} b_k(x-a)^k$$

are power series about the same basepoint $x = a$*, with respective intervals of convergence* I_a *and* I_b*.*
Let

$$r(x) = \sum_{k=0}^{\infty} c_k(x-a)^k = \sum_{k=0}^{\infty} (a_k + b_k)(x-a)^k = \sum_{k=0}^{\infty} \left(a_k(x-a)^k + b_k(x-a)^k\right)$$

be the series obtained formally by adding $p(x)$ and $q(x)$ termwise. Then $r(x)$ converges at least on the interval $I = I_a \cap I_b$, and for $x \in I$

$$r(x) = p(x) + q(x)$$

—that is,

$$\sum_{k=0}^{\infty} \left(a_k(x-a)^k + b_k(x-a)^k\right) = \sum_{k=0}^{\infty} a_k(x-a)^k + \sum_{k=0}^{\infty} b_k(x-a)^k \quad x \in I.$$

Proof. The corresponding statement for partial sums is true (by the commutative law of addition), and we take the limit: if $p(x)$ and $q(x)$ both converge at a particular value of x, then

$$\begin{aligned} r(x) = \sum_{k=0}^{\infty} c_k(x-a)^k &= \lim_{N\to\infty} \sum_{k=0}^{N} c_k(x-a)^k = \lim_{N\to\infty} \sum_{k=0}^{N} (a_k + b_k)(x-a)^k \\ &= \lim_{N\to\infty} [\sum_{k=0}^{N} a_k(x-a)^k + \sum_{k=0}^{N} b_k(x-a)^k] \\ &= \lim_{N\to\infty} \sum_{k=0}^{N} a_k(x-a)^k + \lim_{N\to\infty} \sum_{k=0}^{N} b_k(x-a)^k \\ &= \sum_{k=0}^{\infty} a_k(x-a)^k + \sum_{k=0}^{\infty} b_k(x-a)^k = p(x) + q(x)\,. \end{aligned}$$

□

As a quick application of this result, we consider

$$\sum_{k=0}^{\infty} (3 + 2^k)x^k = 4 + 5x + 7x^2 + 9x^3 + \ldots$$

which has the form $\sum_{k=0}^{\infty} c_k(x-a)^k$ where $a = 0$, $c_k = a_k + b_k$ with $a_k = 3$, $b_k = 2^k$. Let us examine the two power series about $x = 0$ whose coefficients are a_k (*resp.* b_k). The first series is geometric

$$\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} 3x^k = 3\sum_{k=0}^{\infty} x^k = 3\left(\frac{1}{1-x}\right) = \frac{3}{1-x} \quad |x| < 1.$$

The second can be viewed as

$$\sum_{k=0}^{\infty} b_k x^k = \sum_{k=0}^{\infty} 2^k x^k = \sum_{k=0}^{\infty} (2x)^k = \frac{1}{1-2x}, \quad |2x| < 1\text{—}i.e., \quad |x| < \frac{1}{2}.$$

Thus we know that at least for $|x| < \frac{1}{2}$,

$$\sum_{k=0}^{\infty}(3+2^k)x^k = \frac{3}{1-x} + \frac{1}{1-2x} = \frac{4-7x}{(1-x)(1-2x)} \qquad |x| < \frac{1}{2}.$$

We also differentiate polynomials termwise, and this carries over to power series, provided we stay away from the endpoints of the interval of convergence. Recall that the derivative of a polynomial

$$p_N(x) = c_0 + c_1x + c_2x^2 + \ldots + c_Nx^N = \sum_{k=0}^{N} c_kx^k$$

is

$$p_N'(x) = c_1 + 2c_2x + \ldots + Nc_Nx^{N-1} = \sum_{k=1}^{N} kc_kx^{k-1}.$$

For a power series at $x = a$

$$p(x) = \sum_{k=0}^{\infty} c_k(x-a)^k$$

we take its **formal derivative** by differentiating termwise:

$$p'(x) = \sum_{k=1}^{\infty} kc_k(x-a)^{k-1} = \sum_{i=0}^{\infty} c_i'(x-a)^i,$$

where

$$c_i' = (i+1)c_{i+1}, \quad i = 0, 1, \ldots.$$

If $p(x)$ has a positive radius of convergence $R > 0$, then it is a well-defined function, at least on the open interval $(a-R, a+R)$, and at any point $x_0 \in (a-R, a+R)$ we can compare the formal derivative $p'(x_0)$ at $x = x_0$ with the actual derivative of p at $x = 0$:

$$\left.\frac{d}{dx}\right|_{x=x_0} [p(x)] = \lim_{x \to x_0} \frac{p(x) - p(x_0)}{x - x_0}.$$

There are two issues here: whether the series $p'(x)$ given by the formal derivative converges at $x = x_0$, and if so, whether the sum of this series equals the actual derivative. It turns out that both questions are answered in the affirmative.

Theorem 6.5.3 (Differentiation of Power Series)**.** [11] *Suppose*

$$p(x) = \sum_{k=0}^{\infty} c_k(x-a)^k = c_0 + c_1(x-a) + c_2(x-a)^2 + \dots$$

is a power series about $x = a$ with positive radius of convergence R $(0 < R \leq \infty)$.
Then

1. *the formal derivative*

$$p'(x) = \sum_{k=1}^{\infty} kc_k(x-a)^{k-1} = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \dots$$
$$= \sum_{i=1}^{\infty} c_i'(x-a)^i$$

has the same radius of convergence R;

2. *at every $x = x_0$ with $|x_0 - a| < R$, the function p is differentiable, and its derivative at $x = x_0$ equals the sum (at $x = x_0$) of the formal derivative:*

$$p'(x_0) = \left.\frac{d}{dx}\right|_{x=x_0} [p(x)] \quad |x_0 - a| < R.$$

We separate the proofs of these two statements into two lemmas. The first deals with convergence:

Lemma 6.5.4. *If*

$$p(x) = \sum_{k=0}^{\infty} c_k(x-a)^k$$

has radius of convergence R, then its formal derivative

$$p'(x) = \sum_{k=1}^{\infty} kc_k(x-a)^{k-1}$$

also has radius of convergence R.

[11] Cauchy stated and gave a proof of this result in 1821, but the proof was incorrect. I have not been able to determine who gave the first correct proof. The more fundamental question of *continuity* of p was noted and handled by Abel in 1826: see Proposition 6.5.6 below.

Proof of Lemma 6.5.4. Instead of dealing directly with $p'(x)$, we deal with the related series

$$(x-a)p'(x) = \sum_{k=1}^{\infty} kc_k(x-a)^k$$

which by Remark 6.5.1 has the same radius of convergence as $p'(x)$. Recall from Theorem 6.4.1 that the radius of convergence can be found using the generalized root test on the coefficients. We know (applying the generalized root test to $p(x)$) that

$$\frac{1}{R} = \limsup \sqrt[k]{|c_k|}.$$

Now, applying the generalized root test to $(x-a)p'(x)$, we have

$$\limsup \sqrt[k]{|kc_k|} = \lim \sqrt[k]{k} \limsup \sqrt[k]{|c_k|}$$

and since $k^{1/k} \to 1$,

$$\limsup \sqrt[k]{|kc_k|} = \frac{1}{R}.$$

□

The proof of the second statement is more involved.

Lemma 6.5.5. *If $|x_0 - a| < R$, where $0 < R \leq \infty$ is the radius of convergence of the power series*

$$p(x) = \sum_{k=0}^{\infty} c_k(x-a)^k$$

then the formal derivative

$$p'(x) = \sum_{i=1}^{\infty} c'_i(x-a)^i = \sum_{k=1}^{\infty} kc_k(x-a)^{k-1}$$

converges, at $x = x_0$, to the actual derivative of $p(x)$ at $x = x_0$:

$$p'(x_0) = \left.\frac{d}{dx}\right|_{x=x_0}[p(x)] = \lim_{x\to x_0} \frac{p(x)-p(x_0)}{x-x_0}.$$

Proof of Lemma 6.5.5. The calculations will be easier to carry out if we assume the basepoint is $a = 0$. This amounts to proving the result for

$$q(x) = p(x+a)$$

at $x = x_0 - a$; the series for $q(x)$ (*resp.* $q'(x)$) have the same coefficients and radius of convergence as $p(x)$ (*resp.* $p'(x)$), and the general version follows easily (Exercise 41). So we assume, without loss of generality, that $a = 0$. We will prove the result using Proposition 6.1.2. Set

$$\phi(x) = \alpha_0 + \alpha_1(x - x_0)$$

where

$$\begin{aligned}\alpha_0 &= p(x_0) = \sum_{k=0}^{\infty} c_k x_0^k \\ \alpha_1 &= p'(x_0) = \sum_{k=0}^{\infty} c'_k x_0^k \quad c'_k = (k+1)c_{k+1}.\end{aligned}$$

Note that $\phi(x)$ can be written (for $|x| < R$) as

$$\begin{aligned}\phi(x) &= \sum_{k=0}^{\infty} c_k x_0^k + (x - x_0) \sum_{k=1}^{\infty} k c_k x_0^{k-1} \\ &= c_0 + \sum_{k=1}^{\infty} c_k \big(x_0^k + (x - x_0) k x_0^{k-1}\big).\end{aligned}$$

Our goal is to prove that

$$p(x) - \phi(x) = \mathfrak{o}(|x - x_0|) \text{ as } x \to x_0$$

which we will do by establishing an estimate of the form

$$|p(x) - \phi(x)| \leq M\,|x - x_0|^2$$

for x sufficiently close to x_0.
To obtain this estimate, we pick x_1 with

$$|x_0| < |x_1| < R$$

so that $p(x_1)$ and $p'(x_1)$ both converge absolutely. All our estimates will assume that x is near enough to x_0 that

$$|x| < |x_1|$$

also.

We can write (using unconditional convergence of all the series)

$$\begin{aligned} p(x) - \phi(x) &= \sum_{k=0}^{\infty} c_k x^k - \left(c_0 + \sum_{k=1}^{\infty} c_k \big(x_0^k + (x - x_0) k x_0^{k-1} \big) \right) \\ &= \left(c_0 + \sum_{k=1}^{\infty} c_k x^k \right) - \left(c_0 + \sum_{k=1}^{\infty} c_k \big(x_0^k + (x - x_0) k x_0^{k-1} \big) \right) \\ &= \sum_{k=1}^{\infty} c_k \left(x^k - x_0^k - (x - x_0) k x_0^{k-1} \right). \end{aligned}$$

Note that for $k = 1$, the quantity in large parentheses is zero, while for each $k = 2, ...$, it is the error $f_k(x) - T^1_{x_0} f_k\,(x)$ for

$$f_k(x) = x^k.$$

Thus by Taylor's theorem (Theorem 6.1.7), given x, there exists s_k between x and x_0 for which this error equals

$$\frac{(x - x_0)^2}{2} f_k''(s_k) = \frac{k(k-1)}{2} s_k^{k-2} (x - x_0)^2.$$

Since $|s_k| < |x_1|$, we can guarantee that, for $k = 1, 2, ...$

$$\begin{aligned} \left| x^k - x_0^k - (x - x_0) k x_0^{k-1} \right| &= \frac{k(k-1)}{2} |s_k|^{k-2} |x - x_0|^2 \\ &\leq \frac{k(k-1)}{2} |x_1|^{k-2} |x - x_0|^2 . \end{aligned}$$

Applying termwise absolute values, we obtain (for $|x| < |x_1|$)

$$\begin{aligned} |p(x) - \phi(x)| &\leq \sum_{k=1}^{\infty} |c_k| \left| x^k - x_0^k - (x - x_0) k x^{k-1} \right| \\ &\leq \left(\sum_{k=2}^{\infty} |c_k| \frac{k(k-1)}{2} |x_1|^{k-2} \right) |x - x_0|^2 . \end{aligned}$$

Now, this last series looks like the absolute convergence test applied to the formal derivative $p''(x)$ of $p'(x)$ at $x = x_1$; by Lemma 6.5.4, this new series $p''(x)$ also has radius of convergence R, and since $|x_1| < R$, we conclude that

$$\sum_{k=2}^{\infty} \frac{k(k-1)}{2} |c_k| |x_1|^{k-2} = M$$

for some $M \in \mathbb{R}$. So we have shown that, for every x with $|x| < |x_1|$,

$$|p(x) - \phi(x)| \leq M\,|x - x_0|^2$$

and in particular

$$\frac{|p(x) - \phi(x)|}{|x - x_0|} \leq M\,|x - x_0| \to 0 \text{ as } x \to x_0.$$

But this amounts to the required condition

$$p(x) - \phi(x) = \mathfrak{o}(|x - x_0|) \text{ as } x \to x_0.$$

□

Taken together, these two lemmas prove Theorem 6.5.3.
A simple corollary of Theorem 6.5.3 is the continuity of $p(x)$ on the *interior* of its interval of convergence (why?). However, as we shall see below, it can be useful to have continuity as well at either endpoint of this interval, provided the series also converges there. This was established by Niels Henrik Abel (1802-1829) in 1826 [1]. We work out a proof in Exercise 47.

Proposition 6.5.6 (Abel's Theorem)**.** *If the power series*

$$p(x) = \sum_{k=0}^{\infty} c_k (x - a)^k$$

has radius of convergence $R > 0$, then $p(x)$ is continuous on its whole interval of convergence.

Another corollary of Theorem 6.5.3 is the termwise integration of power series.

Lemma 6.5.7. *If the power series*

$$p(x) := \sum_{k=0}^{\infty} c_k (x - a)^k$$

has radius of convergence $0 < R \leq \infty$, then the formal indefinite integral

$$P(x) := \sum_{k=0}^{\infty} \frac{c_k}{k+1} (x - a)^{k+1}$$

also has radius of convergence R, and is the antiderivative of p which vanishes at x = a:

$$P'(x) = p(x) \text{ for } |x-a| < R, \text{ and } P(a) = 0.$$

That is,

$$P(x) = \int_a^x p(t)\ dt.$$

Proof. The argument that $P(x)$ has the same radius of convergence as $p(x)$ is the same as the proof of Lemma 6.5.4. But then, we can apply Lemma 6.5.5 to $P(x)$ to prove that $p(x) = P'(x)$ in the interval of convergence. □

These results can be used to greatly extend our repertoire of power series whose sum we can evaluate explicitly.
As a first example, we start with $c_k = 1$, $k = 0, ...$

$$p(x) = \sum_{k=0}^{\infty}(x-a)^k = 1 + (x-a) + (x-a)^2 + \dots$$

which is geometric with ratio $|x-a|$ and first term 1, so that

$$p(x) = \sum_{k=0}^{\infty}(x-a)^k = \frac{1}{1-(x-a)} = \frac{1}{1+a-x} \text{ for } |x-a| < 1.$$

Now, the formal derivative is

$$p'(x) = \sum_{k=1}^{\infty} k(x-a)^{k-1} = 1 + 2(x-a) + 3(x-a)^2 + \dots$$

which, by Theorem 6.5.3, converges for $|x-a| < 1$ to the derivative of p,

$$\frac{dp}{dx} = \frac{1}{(1+a-x)^2}$$

so we have

$$\sum_{k=1}^{\infty} k(x-a)^{k-1} = \sum_{i=0}^{\infty}(i+1)(x-a)^i = \frac{1}{(1+a-x)^2} \quad |x-a| < 1.$$

We can, of course, repeat the process: for example, differentiation of this last series gives

$$\sum_{k=2}^{\infty} k(k-1)(x-a)^{k-2} = 2 + 6(x-a) + 12(x-a)^2 + 20(x-a)^3 + \dots$$
$$= \frac{2}{(1+a-x)^3} \quad |x-a| < 1.$$

A second example takes us beyond rational functions. The series about $a = 1$

$$\sum_{k=0}^{\infty} (-1)^k (x-1)^k = \sum_{k=0}^{\infty} (1-x)^k$$

is geometric, with first term 1 and ratio $r = 1 - x$, so

$$\sum_{k=0}^{\infty} (1-x)^k = \frac{1}{1-(1-x)} = \frac{1}{x} \quad |x-1| < 1 \ (i.e., \ \ 0 < x < 2).$$

Now, $1/x$ is the derivative of $\ln x$, so *if* we had a power series about $a = 1$ for $\ln x$

$$\sum_{k=0}^{\infty} c_k (x-1)^k$$

then its formal derivative would have to give $1/x$, which is to say

$$\sum_{k=1}^{\infty} k c_k (x-1)^{k-1} = \sum_{i=0}^{\infty} (i+1) c_{i+1} (x-1)^i = \sum_{i=0}^{\infty} (-1)^i (x-1)^i.$$

Matching coefficients, we have

$$(i = 1) c_{i+1} = (-1)^i \quad i = 0, 1, ...$$

or

$$c_k = \frac{(-1)^{k-1}}{k} \quad k = 1, 2, ...$$

This determines all coefficients except the constant term. But c_0 is the value at the basepoint, so

$$c_0 = \ln 1 = 0.$$

Thus, the series

$$p(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (x-1)^k = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} + \dots$$

is a candidate for a series giving $p(x) = \ln x$ when $|x - 1| < 1$. What we actually know is that

$$p(x) = 0 \quad = \ln 1$$
$$p'(x) = \frac{dp}{dx} = \frac{1}{x} \quad (|x - 1| < 1).$$

But this means $p(x) - \ln x$ has derivative zero and hence is constant—and since $p(1) = \ln 1$, we have

$$p(x) = \ln x \quad 0 < x < 2$$

or the series representation

$$\begin{aligned}\ln x &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}(x-1)^k \\ &= (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \dots \quad 0 < x < 2.\end{aligned}$$

From this we can easily get also

$$\begin{aligned}\ln(x+1) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \\ &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k \quad |x| < 1,\end{aligned}$$

a series representation known to Nicolaus Mercator (1620-1687) in 1668. Note that this series also converges at $x = 1$ by the alternating series test—in fact, it is the alternating harmonic series. It follows from Abel's Theorem (Proposition 6.5.6) that

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots = \ln 2.$$

See Exercise 48 for some other series representations of $\ln 2$.

Similar arguments can be used to find power series for other functions, notably $\arctan x$ (see Exercise 37).

A third example is even more subtle, but gives us other new functions. Let us try to find a power series about $x = 0$ for

$$f(x) = e^x.$$

Formally, if we have

$$f(x) = \sum_{k=0}^{\infty} c_k x^k$$

then

$$f'(x_) = \sum_{k=1}^{\infty} k c_k x^{k-1} = \sum_{i=0}^{\infty} (i+1) c_{i+1} x^i.$$

But for $f(x) = e^x$ we have $f'(x) = e^x = f(x)$, so we expect the two series to agree; equating corresponding terms

$$(i+1)c_{i+1} = c_i \quad i = 0, 1, ...$$

gives us the recursive relation

$$c_{i+1} = \frac{c_i}{i+1}.$$

Now,

$$c_0 = f(0) = e^0 = 1$$

so we find

$$c_k = \frac{1}{k!} \quad k = 1, ...$$

(and also, by convention, for $k = 0$).
Does this series converge to e^x, and if so, for which values of x? Note that the series

$$f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

has radius of convergence $R = \infty$ (Exercise 42), so converges for *all* x to a function f for which

$$f'(x) = f(x)\,.$$

If we consider

$$g(x) = \ln f(x)$$

then

$$g'(x) = \frac{f'(x)}{f(x)} = 1$$

for all x, and hence

$$g(x) = x + C$$

for some constant C. But $f(0) = c_0 = 1$, so

$$g(0) = \ln 1 = 0 = 0 + C$$

forcing $C = 0$; we thus have

$$g(x) = x$$

for all x, giving us the series representation[12]

$$e^x = f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad -\infty < x < \infty. \tag{6.4}$$

In particular, substituting $x = 1$, we see that

$$\sum_{k=0}^{\infty} \frac{1}{k!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots = e^1 = e. \tag{6.5}$$

Similar arguments can be used to determine the series representations

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \ldots \tag{6.6}$$

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots \tag{6.7}$$

for all x (Exercise 35).

Finally, we consider multiplication of series. Multiplication of polynomials is already complicated, since (by the distributive law) the product of two sums is obtained by adding up all possible products of a term from the first sum with a term from the second.

For example, the product of two quadratic polynomials

$$p_2(x) = a_0 + a_1 x + a_2 x^2$$
$$q_2(x) = b_0 + b_1 x + b_2 x^2$$

is

$$p_2(x)\, q_2(x) = (a_0 + a_1 x + a_2 x^2)(b_0 + b_1 x + b_2 x^2)$$

which can be written, using the distributive law, as

$$\begin{gathered}(a_0 b_0) + (a_1 x)(b_0) + (a_2 x^2)(b_0) \\ +(a_0)(b_1 x) + (a_1 x)(b_1 x) + (a_2 x^2)(b_1 x) \\ +(a_0)(b_2 x^2) + (a_1 x)(b_2 x^2) + (a_2 x^2)(b_2 x^2)\end{gathered}$$

[12]This could also be proved using Taylor's theorem—see Exercise 38.

or, grouping together the terms with like powers of x,

$$\begin{aligned}\bigl(a_0b_0\bigr) + \bigl(a_0b_1 + a_1b_0\bigr)x + \bigl(a_0b_2 + a_1b_1 + a_2b_0\bigr)x^2 \\ +\bigl(a_1b_2 + a_2b_1\bigr)x^3 + \bigl(a_2b_2\bigr)x^4\end{aligned}$$

Notice that in this product, the coefficient of any power of x, say x^k, is the sum of all possible products a_ib_j for which $i + j = k$.
The same pattern holds for the product of two polynomials of higher degree: if

$$\begin{aligned}p_N(x) &= a_0 + a_1x + \ldots + a_Nx^N = \sum_{i=0}^{\infty} a_ix^i \\ q_N(x) &= b_0 + b_1x + \ldots + b_Nx^N = \sum_{j=0}^{\infty} b_jx^j\end{aligned}$$

then $p_N(x)\, q_N(x)$ is a polynomial of degree $2N$ in which the coefficient of x^k is the sum of all possible products a_ib_j with $i + j = k$. This has two useful consequences:

Observation 1: the coefficient of x^k in $p_N(x)\, q_N(x)$ does not involve any a_i or b_j with index greater than k;

Observation 2: for $k \leq N$, the coefficient of x^k in the product is

$$c_k = a_0b_k + a_1b_{k-1} + \ldots + a_kb_0 = \sum_{i=0}^{k} a_ib_{k-i}.$$

Now, suppose $p(x)$ and $q(x)$ are power series with partial sums $p_N(x)$ and $q_N(x)$, respectively, which both converge at x:

$$\begin{aligned}p(x) &= \sum_{i=0}^{\infty} a_ix^i = \lim_{N\to\infty} \left(p_N(x) = \sum_{i=0}^{N} a_ix^i \right) \\ q(x) &= \sum_{j=0}^{\infty} b_jx^j = \lim_{N\to\infty} \left(q_N(x) = \sum_{j=0}^{N} b_jx^j \right).\end{aligned}$$

Then

$$p(x)\, q(x) = \lim_{N\to\infty} p_N(x)\, q_N(x)$$

by Theorem 2.4.1. Now for $N = 0, 1, ...$, let $r_N(x)$ be the polynomial of degree N obtained by truncating the product $p_N(x)\, q_N(X)$ after the x^N term:

$$r_N(x) = \sum_{k=0}^{N} c_k x^k = c_0 + c_1 x + c_2 x^2 + \ldots + c_N x^N$$

where

$$c_k = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0 = \sum_{i=0}^{k} a_i b_{k-i}, \quad k = 0, \ldots, N.$$

Note that $r_N(x)$ forms the N^{th} partial sum of a new power series, with coefficients given by the formula above:

$$r(x) = \sum_{k=0}^{\infty} c_k x^k$$

where

$$c_k = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0 = \sum_{i=0}^{k} a_i b_{k-i}, \quad k = 0, 1, ...$$

We want to show that $r(x) = p(x)\, q(x)$:

Proposition 6.5.8. *Suppose*

$$p(x) = \sum_{i=0}^{\infty} a_i x^i, \quad q(x) = \sum_{j=0}^{\infty} b_j x^j$$

are both absolutely convergent at $x = x_0$*. Let*

$$r(x) = \sum_{k=0}^{\infty} c_k x^k$$

be the power series with coefficients

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0, \quad k = 0, 1, ...$$

Then $r(x)$ *converges absolutely at* $x = x_0$*, and*

$$r(x_0) = p(x_0)\, q(x_0)\,.$$

Proof. First, we need to show that the series $r(x)$ converges absolutely at $x = x_0$.
Denote by $p_N(x)$, $q_N(x)$ and $r_N(x)$ respectively the N^{th} partial sums of $p(x)$, $q(x)$, and $r(x)$; also, we will use the notation $|p_N|\,|x|$ (*resp.* $|q_N|\,|x|$), $|r_N|\,|x|$) to denote the partial sums obtained by applying the absolute convergence test:

$$|p_N|\,|x| = \sum_{i=0}^{N} |a_i|\,|x|^i$$

$$|q_N|\,|x| = \sum_{j=0}^{N} |b_j|\,|x|^j$$

$$|r_N|\,|x| = \sum_{k=0}^{N} |c_k|\,|x|^k\,.$$

Finally, let

$$\bar{r}_N(x) = \sum_{k=0}^{N} \bar{c}_k\,|x|^k\,,$$

where

$$\bar{c}_k = |a_0 b_k| + |a_1 b_{k-1}| + \ldots + |a_k b_0|\,.$$

It is easy to check the following inequalities:

$$|r_N|\,|x| \leq \bar{r}_N(x) \leq \big(|p_N|\,|x|\big)\cdot\big(|q_N|\,|x|\big) \leq \bar{r}_{2N}(x) :$$

the first inequality follows from $|c_k| \leq \bar{c}_k$, the second from the fact that $\bar{r}_N(x)$ is a truncation of $\big(|p_N|\,|x|\big)\cdot\big(|q_N|\,|x|\big)$ (and all terms being truncated are non-negative), while the last follows from observation 1 above applied with $k = N$. Note that each of the sequences of partial sums $|r_N|\,|x|$, $\bar{r}_N(x)$, and $\big(|p_N|\,|x|\big)\cdot\big(|q_N|\,|x|\big)$ is non-decreasing, and absolute convergence of $p(x)$ and $q(x)$ at $x = x_0$ means

$$\lim\,(|p_N|\,|x_0|)\,(|q_N|\,|x_0|) = \lim\,(|p_N|\,|x_0|\,|q_N|\,|x_0|) = L < \infty.$$

But since the sequences $\{|r_N|\,|x_0|\}$ and $\{\bar{r}_N(x_0)\}$ are intertwined with $\{|p_N|\,|x|\,|q_N|\,|x|\}$, we have also

$$|r_N|\,|x_0| \to L$$
$$\bar{r}_N(x_0) \to L.$$

In particular, convergence of $|r_N||x_0|$ means $r(x)$ converges absolutely at $x = x_0$.
Now, we need to show $r(x)$ has the right sum. We saw that $r_N(x)$ is a truncation of $p_N(x)\,q_N(x)$, and the terms that are missing all have the form $a_i b_j x^k$ with $i + j = k$, $N < k \leq 2N$. But $r_{2N}(x)$ contains *all* such terms, and we have

$$\begin{aligned} |p_N(x)\,q_N(x) - r_N(x)| &\leq \sum_{k=N+1}^{2N} \bar{c}_k\,|x|^k \\ &\leq \sum_{k=N+1}^{\infty} \bar{c}_k\,|x|^k\,. \end{aligned}$$

For $x = x_0$, this last quantity is the "tail" of the convergent series $\sum_{k=0}^{\infty} \bar{c}_k\,|x_0|^k$, so

$$\lim_{N\to\infty} \sum_{k=N+1}^{\infty} \bar{c}_k\,|x_0|^k = 0.$$

Thus,

$$|p_N(x_0)\,q_N(x_0) - r_N(x_0)| \to 0$$

and therefore

$$r(x_0) = \lim_{N\to\infty} r_N(x_0) = \lim_{N\to\infty} p_N(x_0)\,q_N(x_0) = p(x_0)\,q(x_0)\,.$$

□

Although we worked in the preceding with series in powers of x (*i.e.*, basepoint 0), the arguments work as well if we move the basepoint of all three series to the new position a. We have then the following more general statement.

Theorem 6.5.9 (Product Series)**.** *Suppose*

$$\begin{aligned} p(x) &= \sum_{i=0}^{\infty} a_i (x-a)^i \\ q(x) &= \sum_{j=0}^{\infty} b_j (x-a)^j \end{aligned}$$

are power series about $x = a$ with radii of convergence R_p and R_q, respectively. Let

$$R = \min\{R_p, R_q\}$$

and set

$$r(x) = \sum_{k=0}^{\infty} c_k (x-a)^k$$

where

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0, \quad k = 0, 1, ...$$

Then $r(x)$ *converges absolutely at every* x *with* $|x-a| < R$*, and*

$$r(x) = p(x)\, q(x) \quad |x-a| < R.$$

As an example, recall that

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \quad |x| < 1.$$

If we take $p(x) = q(x)$ in Proposition 6.5.8, the "product series" is

$$r(x) = \sum_{k=0}^{\infty} c_k x^k$$

where for $k = 0, 1, ...$

$$c_k = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0 = 1 \cdot 1 + 1 \cdot 1 + \ldots + 1 \cdot 1 = k+1.$$

That is,

$$r(x) = \sum_{k=0}^{\infty} (k+1) x^k = \frac{1}{(1-x)^2} \quad |x| < 1.$$

This of course was what we found earlier using differentiation.
As a second example, consider the two series about $x = 1$

$$p(x) = \sum_{i=0}^{\infty} 2^i (x-1)^i$$
$$q(x) = \sum_{j=0}^{\infty} (-1)^j (x-1)^j$$

Both are geometric; we have

$$p(x) = \frac{1}{1-(2x-2)} = \frac{1}{3-2x} \quad |x-1| < \frac{1}{2}$$
$$q(x) = \frac{1}{1-(1-x)} = \frac{1}{x} \quad |x-1| < 1.$$

The product series

$$r(x) = \sum_{k=0}^{\infty} c_k (x-1)^k$$

has coefficients

$$\begin{aligned} c_k &= (2^0)(-1)^k + (2^1)(-1)^{k-1} + \ldots + (2^{k-1})(-1)^0 \\ &= (-1)^k \left(1 - 2 + 4 - 8 + \ldots + (-1)^k (2^k)\right). \end{aligned}$$

This has no easy simplification: the first few terms are

$$1, 1, 3, 5, 11, 21, 43, \ldots$$

and Theorem 6.5.9 tells us that for $|x-1| < \frac{1}{2}$,

$$\begin{aligned} r(x) = \sum_{k=0}^{\infty} c_k (x-1)^k &= 1 + (x-1) + 3(x-1)^2 + 5(x-1)^3 + \ldots \\ &= \left(\frac{1}{3-2x}\right)\left(\frac{1}{x}\right) = \frac{1}{3x - 2x^2}. \end{aligned}$$

For a final example, we compute the first few terms of a series representation for

$$f(x) = e^x \cos x.$$

Recall that we already have the series representations

$$\begin{aligned} e^x &= \sum_{i=0}^{\infty} \frac{x^i}{i!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \\ \cos x &= \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m}}{(2m)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots \end{aligned}$$

both of which hold for all x. The coefficient of x^i in the first series is

$$a_i = \frac{1}{i!}$$

so the first few coefficients are

$$a_0 = 1, \quad a_1 = 1, \quad a_2 = \frac{1}{2}, \quad a_3 = \frac{1}{6}, \ldots$$

In the second series, the coefficient of x^j is given by

$$b_j = \begin{cases} 0 & \text{if } j \text{ is even} \\ \frac{(-1)^m}{(2m)!} = \frac{(-1)^{j/2}}{j!} & \text{if } j = 2m \text{ is even} \end{cases}$$

that is

$$b_0 = 1, \quad b_1 = 0, \quad b_2 = -\frac{1}{2}, \quad b_3 = 0, \ldots$$

While it is possible to work out a general formula for the coefficient c_k of x^k in the product series, we will simply calculate the first few values directly:

$$\begin{aligned}
c_0 &= a_0 b_0 &&= (1)(1) &&= 1 \\
c_1 &= a_0 b_1 + a_1 b_0 &&= (1)(0) + (1)(1) &&= 1 \\
c_2 &= a_0 b_2 + a_1 b_1 + a_2 b_0 &&= (1)(-\frac{1}{2}) + (1)(0) + (\frac{1}{2})(1) &&= 0 \\
c_3 &= a_0 b_3 + a_1 b_2 + a_2 b_1 + a_3 b_0 &&= (1)(0) + (1)(-\frac{1}{2}) + (\frac{1}{2})(0) + (\frac{1}{6})(1) &&= -\frac{1}{3}
\end{aligned}$$

so we have the series representation

$$e^x \cos x = 1 + x - \frac{1}{3}x^3 + \ldots.$$

In Table 6.1 we give a few useful basic power series established in this section, which are worth memorizing.

Table 6.1: Basic Power Series

$\frac{1}{1-x}$	$= \sum_{k=0}^{\infty} x^k,$	$-1 \leq x < 1$
e^x	$= \sum_{k=}^{\infty} \frac{x^k}{k!}$	$-\infty < x < \infty$
$\ln(x+1)$	$= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k,$	$-1 < x \leq 1$
$\cos x$	$= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!},$	$-\infty < x < \infty$
$\sin x$	$= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!},$	$-\infty < x < \infty$

Exercises for § 6.5

Answers to Exercises 1-33 (odd only), 36 are given in Appendix B.

Practice problems:

In problems 1-17, use the known power series for basic functions, together with the manipulation rules given in this section, to express the sum of each power series below as an explicit function of x. Give the center and radius of convergence for each series (*i.e.*, you do not need to determine convergence at the endpoints of the interval of convergence).

1. $\sum_{k=0}^{\infty} x^{k+2}$
2. $\sum_{k=0}^{\infty} x^{3k}$
3. $\sum_{k=1}^{\infty} (x^k + x^{k-1})$
4. $\sum_{k=0}^{\infty} x^{2k+1}$
5. $\sum_{k=1}^{\infty} (-1)^{k-1} x^{2k+1}$
6. $\sum_{k=1}^{\infty} (2 - 3^k) x^k$
7. $\sum_{k=0}^{\infty} \frac{(x-1)^k}{k!}$
8. $\sum_{k=1}^{\infty} \frac{(-1)^k (x-2)^k}{k}$
9. $\sum_{k=0}^{\infty} (4-x)^k$
10. $\sum_{k=1}^{\infty} \frac{x^{k+1}}{2^k}$
11. $\sum_{k=3}^{\infty} k x^k$
12. $\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}$
13. $\sum_{k=1}^{\infty} \frac{3^k}{k} x^k$
14. $\sum_{k=2}^{\infty} \frac{(-1)^k x^k}{k(k-1)}$
15. $\sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k)!}$
16. $1 + x - x^2 - x^3 + x^4 + x^5 - ...$ (signs switch every two terms)
17. $1 + x + x^2 - x^3 + x^4 + x^5 + x^6 - x^7 + x^8 + ...$ (every fourth term has a minus) (*Hint:* Express as a difference of two series, each with all terms added.)

In problems 18-32 find a series in powers of x expressing the given function; give the radius of convergence.

18. $\frac{x}{1-x}$
19. $\frac{x+1}{x-1}$
20. $\frac{2x}{3x+1}$
21. $\frac{1}{1-x^2}$
22. $\frac{1}{1+4x^2}$
23. $\frac{2x}{x^2+1}$
24. $\ln(2x+2)$
25. $\ln(x+2)+1$
26. $x\ln(x+1)$
27. e^{-3x}
28. e^{-x^2}
29. xe^{2x}

30. $\sin 3x$ 31. $x^2 \sin x$ 32. $x \cos x^2$

33. Find the first five nonzero terms for the power series about $x = 0$ of the function $f(x) = e^{2x} \cos 3x$.

34. Find the first three nonzero terms for the power series about $x = 0$ of the function $f(x) = \int_0^x \sin t^2 \, dt$.

35. Use the following information to find the power series about $x = 0$ for $\sin x$ and $\cos x$, following the kind of argument in the derivation of the series for e^x on Pages 534-536:

- $\sin 0 = 0$ and $\cos 0 = 1$
- $\frac{d}{dx} [\sin x] = \cos x$ and $\frac{d}{dx} [\cos x] = \sin x$
- For each of these functions, $f''(x) = -f(x)$.

36. (a) Substitute x^2 for x in the power series for $\sin x$ to get a series for $\sin x^2$.

    (b) Use Lemma 6.5.7 to find a power series for an antiderivative of $\sin x^2$. For which values of x is this a valid representation?

    (c) Write down the first three nonzero terms in a series giving the number

$$\int_0^1 \sin x^2 \, dx.$$

Note that the series appears to converge extremely quickly.

37. Find the power series for $\arctan x$ about $x = 0$ as follows:

    (a) Use geometric series to find the power series about $x = 0$ for $\dfrac{1}{x^2 + 1}$.

    (b) Formally integrate this series to get the series for $\arctan x$.

    (c) What is the interval of convergence for these series?

38. Show that $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ for all $x \in \mathbb{R}$ directly, by first calculating the coefficients of the Taylor series at $x = 0$ and then using Theorem 6.1.7 (Taylor's Theorem, Lagrange Form) to prove convergence to e^x at every $x \in \mathbb{R}$.

Theory problems:

39. We give here an example due to Cauchy (1821) [30, p. 196] which shows that the formal product of two convergent series need not itself converge.

    (a) Show that the series

    $$1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{5}} - \cdots$$

    converges.

    (b) Show that for $0 \leq x \leq n$

    $$(n + 1 - x)(x + 1) \leq \left(1 + \frac{n}{2}\right)^2$$

    (*Hint:* Find the maximum of the left side.)

    (c) Use this to show that

    $$\sum_{j=0}^{n} \frac{1}{\sqrt{n+1-j}\sqrt{j+1}} \geq \frac{2n+2}{n+2}.$$

    (*Hint:* Bound each term of the sum below.)

    (d) Use the Divergence Test to show that the product of our series with itself diverges.

40. Mimic the proof of the first statement in Remark 6.5.1 to prove the second statement, that for $n = 1, 2, ...,$ $\sum_{k=0}^{\infty} c_k(x-a)^k$ and $\sum_{k=0}^{\infty} c_k(x-a)^{k+n}$ have the same interval of convergence and

    $$\sum_{k=0}^{\infty} c_k(x-a)^{k+n} = (x-a)^n \sum_{k=0}^{\infty} c_k(x-a)^k.$$

41. Use the Chain Rule and basic properties of power series to show that if $p(x) = \sum_{k=0}^{\infty} c_k(x-a)^k$ has radius of convergence $R > 0$ and $q(x) = p(x+a)$ for all x in the interval of convergence of $p(x)$, then

(a) $q(x) = \sum_{k=0}^{\infty} c_k x^k$ for all x in the interval $(-R, R)$

(b) $q'(x) = p'(x+a)$ for every such x.

(c) Conclude that the general case of Lemma 6.5.5 follows from the special case proved on Pages 528-531, and that Theorem 6.5.3 follows.

42. Show that the series $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ converges for all x.

Challenge problem:

43. **Irrationality of *e*:** Show that e is irrational, as follows:

(a) Show that

$$\frac{1}{e} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}.$$

(b) Use Theorem 6.1.7 to show that for each $n = 2, 3, ...,$

$$\left| \frac{1}{e} - \sum_{k=0}^{n-1} \frac{(-1)^k}{k!} \right| < \frac{1}{n!}.$$

(c) Now, suppose e is rational. Find an integer N such that for all $n \geq N$

$$(n-1)!\frac{1}{e} \text{ is an integer.}$$

(d) Show that for all n,

$$(n-1)! \sum_{k=0}^{n-1} \frac{(-1)^k}{k!} \text{ is an integer.}$$

(e) Conclude that for all $n \geq N$

$$(n-1)! \left| \frac{1}{e} - \sum_{k=0}^{n-1} \frac{(-1)^k}{k!} \right| < \frac{(n-1)!}{n!} = \frac{1}{n}$$

is an integer, which is an absurdity.

History notes:

44. **Newton's Binomial Series:**

 (a) Show that for any two nonzero integers m and n, the Taylor series about $x = 0$ of the function $(1+x)^{m/n}$ is given by

$$(1+x)^{m/n} = 1 + \frac{m}{n}x + \frac{\frac{m}{n}\left(\frac{m}{n}-1\right)}{2!}x^2 + \frac{\frac{m}{n}\left(\frac{m}{n}-1\right)\left(\frac{m}{n}-2\right)}{3!}x^3 + ...$$

 (b) Verify that if $n = 1$ (that is, the function is $(x+1)^m$), the terms of this series involving powers of x above m are zero, and this is just the Binomial Theorem—that is, the expansion of an integer power of $(x+1)$ in powers of x.

 (c) Verify that the series agrees with the form given by Newton in the 1660's, in the recursive form [18, pp. 6-9], [51, pp. 285-7]:

$$(P+PQ)^{m/n} = P^{m/n} + \frac{m}{n}AQ + \frac{m-n}{2n}BQ + \frac{m-2n}{3n}CQ + \frac{m-3n}{4n}DQ + \text{etc.}$$

 where each of the coefficients $A, B, C, D, ...$ denotes the whole immediately preceding term.

 (This formula was rediscovered independently by James Gregory (1638-1675) a few years later.)

45. **Wallis' representation of π:** In his *Arithmetica infinitorum*, Wallis gave the following representation for π as an infinite sum:

$$\frac{\pi}{2} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdot \cdots$$

 by an interpolation method. Here is a "modern" derivation [20, pp. 169-70]:

 (a) Let

$$I_n := \int_0^{\pi/2} \sin^n x\, dx.$$

Show that

$$I_0 = \frac{\pi}{2}$$
$$I_1 = 1$$
$$I_n = \frac{n-1}{n} I_{n-2} \text{ for } n \geq 2.$$

(*Hint:* Integration by parts for the last one.)

(b) Conclude that

$$I_{2n} = \frac{\pi}{2} \cdot \frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n}$$
$$I_{2n+1} = \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdots \frac{2n}{2n+1}$$

and so

$$\frac{\pi}{2} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \cdots \frac{2n}{2n-1} \cdot \frac{2n}{2n+1} \cdot \frac{I_{2n}}{I_{2n+1}}.$$

(c) Show that

$$1 < \frac{I_{2n}}{I_{2n+1}} \leq \frac{I_{2n-1}}{I_{2n+1}} = 1 + \frac{1}{2n}.$$

(*Hint:* Compare the integrands in I_k and I_{k+1} for arbitrary k.)

(d) Derive Wallis' product from this.

46. **Leibniz' Series:**[13] [20, pp. 247-8] Leibniz derived the summation formula

$$\frac{\pi}{4} = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + ... \tag{6.8}$$

which can be stated in purely geometric terms as saying that the sum of this series is the ratio between the area of a circle and that of a circumscribed square. Leibniz regarded this as one of his finest achievements. He derived Equation (6.8) using his "Transmutation Theorem" (Exercise 4 in § 5.4),

$$\int_a^b y\,dx = \frac{1}{2}\left((xy)\Big|_a^b + \int_a^b z\,dx \right)$$

[13] This formula was actually discovered somewhat earlier by James Gregory (1638-1675) and, apparently, even earlier by the Indian mathematician Kerala Gargya Nilakantha (*ca.* 1500). See [45].

where

$$z = y - x\frac{dy}{dx}$$

as follows:

(a) Consider the semicircle of radius 1 over the x-axis tangent to the y-axis, with equation

$$y = \sqrt{2x - x^2}.$$

Show that (assuming a circle with radius 1 has area π)

$$\int_0^1 y\,dx = \frac{\pi}{4}.$$

(b) Show that the quantity z in the Transmutation Theorem has the form

$$z = \sqrt{\frac{x}{2-x}}$$

(*Hint:* First note that $\frac{dy}{dx} = \frac{1-x}{y}$.)

(c) Invert this last equation to get

$$x = \frac{2z^2}{1+z^2}.$$

(d) Show that

$$\int_0^1 z\,dx = 1 - \int_0^1 x\,dz$$

by dividing the square $0 \le x \le 1$, $0 \le z \le 1$ into two regions using the graph of z versus x, and slicing one of these regions vertically and the other horizontally, in the manner of § 5.8.

(e) Now, substitute in the Transmutation Theorem and simplify to get

$$\frac{\pi}{4} = \int_0^1 y\,dx = 1 - \int_0^1 \frac{z^2}{1+z^2}dz.$$

(f) Use geometric series to show that the integral on the right is

$$\int_0^1 (z^2 - z^4 + z^6 - ...)\,dz.$$

(g) Ignoring for a moment the issue of convergence, note that the formal integration and substitution gives

$$\int_0^1 (z^2 - z^4 + z^6 - ...)dz = \frac{1}{3} - \frac{1}{5} + \frac{1}{7} - .... \qquad (6.9)$$

(h) Show how this gives the summation formula (6.8).

Leibniz did not worry about the validity of the integration formula (6.9), but since the geometric series for the integrand diverges at $z = 1$, we *should* worry. Here is a rigorous way to justify the derivation of Equation (6.9) from term-by-term integration [47, pp. 254-5]:

(i) The difference between the function $\frac{z^2}{1+z^2}$ and a partial sum of the series with an even number of terms is given by

$$z^2 - z^4 + z^6 - z^8 - ... + z^{4n} - \frac{z^2}{1+z^2} = \frac{z^{4n+2}}{1+z^2}$$

(Verify this by dividing $z^2 - z^{4n+2}$ by $1 + z^2$.) Now integrate both sides, to get

$$\frac{1}{3} - \frac{1}{5} + \frac{1}{7} - \frac{1}{9} + ... + \frac{1}{4n+1} - \int_0^1 \frac{z^2}{1+z^2}dz = \int_0^1 \frac{z^{4n+2}}{1+z^2}dz.$$

But since $1 + z^2 \geq 1$, we can write

$$0 < \int_0^1 \frac{z^{4n+2}}{1+z^2}dz \leq \int_0^1 z^{4n+2}\, dz = \frac{1}{4n+3}$$

and since the last fraction goes to zero with n, the partial sums of our series converge to the given integral.

47. **Abel's Theorem** In [1], Abel was concerned with the Binomial Series (first investigated by Newton–see Exercise 44), but his method generalizes to give Proposition 6.5.6. The proof we sketch below is based on [30, pp. 250-1].

We already know by Theorem 6.5.3 that $p(x)$ is differentiable, hence continuous, on the open interval $(a - R, a + R)$. Thus all we need to show is that, if $R < \infty$ and $p(x)$ also converges at $x = a + R$ (*resp.* $x = a - R$), then $p(x) \to p(a + R)$ as $x \to (a + R)^-$ (*resp.* $p(x) \to p(a - R)$ as $x \to (a - R)^+$).

(a) Use a substitution to show that we can assume without loss of generality that $a = 0$ and $a + R = 1$—that is, that it suffices to show that

Claim: *If*

$$f(x) = \sum_{k=0}^{\infty} c_k x^k$$

has radius of convergence $R = 1$*, and also converges at* $x = 1$*, then*

$$\lim_{x \to 1^-} f(x) = f(1) = \sum_{k=0}^{\infty} c_k.$$

(b) **Abel's Partial Summation:** Denote the n^{th} partial sum of the series $f(x)$ by

$$f_n(x) := \sum_{k=0}^{n} c_k x^k.$$

We will need to estimate $|f_{n+k}(x) - f_n(x)|$ for $k > 0$.
Given n and $j > 0$, let

$$C_{n,j} := c_{n+1} + \cdots + c_{n+j}.$$

Show that for each $x \in [0, 1]$,

$$\begin{aligned} f_{n+k}(x) - f_n(x) &= c_{n+1}x^{n+1} + c_{n+2}x^{n+2} + \cdots + c_{n+k}x^{n+k} \\ &= C_{n,k}x^{n+k} + C_{n,k-1}\left(x^{n+k-1} - x^{n+k}\right) \\ &\quad + C_{n,k-2}\left(x^{n+k-2} - x^{n+k-1}\right) + \cdots + C_{n,1}\left(x^{n+1} - x^{n+2}\right). \end{aligned}$$

(c) **Show** that given $\varepsilon > 0$ we can find N such that whenever $n \geq N$ and $k > 0$

$$|C_{n,k}| < \frac{\varepsilon}{3}.$$

(d) As a consequence, **show** that for $0 \leq x \leq 1$, $n \geq N$ and $k > 0$

$$|f_{n+k}(x) - f_n(x)| < \frac{\varepsilon}{3}x^{n+1} \leq \frac{\varepsilon}{3}$$

(*hint:* combine the the triangle inequality and the inequality in (c) to get a telescoping sum) and hence, for $0 \leq x \leq 1$

$$|f(x) - f_n(x)| \leq \frac{\varepsilon}{3}.$$

(e) Pick $n > N$. Since $f_n(x)$ is a polynomial, we can find $\delta > 0$ such that $1 - \delta < x < 1$ guarantees

$$|f_n(x) - f_n(1)| < \frac{\varepsilon}{3}.$$

(f) Now suppose $1 - \delta < x < 1$. Then identify reasons why

$$|f(1) - f_n(1)| \leq \frac{\varepsilon}{3}$$
$$|f_n(1) - f_n(x)| < \frac{\varepsilon}{3}$$

and

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{3}$$

and show how this proves

$$|f(1) - f(x)| < \varepsilon.$$

48. **Some series for ln 2** [30]: Here are some other series which can be shown to converge to $\ln 2$:

(a) Use the MacLaurin series for $f(x) = \ln(1 - x)$ to show that

$$\ln 2 = \sum_{k=1}^{\infty} \frac{1}{k \cdot 2^k} = \frac{1}{2} + \frac{1}{2 \cdot 2^2} + \frac{1}{3 \cdot 2^3} + ...$$

(b) Subtract the series for $\ln(1 - x)$ from that for $\ln(1 + x)$ to obtain the following series, due to James Gregory (1638-1675):

$$\ln\left(\frac{1+x}{1-x}\right) = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + ...\right).$$

(c) Substitute $x = \frac{1}{3}$ into Gregory's series to obtain

$$\ln 2 = 2\left(\frac{1}{3} + \frac{1}{3 \cdot 3^3} + \frac{1}{5 \cdot 3^5} + ...\right).$$

Can you justify that this is correct?

(d) Newton (1671) used the identity

$$2 = \frac{3}{2} \cdot \frac{4}{3}$$

to write

$$\ln 2 = \ln\left(\frac{1+\frac{1}{5}}{1-\frac{1}{5}}\right) + \ln\left(\frac{1+\frac{1}{7}}{1-\frac{1}{7}}\right).$$

Use this together with Gregory's series to find a series converging to $\ln 2$.

49. **Mercator's series:** Nicolaus Mercator (1620-1687) in his *Logarithmotechnica* (1668) calculated extensive tables of logarithms, and in the third part gave his series for the logarithm (which agrees with the MacLaurin series):

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} - \dots$$

In his time, the function $\ln b$ was defined as the area under the hyperbola $xy = 1$ between $x = 1$ and $x = b$, and he derived this series by a Cavalieri-style calculation of this area. In his review of Mercator's book, John Wallis (1616-1703) gave more details [20, pp. 162-3]. We use the terminology of Chapter 5.

(a) Consider the partition $\mathcal{P}_n$ of $[1, x]$ into n equal parts of width $h = \frac{x}{n}$; the circumscribed rectangles over the component intervals have heights

$$1,\ \frac{1}{1+h},\ \frac{1}{1+2h}, \dots, \frac{1}{1+(n-1)h}.$$

Each of these numbers is the sum of a geometric series, and substituting into the upper sum we get an upper estimate for the area

$$\begin{aligned}\mathcal{U}(\mathcal{P}, f) &= h + \sum_{j=1}^{n-1} \frac{h}{1+jh} \\ &= h + h\left(\sum_{k=0}^{\infty}(-1)^k h^k\right) + h\left(\sum_{k=0}^{\infty}(-1)^k (2h)^k\right) \\ &\qquad + \dots + h\left(\sum_{k=0}^{\infty}(-1)^k ((n-1)h)^k\right).\end{aligned}$$

(b) Collect terms involving the same powers of h to get

$$\begin{aligned}\mathcal{U}(\mathcal{P}, f) = nh - h[h + 2h + ... + (n-1)h] \\ + h[h^2 + (2h)^2 + ... + ((n-1)h)^2] + \\ ... + (-1)^k h[h^k + (2h)^k + ... + ((n-1)h)^k] + ...\end{aligned}$$

and substitute $h = \dfrac{x}{n}$ to get

$$\mathcal{U}(\mathcal{P}, f) = x - \frac{x^2}{n^2}\left(\sum_{i=1}^{n-1} i\right) + \frac{x^3}{n^3}\left(\sum_{i=1}^{n-1} i^2\right) + ... + (-1)^k \frac{x^{k+1}}{n^{k+1}}\left(\sum_{i=1}^{n-1} i^k\right) + ...$$

(c) Use the formula, stated by Wallis in his *Arithmetica infinitorum* (1656) on the basis of calculation up to $n = 10$, that

$$\lim_{n\to\infty} \frac{1^k + 2^k + ... + n^k}{n^{k+1}} = \frac{1}{k+1}$$

to take the limit of each term above as $n \to \infty$ to conclude

$$\lim_{n\to\infty} \mathcal{U}(\mathcal{P}, f) = x - \frac{x^2}{2} + \frac{x^3}{3} - ...$$

as required.

(d) Where are the gaps in the argument, from our modern point of view?

50. **Euler on e:** In his *Introductio in Analysin Infinitorum* (1748), Euler introduced the exponential function and defined the logarithm as the inverse of the exponential function (in contrast to the earlier definitions of the logarithm as an area—see § 3.6). He then found series expansions for both. Note the abandon with which he ignores some very subtle difficulties (but comes to the correct conclusions). We follow the notation of [20, pp. 272-4].

(a) Given that $a^0 = 1$, write $a^\varepsilon = 1 + k\varepsilon$ for an infinitely small number ε. Now given x, consider the infinitely large number $N = x/\varepsilon$ and write

$$\begin{aligned}a^x = a^{N\varepsilon} &= (a^\varepsilon)^N \\ &= (1 + k\varepsilon)^N = \left(1 + \frac{kx}{N}\right)^N\end{aligned}$$

which can be expanded via Newton's binomial series (Exercise 44) as

$$a^x = 1 + \frac{N}{N}kx + \frac{1}{2!}\frac{N(N-1)}{N^2}k^2x^2 + \frac{1}{3!}\frac{N(N-1)(N-2)}{N^3} + \cdots$$

(b) Since N is infinitely large, every fraction of the form $\dfrac{N-n}{N}$ equals 1, so

$$a^x = 1 + \frac{kx}{1!} + \frac{k^2x^2}{2!} + \frac{k^3x^3}{3!} + \cdots$$

(c) Substituting $x = 1$, we have in particular the relationship between a and k

$$a = 1 + \frac{k}{1!} + \frac{k^2}{2!} + \frac{k^3}{3!} + \cdots$$

Define e as the value of a for which $k = 1$. This gives a series for e

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots$$

(d) In particular, setting $a = e$ in the equation two items back, we have

$$e^x = \left(1 + \frac{x}{N}\right)^N$$

which we can interpret as the instantaneous compounding formula. Of course, we also have the series

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

(e) Now for the logarithm. Set

$$1 + y = a^x = a^{N\varepsilon} = (1 + k\varepsilon)^N$$

or

$$\log_a(1 + y) = N\varepsilon.$$

(f) Then

$$1 + k\varepsilon = (1 + y)^{1/N}$$

or

$$\varepsilon = \frac{(1 + y)^{1/N} - 1}{k}$$

so

$$\log_a(1 + y) = N\varepsilon = \frac{N}{k}[(1 + y)^{1/N} - 1].$$

(g) Now replace a with e (and k with 1) and y with x to write

$$\ln(1+x) = N[(1+x)^{1/N} - 1].$$

(h) Now expand $(1+x)^{1/N}$ using Newton's binomial series (Exercise 44)

$$\ln(1+x) = x - \frac{1}{2!}\frac{N-1}{N}x^2 + \frac{1}{3!}\frac{(N-1)(2N-1)}{N^3}x^3 + \cdots$$

or, since N is infinitely large, the fractions involving N can be replaced with the ratio of the coefficients of N, giving

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots$$

6.6 Complex Numbers and Analytic Functions (Optional)

The *complex numbers* are formed by creating an "imaginary" square root of -1—that is, we posit a new number $\boldsymbol{i}$ satisfying

$$i^2 = -1 \tag{6.10}$$

and allow all the usual arithmetic involving this and real numbers; this leads to the collection of **complex numbers** (often denoted $\mathbb{C}$), or expressions of the form

$$z = a + bi$$

where $a, b \in \mathbb{R}$. The real numbers a (*resp.* b) are called the **real part** (*resp.* **imaginary part**) of z, and denoted

$$\Re(z) = a$$
$$\Im(z) = b.$$

Arithmetic among complex numbers proceeds as we would expect it to, using Equation (6.10) to reduce any power of i above the first. Two features of complex arithmetic which deserve comment are the change in the meaning of absolute value and the existence of conjugates.

The "size" of a complex number can be thought of this way: we associate the complex number $z = a + bi$ with the point in the plane whose Cartesian coordinates are (a, b) (this is called the **Argand diagram** for the complex numbers), and then denote by $|z|$ the distance from that point

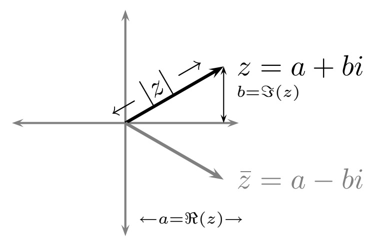


Figure 6.5: Argand diagram

to the origin (in keeping with the meaning of absolute value in $\mathbb{R}$):

$$|z| = |a + bi| = \sqrt{a^2 + b^2}. \tag{6.11}$$

In the complex context, this is referred to as the **modulus** of z. We think of the real numbers as those with imaginary part zero, and with this interpretation the modulus $|z|$ agrees with the absolute value when $z = a + 0i$.

The **conjugate** of a complex number $z = a + bi$ is its reflection across the real (*i.e.*, x-) axis, denoted $\bar{z}$:

$$\bar{z} = \overline{a + bi} := a - bi. \tag{6.12}$$

Note that $\bar{z} = z$ precisely if z is real ($\Im(z) = 0$), and the conjugate of $\bar{z}$ is z itself. Also, we note that for any complex number $z = a + bi$

$$z \cdot \bar{z} = (a + bi)(a - bi) = a^2 - abi + bai - b^2 i^2 = a^2 + b^2 = |z|^2 . \tag{6.13}$$

Now, polynomials—that is, functions whose definition involves only addition, subtraction and multiplication—immediately and naturally extend to the complex numbers; thus we can calculate for example $f(a + bi)$ for $f(x) = x^2$ as

$$f(a + bi) = (a + bi)(a + bi) = a^2 + 2abi + b^2 i^2 = (a^2 + b^2) + (2ab)i$$

while for $f(x) = x^2 - x$ we get

$$f(a + bi) = (a^2 + b^2) + 2abi - a - bi = (a^2 + b^2 - a) + (2ab - b)i.$$

This has important consequences: for example, Carl Friedrich Gauss (1777-1855) showed that with this interpretation *every* nonconstant polynomial function has (complex) points where it equals zero; for example, if $f(x) = x^2 + 1$, then $f(z) = 0$ for $z = \pm i$. This is known as the **Fundamental Theorem of Algebra**.

To extend this to rational functions, we need to be able to divide complex numbers. This is done with the help of Equation (6.13). Note that the modulus of a nonzero complex number is a positive real number, so we can always divide by $|z|$ if $z \neq 0$. But then, by Equation (6.13), for $z \neq 0$,

$$\frac{1}{z} = \frac{\bar{z}}{|z|}. \tag{6.14}$$

Using this, every rational function makes sense for (most) complex numbers. We say most, because if the denominator is nonconstant (*i.e.*, the function is not a polynomial) then it equals zero *somewhere* in the complex plane; a place where the denominator equals zero (and the numerator does not) is called a **pole** of the function.

Finally, since we can extend the calculation of polynomials to complex inputs, we can calculate all of the partial sums of any power series. To calculate the value of the series, there remains the question of convergence: in this context, let us agree that a sequence of *complex* numbers $z_n = a_n + b_n i$ converges to the complex number $z = a + bi$ if the sequence of real (*resp.* imaginary) parts converges:

$$z_n = a_n + b_n i \to z = a + bi \;\Leftrightarrow\; a_n \to a \text{ and } b_n \to b. \tag{6.15}$$

Then we can ask about the convergence of the partial sums of a power series at any *complex* number. It turns out (you can look back at the proof and convince yourself that everything carries over) that if the power series has radius of convergence $R > 0$ as a *real* power series, then in fact it converges at all *complex* points whose distance from the basepoint is $< R$, and diverges at all complex points at distance $> R$ from the base. (See Figure 6.6.) This helps explain the phenomenon that, for example, the rational function $f(x) = \frac{1}{x^2+1}$ has a power series (around $x = 0$) whose radius of convergence is only $R = 1$, even though $f(x)$ is defined for *all* real numbers x: we see that f has poles at $x = \pm i$, which are at distance 1 from the origin, so the series has no chance to converge further out.

Now using these observations, we see also that if a function is equal to the sum of its Taylor series about some point, then we can use this series to extend the function to all *complex* numbers whose distance from the basepoint is less than the radius of convergence of the series. A remarkable fact is the rigidity of complex functions. We can define the notion of the derivative of a function of a complex variable by just formally using the

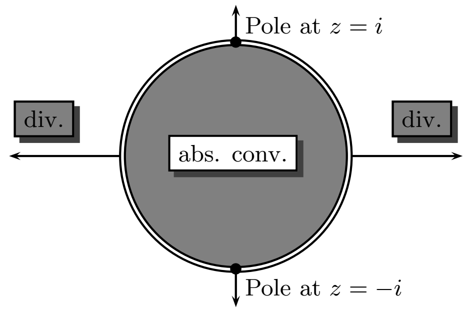


Figure 6.6: Domain of convergence for $\frac{1}{z^2+1} = \sum_{k=0}^{\infty}(-1)^k z^{2k}$

usual (real) definition, but allowing everything to be complex (and interpreting convergence of limits as convergence of real and imaginary parts). We have seen examples of differentiable functions of a *real* variable which are differentiable but whose Taylor series does not converge to the function (in fact which may not even have a Taylor series at some point because of the lack of higher derivatives). But in the *complex* context, this never happens: any function which is (just once) differentiable in the complex sense is in fact infinitely differentiable; moreover its Taylor series at that point has a positive radius of convergence, and converges to the function itself within that radius. An **analytic function** is a function whose Taylor series at each point converges with some positive radius of convergence to the function itself; thus a *complex*-differentiable function is automatically analytic.

In particular, let us examine the extension of the exponential function to complex numbers. We saw in § 6.5 (Equation (6.4) that the power series expansion

$$e^x = f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

converges to e^x for all real x—that is, the radius of convergence is $R = \infty$. It follows that we can define e^z for any *complex* number $z = a + bi$ by means of this series. Let us consider the case of a *pure imaginary* z, that is, $\Re(z) = a = 0$. The terms of the series then have the form

$$\frac{z^k}{k!} = \frac{(bi)^k}{k!} = i^k \frac{b^k}{k!},$$

which is *real* for k *even* and *pure imaginary* for k *odd*. Let us separate these cases by writing $k = 2n$ (*resp.* $k = 2n + 1$). Then we note that

$$i^{2n} = (i^2)^n = (-1)^n, \quad i^{2n+1} = (i^{2n}) \cdot i = (-1)^n i$$

so that we can separate the real terms of the series from the pure imaginary ones to get

$$\begin{aligned}\sum_{k=0}^{\infty} \frac{(bi)^k}{k!} &= \sum_{n=0}^{\infty} \frac{(bi)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(bi)^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{(b)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} (-1)^n i \frac{(b)^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{(b)^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{(b)^{2n+1}}{(2n+1)!}.\end{aligned}$$

But a quick comparison of this with Equation (6.6) shows that the two series in the last expression are, respectively, the power series for $\cos b$ and $\sin b$. Since they also converge to their respective functions with radius of convergence $R = \infty$, we can replace the series with these honest (real) trigonometric functions. Thus

$$e^{bi} = \cos b + i \sin b. \tag{6.16}$$

But we also know that $e^{a+b} = e^a + e^b$ for any two real numbers; it follows that this is true on the level of power series, and hence extends to complex exponents. Thus, we have **Euler's formula**

$$e^{a+bi} = e^a(\cos b + i \sin b) \tag{6.17}$$

which relates the exponential and trigonometric functions. (See Figure 6.7.) Furthermore, note that the right hand side of Equation (6.16)

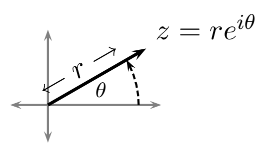


Figure 6.7: Euler's formula

is precisely the point we used to define the trigonometric functions $\sin \theta$ and $\cos \theta$ in § 3.1 (p. 94). In different language, this is the point with polar coordinates $r = 1$, $\theta = b$. By substituting $a = \ln r = \ln |z|$ in Equation (6.17), we see that the point with polar coordinates $(r, \theta))$ can be expressed as

$$z = re^{i\theta} = e^{\ln|z|+i\theta}.$$

In the complex context, the angle θ, which is determined only up to addition of integer multiples of 2π, is called the **argument** of z, denoted $\arg z$. This, incidentally, gives us a definition of the natural logarithm of a complex number,

$$\ln z = \ln |z| + i \arg z$$

but notice that it is ambiguous in the same way as inverse trigonometric functions are ambiguous, until we choose where the values are to lie. The situation in the complex plane is, however, more complex (no pun intended :-)) than on the line. To learn more, take a course in complex variables!

Exercises for § 6.6

Answers to Exercises 1acegikmo, 2-3ac, 4d, 5b are given in Appendix B.

Practice problems:

1. Use the results of this section to calculate each of the following quantities, expressing your answer in the form $a + bi$:

(a) $\overline{i^3}$	(b) i^4	(c) i^5
(d) $\dfrac{1}{i}$	(e) $(2 + 3i)(2 - 3i)$	(f) $\overline{(2 + 3i)^2}$
(g) $\dfrac{2 + 3i}{2 - 3i}$	(h) $e^{i\pi/4}$	(i) $e^{i\pi/3}$
(j) $e^{i\pi/2}$	(k) $e^{4\pi i/3}$	(l) $e^{1+i\pi/4}$
(m) $e^{1-i\pi/4}$	(n) $\left\|e^{1+i}\right\|$	(o) $\overline{e^{1+i}}$

2. Express each number below in the form $re^{i\theta}$:

(a) $1 + i$	(b) $i - 1$	(c) $\sqrt{3} + i$	(d) i

3. Find all (complex) solutions of each equation below. Express your answer(s) in the form $a + bi$, with a and b real.

(a) $z^2 + 1 = 0$	(b) $z^2 = 4i$	(c) $z^3 - 1 = 0$	(d) $z^4 + 1 = 0$

Theory problems:

4. (a) Show that a complex number z has modulus $|z| = 1$ precisely if it can be expressed in the form $z = e^{i\theta}$ for some real number θ.

   (b) If you know one value of θ in the above, what are the other possible values that give the same number z?

   (c) Use Euler's Formula to prove **DeMoivre's Theorem:** For any integer n,
   $$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta.$$

   (d) Use DeMoivre's Theorem to calculate a formula for $\sin^5\theta$ in terms of $\cos\theta$ and $\sin\theta$.

5. (a) Use Euler's Formula to derive the formal identities
   $$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$$
   $$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$
   $$i\tan\theta = \frac{e^{i\theta} - e^{-i\theta}}{e^{i\theta} + e^{i\theta}}.$$
   Note that these are actually identities relating power series, and so can be automatically extended to complex values of θ, thus defining "trigonometric functions" for complex "angles".

   (b) Use these identities to calculate $\sin i$, $\cos i$, and $\tan i$. (Some of your answers may be complex rather than real.)

Mathematical proofs, like diamonds, are hard as well as clear.

John Locke

A

The Rhetoric of Mathematics (Methods of Proof)

Unlike politics, where an unchallenged assertion quickly becomes a fact, mathematical assertions require proof. This is particularly true of statements involving very abstract situations: we must establish that the conclusion holds every time the hypotheses apply, and not just in the situation we may have had in mind when we formulated our assertion. Beginning math students often freeze at the sound of *show that...* or *prove the following*. Relax! There is no reason to panic. Many proofs are straightforward calculations or derivations. Others involve a fairly natural set of easy deductions. As you learn more, you will begin to understand certain standard tricks of the trade, and begin to incorporate them into your own arguments.

A mathematical proof is an argument for the truth of some assertion, subject to certain standards of *rigor*. What constitutes a rigorous proof has evolved over time. Babylonian and Egyptian mathematics was generally pragmatic and empirical.[1] The Greeks, starting with Thales of Miletus (*ca.* 625-*ca.* 547 BC) and Pythagoras of Samos (*ca.* 580-500 BC), gradually developed a more abstract mathematics built on a systematic approach to reasoning. The culmination of this development came in the

[1]See [11] for an extensive discussion of what we know about Babylonian and Egyptian mathematics, and how we know it.

Elements by Euclid of Alexandria (*ca.* 300 BC), which presented the mathematical results of the members of Plato's Academy and their contemporaries. The *Elements* served as the gold standard of rigor well into the nineteenth century. It was built on the idea that a mathematical theory starts from some basic shared assumptions (or **postulates**), and builds up to sophisticated results by the accumulation of individual arguments, organized as **lemmas** and **propositions**.

In the context of the present book, we take for granted such items as the rules of arithmetic or the algebra of polynomials. While it would be difficult to explicitly list *all* the assumptions we start from, one of the ingredients in learning a mathematical subject is to gain a sense of what can be taken for granted and what needs proof.

In the early development of the calculus, the geometric and logical language of Euclid, along with algebraic calculation, was the mode of argument, but often the objects being studied (fluxions, infinitesimals, differentials) were only vaguely understood in highly intuitive terms (at least to modern eyes). In fact, a brilliant attack on the legitimacy of the arguments in calculus, published in 1734 by Bishop George Berkeley (1685-1753) in a pamphlet called *The Analyst*, forced many of the followers of Newton and Leibniz to reformulate their arguments with greater care. During the eighteenth and early nineteenth centuries, the notions of *limit*, *derivative*, and *integral* were formulated and reformulated in ever more careful terms by among others Leonard Euler (1701-1783), Joseph Louis Lagrange (1736-1813), Bernhard Bolzano (1781-1848) and Augustin-Louis Cauchy (1789-1857). In the middle of the nineteenth century, mathematics was shaken by a number of discoveries that led to a thorough re-examination of the foundations of the subject, including the notions of *number*, *function*, and even *geometry*. Among the important contributors to this re-examination were Karl Theodor Wilhelm Weierstrass (1815-1897), George Cantor (1845-1918), and David Hilbert (1862-1943). Much of mathematics since 1900 exhibits greater generality and abstraction, and imposes more stringent standards of rigor, than what came before.

The best way to learn how to prove things is by doing: as you read this book, pay attention to the kinds of arguments you see, and try them out when proving things on your own. You will gradually gain better fluency in the language and become more comfortable with the rules of the game. Here, to help you organize things in your mind, we will outline a few common strategies of proof which are useful in a great many mathematical situations. All correspond to some basic rules of common sense.

Verification and Example

The most elementary "proofs" are straightforward verifications, by calculation or analogous procedures. For example:

Claim: *$x = 2$ is a solution to the equation*

$$x\cos(\pi x) = \sqrt{x+2}$$

Proof. We calculate: when

$$x = 2$$

then

$$\begin{aligned} \pi x &= 2\pi \\ \cos(\pi x) &= \cos 2\pi \\ &= 1 \\ x\cos(\pi x) &= 2 \\ \sqrt{x+2} &= \sqrt{4} \\ &= 2. \end{aligned}$$

□

Traditionally, the end of a proof was indicated by the letters "Q.E.D.", an abbreviation of the Latin phrase, *quod erat demonstrandum*, which translates "as was to be shown". A modern alternative notation is to put a square □ (or other clever symbol) at the end of a proof[2]. It is a kindness to your reader to explicitly indicate the end of your proof, so that they can switch gears and return to the main line of argument.
Of a slightly different flavor is the following:

Claim: *The equation*

$$x^4 - 13x^2 + 36 = 0$$

has a real solution.

[2] This is the notation we have used throughout the text.

Proof. Let

$$x = 2;$$

then

$$\begin{aligned} x^4 - 13x^2 + 36 &= 16 - (13)(4) + 36 \\ &= 16 - 52 + 36 \\ &= 0. \end{aligned}$$

□

The claim is an **existence** statement (*There exists* some x satisfying $x^4 - 13x^2 + 36$), and the simplest way to prove such a statement is to produce an explicit example of the desired kind of object.
Another direct verification is illustrated by

Claim: *7 is prime.*

Proof. The set of integers strictly between 1 and 7 is $\{2, 3, 4, 5, 6\}$; since 7 is odd, it is not divisible by 2, 4, or 6, and

$$\begin{aligned} 2 \cdot 3 < 7 < 3 \cdot 3 \\ 1 \cdot 5 < 7 < 2 \cdot 5. \end{aligned}$$

□

Here we use the fact that any divisor of a natural number lies between 1 and that number, so to check that 7 has no divisors other than 1 and 7, we need only check the five given numbers. The rules of arithmetic are, again, taken for granted.
How about this?

Claim: *6 is* not *prime*

Proof. $2 \times 3 = 6$. □

Notice the difference between the two preceding proofs. The statement that 7 *is* prime can be restated as: *for any two natural numbers a and b, both different from* 1 *and* 7*, the product ab is not equal to* 7*.* Thus it is a statement about *all* entities of a certain sort. To prove that such a statement is *true*, we need to really take care of *all* the possibilities, not just some. However, to show that such a statement is *false*, it suffices to produce just one *counterexample*, because the negation of "*all* entities have a given property" is "*not all* entities have the property"—that is, "*some* entity *fails* to have the property": this negation is an *existence* statement. Conversely, the negation of an *existence* statement like "*there exists* a real solution to $x^2 + 1 = 0$" is effectively "for *all* real numbers x, $x^2 + 1$ does *not* equal zero": to prove an equation *has* a solution, we can just exhibit one; to prove that it does *not* have a solution we need to eliminate *all* real numbers as possible solutions.

A final word of caution: in mathematics, *showing* something is synonymous with *proving* it, not with *illustrating* it: you don't "show" that $x^2 + 1 = 0$ has no solutions by noting that $x = 2$ is not a solution (otherwise, you could also "show" that $x^2 - 1 = 0$ has no solutions).

Proof by Contradiction

Most of the statements we will try to prove can be formulated as *implications* in the form *if A then B* (in symbols, $A \Rightarrow B$), where A and B are two assertions. The **contrapositive** of $A \Rightarrow B$ is obtained by replacing each statement with its negative, and reversing arrows, thus:

$$not\ A \Leftarrow not B;$$

the contrapositive of "if A then B" is "if B fails, then A fails". A statement is logically equivalent to its contrapositive: a convoluted way to say "if A is true then so is B" is "if B is false, then so is A". Sometimes, it is easier to see how to prove the contrapositive of a statement instead of the statement itself.

This is closely related to **proof by contradiction**: we prove a statement by proving that it is impossible that it be false. The form is almost a kind of mathematical satire (or slapstick): we start with the premise that the original statement is *not* true, and go through a chain of reasoning that ends up with an obviously absurd statement.

A classic example is the proof that $\sqrt{2}$ is irrational on p. 21. Another is the proof given by Euclid (Book IX, Prop. 20 of the *Elements*) that there are infinitely many primes. We will use in this proof the observation that a number which is *not* prime is divisible by *some* prime other than 1.

Claim: *There are infinitely many primes.*

Proof. Suppose not; then we can list *all* the primes, say

$$\{\text{primes}\} = \{p_1, p_2, ...p_k\}. \tag{A.1}$$

Consider the number

$$q = p_1 \cdot p_2 \cdot \cdots \cdot p_k + 1. \tag{A.2}$$

Clearly $q > p_i$ for $i = 1, \ldots, k$, so q is not on the list (A.1), and hence is *not* prime. It follows (using our observation) that some $p_i \neq 1$ divides q. But we can rewrite (A.2) as

$$q - (p_1 \cdot p_2 \cdot \cdots \cdot p_k) = 1,$$

and p_i divides both terms on the left, while it cannot divide 1 on the right. This shows that any *finite* list of primes must be incomplete, so there are infinitely many primes, as asserted. □

Mathematical Induction

Suppose we want to prove that all natural numbers n have a certain property. We can try to check the property directly for $n = 1$, or for any particular value for n, but we can't do this explicitly for absolutely *every* value of n. However, suppose we had a way of proving, *abstractly*, that *whenever* the property holds for *some* particular value of n, it must *also* hold for its *successor*, $n + 1$. Then by the domino effect, we could actually prove the statement for *any* value of n: for any *particular* value, like $n = 100$, we could first use our knowledge that the case $n = 1$ holds, together with our abstract result, to conclude that the case $n = 2$ holds; but then, recycling this into our argument, we could conclude that the case $n = 3$ also holds; by repeatedly recycling each conclusion back into our argument, we could (after 99 iterations) conclude that the property holds for $n = 100$. Now, knowing that something *could* be proved is (like checkmate in chess) enough to convince us that it is true; the *coup de grace* of an *explicit* proof in each individual case is unnecessary.
This strategy is formalized as **proof by induction**. A proof by induction has two formal steps:

- the **initial case** , in which we prove a statement for an initial value of n (usually $n = 1$), and
- an **induction step** in which we prove the abstract conditional statement: if the property holds for some integer[3] n then it holds also for $n + 1$.

Here is an example:

> **Claim:** *If $x > 1$, then every (positive integer) power of x, x^n, is greater than or equal to x.*

Proof. (by induction on n)

Initial case: for $n = 1$, $x^1 = x$ (so $x \leq x^1$).

Induction step: Suppose $x^n \geq x$. We need to show $x^{n+1} \geq x$. First, we can multiply our induction hypothesis by x on both sides to get

$$x^{n+1} = x \cdot x^n \geq x \cdot x.$$

But we also know $x > 1$, and multiplying both sides of *this* by x gives

$$x \cdot x > x \cdot 1 = x.$$

These two inequalities together give the desired one

$$x^{n+1} \geq x.$$

□

Here is another example, perhaps a bit less obvious: we want to show that the sum of the first n *odd* integers equals n^2: the n^{th} odd can be written as $2n - 1$, so we want to show that

$$1 + 3 + 5 + \cdots + (2n - 1)^2 = n^2.$$

It will be easier to use summation notation (see p. 18) for this.

> **Claim:** $\sum_{k=1}^{n}(2k - 1) = n^2$.

---

[3]This is called the **induction hypothesis**. A variation on this allows us to assume that the property holds, not *just* for n, but for *all* integers in the set $\{1, \ldots, n\}$.

Proof. (by induction on n)

Initial case: for $n = 1$,

$$\sum_{k=1}^{1}(2k-1) = (2 \cdot 1 - 1) = 1 = 1^2.$$

Induction step: Suppose

$$\sum_{k=1}^{n}(2k-1) = n^2;$$

we need to show that

$$\sum_{k=1}^{n+1}(2k-1) = (n+1)^2.$$

To do this, we break up the sum on the left as

$$\sum_{k=1}^{n+1}(2k-1) = \sum_{k=1}^{n}(2k-1) + \big(2(n+1)-1\big)$$

which, using our induction hypothesis, equals

$$\begin{aligned} n^2 + \big(2(n+1)-1\big) &= n^2 + \big(2n+2-1\big) \\ &= n^2 + 2n + 1 \\ &= (n+1)^2 \end{aligned}$$

proving our induction step.

□

B
Answers to Selected Problems

1.2

1. $\sqrt{5}-\sqrt{3}>\sqrt{15}-\sqrt{13}>\sqrt{115}-\sqrt{113}$.

3. Up (*resp.* down) when the fraction is <1 (*resp.* >1).

5. $a<-c$ or $a>c$

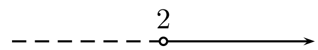
6. (a) **False**
   (c) **False**
   (e) **True**

7. 1. $|x-y|>0$ *unless* $x=y$, *in which case* $|x-y|=0$
   3. $|x-y|\leq|x-z|+|z-y|$

1.3

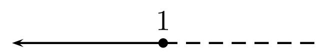
1. (a) $[2,3]$

(c) $(2, \infty)$



(e) $(3, 2]$: Nonsense $(3 > 2)$.

(g) $(-\infty, 1]$



2. (a) $[0, 1)$ and $[1, 2]$
   (c) impossible
   (e) impossible
   (g) $I = (0, 1)$, $J = \left[\frac{1}{2}, \frac{2}{3}\right]$
   (i) $I = [1, \infty)$, $J = (-\infty, -1]$.

5. (a) The union $[0, 1] \cup [2, 3]$ is not an interval.
   (d) Let $I = [0, 1]$ and $J = (1, 2)$; then $I \cap J = \emptyset$, while $I \cup J = [0, 2)$.

2.1

1. $1, 3, 5, 7, 9, 11$

Decimal answers are rounded to two places.

3. $1, \frac{1}{2} = 0.50, \frac{1}{3} \approx 0.33, \frac{1}{4} = 0.25, \frac{1}{5} = 0.20$

5. $1, 1, \frac{1}{2} = 0.50, \frac{1}{6} \approx 0.17, \frac{1}{24} \approx 0.04, \frac{1}{120} \approx 0.01$

7. $1, -\frac{1}{2} = -0.50, \frac{1}{4} = 0.25, -\frac{1}{8} \approx -0.13, \frac{1}{16} \approx 0.06, -\frac{1}{32} \approx -0.03$

9. $0, 1, 0, -1, 0, 1$

11. $0, 1, \frac{\sqrt{3}}{2} \approx 0.87, \frac{1}{\sqrt{2}} \approx 0.71, \sin\frac{\pi}{5} \approx 0.59$

13. $1, 4, 9, 16, 25, 36$

15. $1, 1.50, 1.83, 2.08, 2.28$

17. $1, 2, 2.50, 2.67, 2.71, 2.72$

19. $1, 0.50, 0.75, 0.63, 0.69, 0.66$

21. $100, 105, 110.25, 115.76, 121.55, 127.63$

23. $2, 1.50, 0.83, -0.37, 2.36, 1.94$

25. $1, 1, 2, 3, 5, 8$

27. $0, 1, -0.50, 0.75, -0.63, 0.69$

2.2

In Problems 1-27, α denotes a lower bound and β an upper bound.

1. $2j + 1$: $\alpha = 0$, no upper bound ($a_j > 2j$).

3. $\frac{1}{k}$: $\alpha = 0$, $\beta = 1$ (all terms are positive, and $c_{k+1} < c_k$)

5. $\frac{1}{n!}$: $\alpha = 0$, $\beta = 1$

7. $(-2)^{-k} = (-1)^k \left(\frac{1}{2}\right)^k$: $\alpha = -\frac{1}{2}$, $\beta = 1$ (all powers of $\frac{1}{2}$ are < 1, and signs alternate)

9. $\sin \frac{n\pi}{2}$: $\alpha = -1$, $\beta = 1$ (This holds for the sine of any angle)

11. $\sin \frac{\pi}{n}$: $\alpha = -1$, $\beta = 1$
(Actually, all terms are positive, so $\alpha = 0$ also works)

13. $\sum_{j=0}^{n}(2j + 1)$: $\alpha = 0$, no upper bound

15. $\sum_{k=1}^{n} \frac{1}{k}$: $\alpha = 0$; upper bound is not obvious (turns out to have no upper bound: see §2.3 prob. 9)

17. $\sum_{n=0}^{N} \frac{1}{n!}$: $\alpha = 0$; existence of upper bound not at all clear with what we know now, but later (eg., integral test) will see this converges, in fact to $e < 3$ and again any $\beta \geq e$ works

19. $\sum_{k=0}^{n}(-2)^{-k}$: not immediately obvious, but this is alternating series, so the first two terms are bounds: $\beta = 1$, $\alpha = 1 - \frac{1}{2} = \frac{1}{2}$

21. $\alpha = 100$, no upper bound

23. not clear, although all terms appear to be ≤ 2

25. $\alpha = 0$ (or even $\alpha = 1$, since the first term is 1 and we are always adding something positive to form the next one), no upper bound

27. not obvious, but appears that $\alpha = -1$, $\beta = 1$

28. The examples given below represent some of many possibilities.

    (a) $a_n = n$
    (b) $b_n = (-1)^n n$
    (c) none (convergence implies boundedness)
    (d) $d_n = (-1)^n$
    (e) none (limits are unique)
    (f) $f_n = 3 - n$
    (g) $g_n = 1 - \frac{1}{n}$

31. $x_n = \frac{1}{n}$

2.3

1. (a) nondecreasing (b) unbounded: diverges; no convergent subsequence

3. (a) nonincreasing (b) bounded, nonincreasing: converges

5. (a) nonincreasing (b) bounded, nonincreasing: converges

7. (a) not eventually monotone (b) converges ($|x_n| = \frac{1}{2^k} \to 0$)

9. (a) not eventually monotone (b) bounded, periodic: diverges; the sequence gotten by taking every fourth term is constant, so converges

11. (a) eventually nonincreasing (b) converges: angle gets small, and sine of a small angle is small

13. (a) nondecreasing (b) unbounded($A_N > N+1$): diverges; no convergent subsequence

15. (a) nondecreasing (b) unbounded, so diverges; no convergent subsequence

17. (a) nondecreasing

19. (a) not eventually monotone (b) converges by Geometric Series test (not required)

21. (a) nondecreasing (b) unbounded, so diverges; no convergent subsequence

25. (a) nondecreasing (b) increasing sequence of integers, so diverges

33. (a) $\frac{1}{2}, \frac{3}{4}, \frac{7}{8}$

2.4

1. divergent

3. $c_k = \frac{1}{k} \to 0$

5. $e_n = \frac{1}{n!} \to 0$

7. $g_k = (-2)^{-k} = \left(\frac{-1}{2}\right)^k \to 0$

9. divergent.

11. $\beta_n = \sin\frac{\pi}{n} \to 0$

13. diverges to ∞

15. divergent

19. $G_n = \sum_{k=0}^{n}(-2)^{-k} \to \frac{2}{3}$.

20. (a) $0.111... = \frac{1}{9}$. (c) $0.0101... = \frac{1}{99}$.
    (e) $3.123123... = \frac{3120}{999}$. (g) $1.0123123... = \frac{10113}{9990}$.

21. (a) $\frac{3}{4} = 0.75$ (c) $\frac{5}{6} = 0.8333...$
    (e) $\frac{3}{7} = 0.428571428571...$ (g) $\frac{11}{7} = 1.571428571428...$

22. (a) $a > 0$, $p > 12$ (b) $a = 10$, $p = 12$
    (c) a arbitrary, $p < 12$ (d) $a < 0$, $p > 12$

2.5

1. (a) $\sup S = \max S = 2$, $\inf S = \min S = -3$
   (c) $\sup S = 1$, $\inf S = -1$, $\max S$ and $\min S$ do not exist .
   (e) $\sup S = \max S = 1$, $\inf S = \min S = 0$
   (g) $\sup S = 1$, $\inf S = 0$, $\max S$ and $\min S$ do not exist.
   (i) $\sup S$, $\max S$, $\inf S$ and $\max S$ all fail to exist.
   (k) $\inf S = 0$ and $\min S$ does not exist, $\sup S = \max S = \frac{1}{\pi}$
   (m) $\sup S = 1$, $\inf S = -1$, $\max S$ and $\min S$ do not exist.

2. (a) $\alpha = 0$ is a lower bound, and $\beta = 100$ is an upper bound, for the set of all solutions.
   (c) $\alpha = -1$ and $\beta = 1$ are a lower and upper bound for the set of all such solutions.
   (e) $\beta = 0$ is an upper bound. However, the set of solutions is not bounded below.

7. (a) i. $[-2, -1]$
   iii. $[-2, 1]$
   v. This is impossible.

2.6

1. (a) $I_n = \left[0, 1 = \frac{1}{n}\right]$. (c) No.

3.1

1. (a) $(-\infty, -3) \cup (-3, 3) \cup (3, \infty)$ (c) $(-\infty, -3) \cup [-2, 2] \cup (3, \infty)$
   (e) $\bigcup_{n=-\infty}^{\infty} \left[n\pi, (n + \frac{1}{2})\pi\right)$ (g) $[-1, 1]$

2. (a) $C(A) = 2\sqrt{A\pi}$ (c) $h(s) = s \sin \frac{\pi}{3}$ (e) $V(S) = \frac{S^{3/2}}{6\sqrt{\pi}}$

3. (a) 1 (c) 0 (e) $1 - 4x^2$

4. (a) $(f \circ g)(x) = 2x^2 - 1$ on $(-\infty, \infty)$
   (c) $(f \circ h)(x) = 1 - 2x$ on $(-\infty, 1]$

(e) $(h \circ g)(x) = \sqrt{1 - |x|}$ on $[-1, 1]$

5. (a) $\alpha = 3$ (c) $\alpha = \pm\beta$

7. (a) Not continuous (c) Not continuous

8. (a) $f(x) = \tan \frac{\pi x}{2}$, $0 < x < 1$.

   (c)
$$f(x) = \begin{cases} 1 & -1 \leq x \leq 0 \\ -1 & 0 < x \leq 1. \end{cases}$$

   (e) $y = -\sqrt{1 - x^2}$, $-1 \leq x \leq 1$.

12. (b)
$$f(x) = ax^2 - \frac{2}{a}$$
$$g(x) = a^2x^3 - 3x.$$

3.2

1. (a) $f(1) = -1 < 0$, $f(2) = 3 > 0$, and a polynomial, so continuous on $[1, 2]$.

   (c) $f(0) = 1 > 0$, $\frac{\pi}{2} < 2 < \pi$, so $\cos 2 < 0$, and $f(2) = 2\cos 2 - 1 < 0$; f is continuous on $[0, 2]$.

2. (a) There is a root in $[-2, -1]$

   (c) There is a root in $[-1, 0]$

   (e) There is a root in $[1, 2]$

4. There is a unique solution.

5. (a) Not monotone on $[0, 2]$.

   (c) Increasing on $[0, 1]$.

   (e) Increasing on $[0, \pi]$.

6. (d) $(0, pi)$

   (e) The most natural definition is: for $x \in \mathbb{R}$, $\operatorname{arccot} x$ is the unique angle $\theta \in (0, \pi)$ satisfying $\cot\theta = x$.

7. (b) $\operatorname{arcsec} x$ can only be defined for $x \geq 1$ or $x \leq -1$. For the first case, it is natural to take $0 \leq \theta < \frac{\pi}{2}$ (which amounts to $\operatorname{arcsec} x = \arccos \frac{1}{x}$). In the second case, we can take either $\frac{\pi}{2} < \theta \leq \pi$ or $-\pi < \theta \leq -\frac{\pi}{2}$.
   (c) $\operatorname{arccsc} x$ is defined only for $x \geq 1$ or $x \leq -1$; in the first case, we can take $\operatorname{arccsc} x = \arcsin \frac{1}{x}$, and in the second, either $-\frac{\pi}{2} \leq \theta < 0$ or $-\pi < \theta \leq -\frac{\pi}{2}$.

8. (a) False (c) False (e) False

9. (b) $\cos(\arcsin x) = \sqrt{1-x^2}$. (c) $\sin(\arccos x) = \sqrt{1-x^2}$.

3.3

1. (a) $\min_{(-1,1)} x^2 = f(0) = 0$; $\sup_{(-1,1)} x^2 = 1$; max not achieved.
   (c) $\min_{(-\infty,\infty)} x^2 = f(0) = 0$; not bounded above
   (e) $\min_{[0,1]} f(x) = 0$; $\max_{[0,1]} f(x) = 1$.

2. (a) $f(x) = \arctan x$. (d) $f(x) = \tan x$, $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$
   (e) $f(x) = x$, $x \in (0,1)$ (f) $f(x) = \sin x$

8. **(a)**

$$\begin{array}{lll} \min f(x) = a+2 & \max f(x) = a+2 & \text{for } a \leq -1 \\ \min f(x) = a^2 & \max f(x) = a+2 & \text{for } -1 < a < 0 \\ \min f(x) = 0 & \max f(x) = (a+1)^2 & \text{for } a \geq 0. \end{array}$$

(b)

$$\begin{array}{lll} \min f(x) = a+2 & \max f(x) = a+2 & \text{for } a \leq -1 \\ \min f(x) = a^2 & \max f(x) = a+2 & \text{for } -1 < a < 0 \\ \min f(x) = 0 & \max f(x) \text{ does not exist} & \text{for } a \geq 0. \end{array}$$

3.4

1. (a) $\frac{3}{2}$ (c) 0 (e) 0 (g) $\frac{8}{5}$
   (i) $\frac{\pi}{2}$ (k) 1 (m) diverges (o) $\frac{5}{3}$
   (q) 0 (s) 0 (u) 0

3.5

1. (a) $f(x)$ blows up at $x = 1$: $\lim_{x\to 1^-} f(x) = -\infty$, $\lim_{x\to 1^+} f(x) = \infty$.

   (c) $f(x)$ blows up at $x = 1$: $\lim_{x\to 1^-} f(x) = -\infty$, $\lim_{x\to 1^+} f(x) = \infty$.

   (e) Removable discontinuity at $x = 1$: $\lim_{x\to 1} f(x) = \frac{2}{3}$.

   (g) $f(x)$ blows up at $x = -1$: $\lim_{x\to -1^-} f(x) = \infty$, $\lim_{x\to -1^+} f(x) = -\infty$; removable discontinuity at $x = 1$, $\lim_{x\to 1} f(x) = 0$.

   (i) $f(x)$ blows up at $x = \pm 1$: $\lim_{x\to -1^-} f(x) = -\infty$, $\lim_{x\to -1^+} f(x) = \infty$, $\lim_{x\to 1^-} f(x) = -\infty$, $\lim_{x\to 1^+} f(x) = \infty$.

   (k) $f(x)$ blows up at $x = \pm 1$: $\lim_{x\to -1^-} f(x) = -\infty$, $\lim_{x\to -1^+} f(x) = \infty$, $\lim_{x\to 1^-} f(x) = -\infty$, $\lim_{x\to 1^+} f(x) = \infty$. $\lim_{x\to 0^-} f(x) = -1$, $\lim_{x\to 0^+} f(x) = 1$.

   (m) $f(x)$has removable discontinuity at $x = 0$, $\lim_{x\to 0} f(x) = 1$.

   (o) $f(x)$ has a removable discontinuity at $x = n\pi$, $n = 0, \pm 1, \pm 2, \ldots$; $\lim_{x\to n\pi} f(x) = 0$

   (q) $f(x)$ has a jump discontinuity at $x = 1$: $\lim_{x\to 1^-} f(x) = 0$, $\lim_{x\to 1^+} f(x) = 2$.

2. (a) $f(x) = \dfrac{x^2 + 2x - 3}{x^2 + x - 2}$, $a = 1$; $f(a) = 2$

   (b) $f(x) = \dfrac{1}{x^2}$, $a = 0$ (c) $f(x) = \dfrac{1}{x}$, $a = 0$

3. (a) 0 (c) -4 (e) none exists

4. (c) The only choice satisfying *both* conditions are $\alpha = 2$ and $\beta = 1$.

5. (c) i. $p(a) \neq 0$ and $q(x)$ has a zero of *odd* multiplicity n at $x = a$;

   ii. $p(a)$ and $\tilde{q}(a)$ have the same sign, and n is even;

   iii. $p(a)$ and $\tilde{q}(a)$ have opposite signs, and n is even.

7. $f(x)$ has a removable discontinuity *at every point* $x \in (0, 1)$.

3.6

1. (a) -3 (c) -3

2. (a) For $b = 1$, b^x is bounded above and below. Otherwise, $\inf b^x = 0$, but b^x is not bounded above.

3.7

1. (c) $\delta = 0.02 < 0.025$ works.

7. (a) $\mathbf{\lim_{x\to a} f(x) = \infty}$ if $\operatorname{dom}(f)$ is not bounded above and, given M, there exists $\delta > 0$ such that $|x - a| < \delta$ and $x \in \operatorname{dom}(f)$ guarantees $f(x) > M$.

   (b) $\mathbf{\lim_{x\to\infty} f(x) = \infty}$ if $\operatorname{dom}(f)$ is not bounded above and, given M, there exists N such that $x > N$ and $x \in \operatorname{dom}(f)$ guarantees $f(x) > M$.

4.1

1. (a) 3 (c) $-\frac{3}{2}$ (e) $\frac{3}{2}$

2. (a) $y = 3x - 1$ (c) $y = 3$ (e) $y = -x + 3$

3. (a) $y = 3x + 2$

4. (a) $f'(1) = 6$ (c) $f'(2) = 18$

   (e) $f'(0) = -2$ (g) $f'(1) = 2$

5. (a) 12 (c) 7.5

   (f) $y = 7x - 5$.

6. (a) $y = 1$; $y = -0.5x - 0.5$; $y = -0.9x - 0.1$

   (c) $y + 1 = -(x - 1)$, or $y = -x$ (d) $y + 1 = x - 1$, or $y = x - 2$.

10. (a) $f(x) = x^2 - x$ (b) $f(x) = x$

    (d) impossible (e) $f(x) = \left|x - \frac{1}{2}\right|$

11. (c) $y - 2 = \frac{1}{3}(x + 1)$, or $3y - x = 7$.

12. (a) $f(x) = |x|$, $a = 0$ is an example, or more generally $f(x) = |x - a|$ works for any given a.

4.2

1. (a) 4 (c) 64 (e) $-\frac{19}{192}$ (g) 7

   (i) 22 (k) $\frac{3\pi-9\sqrt{3}}{2\pi^2}$ (m) 3 (o) $\frac{1}{3}$

   (q) 1 (s) 1 (u) $-\dfrac{8}{\pi^2}(\pi+2)$

   (w) 1

2. (a) $y-6=4(x-1)$ (c) $y-64=\frac{191}{4}(x-4)$

   (e) $y-\frac{5}{64}=-\frac{19}{192}(x-64)$ (g) $y=7x-5$

   (i) $y=22x-6$ (k) $y-\frac{3\sqrt{3}}{2\pi}=\left(\frac{3\pi-9\sqrt{3}}{2\pi^2}\right)(x-\frac{\pi}{3})$

   (m) $y=3x+3$ (o) $x=3y+1$

   (q) $y=x-1$ (s) $y=x$

   (u) $\pi^2y+8(\pi+2)x=2\pi(\pi+4)$ (w) $6y-x=3\sqrt{3}-2\pi.$

3. (a) $f'\left(\frac{\pi}{3}\right)=4$ and $g'\left(\frac{\pi}{3}\right)=2\sqrt{3}.$

5. (a)

$$f'(x)=\begin{cases}2x & \text{for } x<1,\\ 2 & \text{for } x=1,\\ 2 & \text{for } x>1.\end{cases}$$

   (b)

$$f'(x)=\begin{cases}2x & \text{for } x<1,\\ 2 & \text{for } x>1.\end{cases}$$

   Not differentiable at $a=1$.

   (e)

$$f'(x)=\begin{cases}\cos x & \text{for } x<0,\\ 1 & \text{for } x=0,\\ 1 & \text{for } x>0.\end{cases}$$

6. (a) all α (c) none (e) $\alpha=-1$

7. (a)

$$
\begin{aligned}
f(x) &= x^2 + 2x + 3 & f'(x) &= 2x + 2 \\
f''(x) &= 2 & f'''(x) &= 0.
\end{aligned}
$$

(c)

$$
\begin{aligned}
f(x) &= x^{1/2} & f'(x) &= \frac{1}{2}x^{-1/2} \\
f''(x) &= -\frac{1}{4}x^{-3/2} & f'''(x) &= \frac{3}{8}x^{-5/2}.
\end{aligned}
$$

8. (a)

$$
\begin{aligned}
\frac{dy}{dx} &= \cos x & \frac{d^2y}{dx^2} &= -\sin x \\
\frac{d^3y}{dx^3} &= -\cos x & \frac{d^4y}{dx^4} &= \sin x
\end{aligned}
$$

(c)

$$
\frac{d^{2k}y}{dx^{2k}} = (-1)^k \sin x \qquad \frac{d^{2k+1}y}{dx^{2k+1}} = (-1)^k \cos x.
$$

4.3

1. (a) $2 + 2\ln 2$ (c) 1 (e) $\frac{1}{2}$

2. (a) $y = (2 + 2\ln 2)x + 2\ln 2$
   (c) $y = x + 1$ (e) $y = x/2$

3. (a)

$$
f'(x) = \begin{cases} e^x & \text{for } x < 0, \\ 1 & \text{for } x = 0, \\ 1 & \text{for } x > 0. \end{cases}
$$

(c)

$$
f'(x) = \begin{cases} e^x & \text{for } x < 0, \\ 1 & \text{for } x > 0. \end{cases}
$$

Not differentiable at $x = 0$.

4.4

1. (a) $f^{-1}(x) = x - 1$, $x \in (-\infty, \infty)$
   (c) $f^{-1}(x) = \sqrt[3]{x+1}$, $x \in (-\infty, \infty)$
   (e) $f^{-1}(x) = \frac{1-2x}{x-1}$, $x \in (-\infty, 1) \cup (1, \infty)$
   (g) $f^{-1}(x) = x^2$, $x \in [0, \infty)$
   (i) $f^{-1}(x) = \ln(x-1)$, $x \in (1, \infty)$

2. (a) 2 (c) 2 (e) $\frac{\pi}{4} + \frac{1}{2}$
   (g) $-\ln\frac{\pi}{2}$ (i) $\frac{1}{8\ln 2}$

3. (a) $y - 1 = \dfrac{1}{5}(x-4)$.

4.5

1. (a) $f'(x) = 6(2x+1)^2$ (c) $f'(x) = \frac{2x}{\sqrt{x^2+1}}$
   (e) $f'(x) = 2\pi\cos(2\pi x)$ (g) $f'(x) = xe^{x^2/2}$
   (i) $f'(x) = -\frac{1}{x^2}\cos\frac{1}{x}$ (k) $f'(x) = e^x(\cos 2x - 2\sin 2x)$
   (m) $f'(x) = x/(x^2+1)\sqrt{\ln(x^2+1)}$ (o) $f'(x) = (2x-1)/2\sqrt{x^2-x}$
   (q) $f'(x) = e^x e^{e^x}$
   (s) $f'(x) = (3x^2+2x+5)^2(x^3-2x)[6(3x+1)(x^3-2x) + (3x^2+2x+5)(6x^2-1)]$
   (u) $f'(x) = \left(3x^{\frac{1}{6}} + 9x^{\frac{5}{6}} + 2\right)/\left(9x^{\frac{2}{3}}\right)(x^{\frac{1}{2}} + x^{\frac{3}{2}} + x^{\frac{1}{3}})^{\frac{1}{3}}$
   (w) $f'(x) = \dfrac{x}{\sqrt{x^2+1}}\sec^2\sqrt{x^2+1}$

2. (a) $2y - 3x = -5$ (c) $y + 4x = -7$ (e) $5y - 4x = 8$

3. (a) $f'(1) = -\frac{1225}{2\sqrt{2}}$ (b) $f'(2) = \frac{2-3\ln 2}{4\sqrt{3}}$

4.6

1. Decreasing at $\frac{8}{25\pi} \approx 0.32$ cubic centimeters per minute.

3. Increasing at $\frac{14}{15} \approx 0.93$ units per second.

6. Decreasing at $\frac{8}{3}$ feet per second.

8. $\frac{dm}{dt} = \frac{3-2m}{x}$.

10. Increasing at $\frac{7}{4}$ inches per second.

12. Decreasing at

    (a) $\frac{12}{\sqrt{13}} \approx 3.3\,ft/sec$ (b) $\frac{6}{\sqrt{5}} \approx 2.7\,ft/sec$

4.7

1. (a) $f(x)$ has stationary points at $x = 1 \pm \sqrt{\frac{1}{3}}$.

   (c) $f(x)$ has a stationary point at $x = 0$, and points of non-differentiability at $x = \pm 1$.

   (e) $f(x)$ has stationary points at $x = \pm 1$.

   (g) $f(x)$ has a point of non-differentiability at $x = -1$ and stationary points at $x = 0$ and $x = 4$.

2. (a) $\min_I f(x) = f(3) = -54$, $\max_I f(x) = f(-2) = 46$.

   (c) $\min_I f(x) = f(0) = 0$, $\max_I f(x)$ does not exist.

   (e) $\min_I f(x) = f(-2) = -\frac{1}{4}$, $\max_I f(x) = f(2) = \frac{1}{4}$.

   (g) $\min_I f(x)$ and $\max_I f(x)$ do not exist.

3. $500\,ft$ (parallel to the river) by $250\,ft$.

5. (a) $r = 10$, $h = 20$ (b) $r = 10$, $h = 20$

   (c) Minimum when $r = 10\sqrt[3]{2}$ and $h = 10$, maximum when $r = 10\sqrt{2}$ and $h = 10\sqrt{2}$.

7. The maximum area is achieved when the upper right corner is located at $(2/\sqrt{3}, \sqrt{2}/\sqrt{3})$.

8. (a) $r = h = \frac{P}{\pi+4}$ (b) $r = \frac{P}{\pi+2}$, $h = 0$

12. The longest ladder that fits has length $\sqrt{w_1^2 + 3w_1^{4/3}w_2^{2/3} + 3w_1^{2/3}w_2^{4/3} + w_2^2}$.

4.8

1. (a) $f(x) = x^2 - 4x$:

   i. **Intercepts:** $(0, 0)$, $(4, 0)$

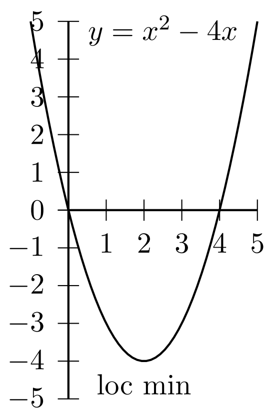
   ii. **Asymptotes:** no horizontal or vertical asymptotes.

   iii. **Monotonicity:** $f(x) \upuparrows$ on $[2, \infty)$, $f(x) \downdownarrows$ on $(-\infty, 2]$.

   iv. **Local Extrema:** local min at $(2, -4)$, no local minimuma.

   v. **Concavity:** concave up everywhere.

   vi. **Points of Inflection:** None.



   (c) $f(x) = x^4 - 8x^3 + 22x^2 - 24x + 8 = (x - 2)^2(x^2 - 4x + 2)$:

   i. **Intercepts:** $(2, 0), (2 \pm \sqrt{2}, 0), (0, 8)$

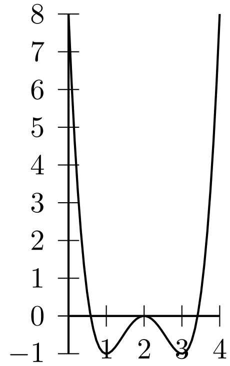
   ii. **Asymptotes:** no horizontal or vertical asymptotes.

   iii. **Monotonicity:** $f(x) \upuparrows$ on $[1, 2]$ and on $[3, \infty)$, $f(x) \downdownarrows$ on $(-\infty, 1]$ and on $[2, 3]$.

   iv. **Local Extrema:** local max at $(2, 0)$, local minima at $(1, -1)$ and $(3, -1)$

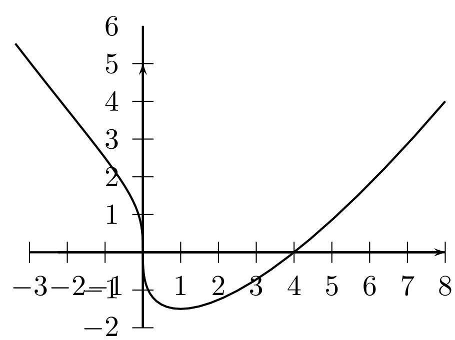
   v. **Concavity:** concave up on $\left(-\infty, 2 - \frac{1}{\sqrt{3}}\right)$ and $\left(2 + \frac{1}{\sqrt{3}}, \infty\right)$, concave down on $\left(2 - \frac{1}{\sqrt{3}}, 2 + \frac{1}{\sqrt{3}}\right)$.

   vi. **Points of Inflection:** $(2 \pm \frac{1}{\sqrt{3}}, -\frac{5}{9})$.



(e) $f(x) = \dfrac{x^{4/3}}{2} - 2x^{1/3} = (x-4)\dfrac{\sqrt[3]{x}}{2}$:

i. **Intercepts:** $(0,0),(4,0)$
ii. **Asymptotes:** no horizontal or vertical asymptotes.
iii. **Monotonicity:** $f(x) \upuparrows$ on $[1,\infty)$ and $f(x) \downdownarrows$ on $(-\infty,1]$.
iv. **Local Extrema:** no local maximum;local minimum at $(1,-\frac{3}{2})$.
v. **Concavity:** $f \smile$ on $(-\infty,-2]$ and $(0,\infty)$ and $f \frown$ on $[-2,0)$.
vi. **Points of Inflection:** $(-2, 3\sqrt[3]{2})$ and $(0,0)$.



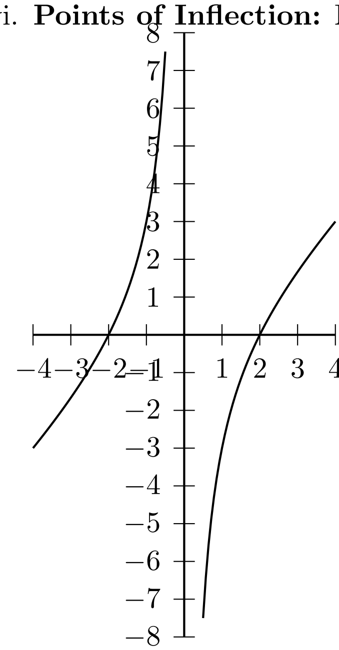
(g) $f(x) = x - \dfrac{4}{x}$:

i. **Intercepts:** $(\pm 2, 0)$; no y-intercept
ii. **Asymptotes:** no horizontal asymptote; VA $x = 0$.
iii. **Monotonicity:** $f(x) \upuparrows$ on $(-\infty,0)$ and on $(0,\infty)$.

iv. **Local Extrema:** No local extrema.

v. **Concavity:** concave up on $(-\infty, 0)$ and concave down on $(0, \infty)$

vi. **Points of Inflection:** No points of inflection.



(i) $f(x) = \dfrac{x^2 + x - 2}{x^2 - x - 2} = 1 + \dfrac{2x}{(x+1)(x-2)}$:

i. **Intercepts:** $(-2, 0), (1, 0), (0, 1)$.

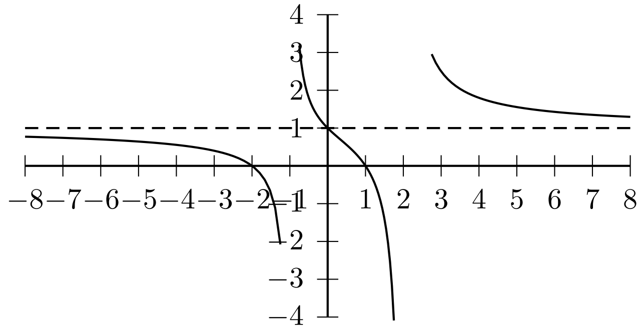
ii. **Asymptotes:** HA: $y = 1$; VA: $x = -1$ and $x = 2$.

iii. **Monotonicity:** $f(x) \downarrow$ on each of the intervals $(-\infty, -1)$, $(-1, 2)$ and $(2, \infty)$.

iv. **Local Extrema:** No local extrema.

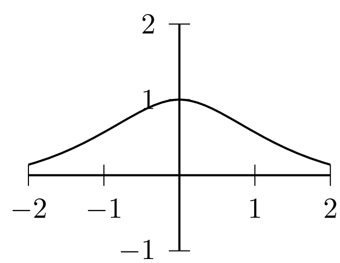
v. **Concavity:** $f''(x) = 0$ at a unique point between $x = 0$ and $x = 1$, call it $x = x_0$ (it is hard to solve the cubic equation $x^3 + 6x - 2 = 0$; concave up on $(-1, x_0)$ and $(2, \infty)$, concave down on $(-\infty, -1)$ and $(x_0, 2)$.

vi. **Points of Inflection:** $(x_0, f(x_0))$ (both coordinates in $(0, 1)$).



(k) $f(x) = e^{-x^2/2}$:

i. **Intercepts:** $(0, 1)$; no x-intercept.
ii. **Asymptotes:** HA: $y = 0$; no VA.
iii. **Monotonicity:** $f(x)\uparrow$ on $(-\infty, 0]$ and $f(x)\downarrow$ on $[0, \infty)$.
iv. **Local Extrema:** local maximum $(0, 1)$; no local minima.
v. **Concavity:** concave up on $(-\infty, -1)$ and $(1, \infty)$, concave down on $(-1, 1)$
vi. **Points of Inflection:** $(\pm 1, e^{-1/2})$.

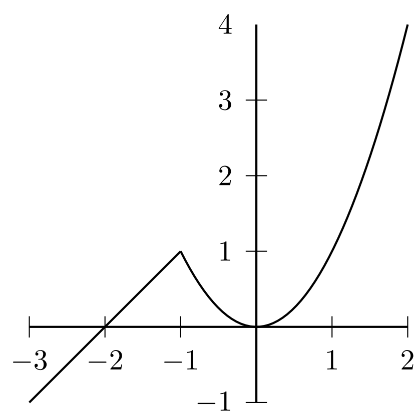


(m)

$$f(x) = \begin{cases} x + 2 & \text{for } x < -1, \\ x^2 & \text{for } x \geq -1. \end{cases}$$

i. **Intercepts:** $(-2, 0), (0, 0)$
ii. **Asymptotes:** None.
iii. **Monotonicity:** $f(x)\uparrow$ on $(-\infty, -1]$ and $[0, \infty)$, and $f(x)\downarrow$ on $[-1, 0]$.

iv. **Local Extrema:** local maximum $(-1, 1)$; local minimum $(0, 0)$.
v. **Concavity:** Neither concave up nor concave down on $(-\infty, -1]$, concave up on $[-1, \infty)$.
vi. **Points of Inflection:** None.



4.9

1. (a) $c = 0$ (c) $c = 1$
   (e) $c = 0$ (g) $c = \pi/4$

3. (a) $c = 1/2$ (c) $c = 2/\sqrt{3}$ (e) $c = 1$

4. (a) $c = \frac{2}{3}$, but $c_f = \frac{1}{2}$, $c_g = \frac{1}{\sqrt{3}}$. (c) $c = \frac{\pi}{4}$, but $c_f = \arccos \frac{3(\sqrt{3}-1)}{\pi}$, $c_g = \arcsin \frac{3(\sqrt{3}-1)}{\pi}$.

6. The derivative of an odd function can have any value.

10. $f(x) = |x|$ is an example.

4.10

1. (a) $\frac{1}{2}$ (c) 2 (e) $\frac{1}{2}$ (g) $-\frac{1}{3}$
   (i) 1 (k) $\frac{1}{2}$ (m) 0 (o) 0

(q) e^2 (s) 0

4. (a) $g(x) = x$, $f(x) = 1$ and $b = 0$

5. Take $f(x)$ to be any nonzero constant multiple of $g(x)$.

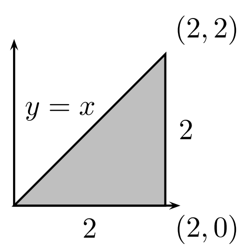
5.1

1. (a) $\mathcal{L}(\mathcal{P}_1, f) = 1.625$, $\mathcal{U}(\mathcal{P}_1, f) = 3.125$.
   (b) $\mathcal{L}(\mathcal{P}_2, f) = \frac{50}{27} \approx 1.852$; $\mathcal{U}(\mathcal{P}_2, f) = \frac{77}{27} \approx 2.852$.
   (c) $\mathcal{L}(\mathcal{P}, f) = \frac{139}{72} \approx 1.931$; $\mathcal{U}(\mathcal{P}, f) = \frac{199}{72} \approx 2.764$.
   (d) $\mathcal{L}(\mathcal{P}', f) = 1.8$; $\mathcal{U}(\mathcal{P}', f) = 2.912$.

2. (a) $\lim \mathcal{L}(\mathcal{P}_n, f) = \lim \frac{n(n-1)}{2n^2} = \frac{1}{2}$;
   $\lim \mathcal{U}(\mathcal{P}_n, f) = \lim \frac{n(n+1)}{2n^2} = \frac{1}{2}$.
   (c) $\lim \mathcal{L}(\mathcal{P}_n, f) = \lim \frac{8}{3}\frac{n(n+1)(2n+1)}{n^3} - \frac{32}{2}\frac{n(n-1)}{n^2} + 14 = \frac{10}{3}$;
   $\lim \mathcal{U}(\mathcal{P}_n, f) = \lim \frac{8}{3}\frac{(n-1)(n)(2n-1)}{n^3} - \frac{32}{2}\frac{(n-1)(n)}{n^2} + 14\frac{n-1}{n} = \frac{10}{3}$

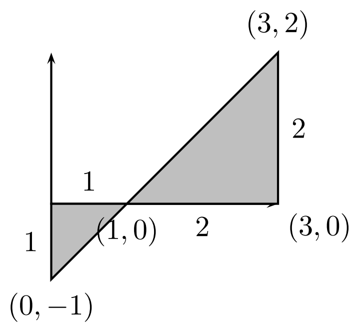
5.2

1. (a) $A = \frac{1}{2}(2)(2) = 2$:



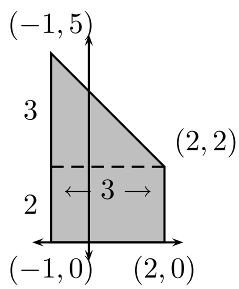
   (c)

$$\int_0^3 (x-1)\,dx = \int_0^1 (x-1)\,dx + \int_1^3 (x-1)\,dx$$
$$= -\frac{1}{2}(1)(1) + \frac{1}{2}(2)(2) = \frac{3}{2}:$$

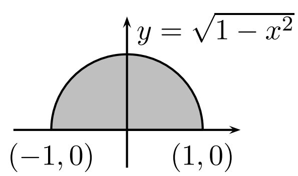


(e)

$$A = \int_{-1}^{2} (4 - x)\, dx$$
$$= (3)(2) + \frac{1}{2}(3)(3) = \frac{21}{2} :$$

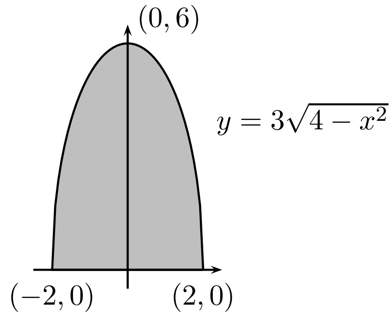


(g) $\int_{-1}^{1} \sqrt{1 - x^2}\, dx = \frac{1}{2}(\pi \cdot 1^2) = \frac{\pi}{2}$:



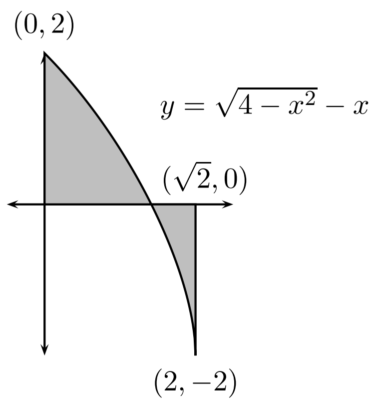
(i)

$$\int_{-2}^{2} 3\sqrt{4 - x^2}\, dx = 3 \int_{-2}^{2} \sqrt{4 - x^2}\, dx$$
$$= (3)(\frac{1}{2})(\pi \cdot 2^2) = 6\pi :$$

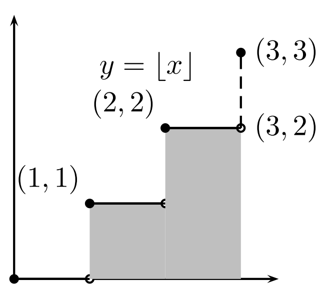


(k)

$$\int_0^2 (\sqrt{4-x^2} - x)\,dx = \int_0^2 \sqrt{4-x^2}\,dx - \int_0^2 x)\,dx$$
$$= \pi - \frac{1}{2}(2)(2) = \pi - 2 :$$

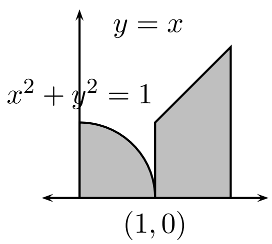


(m) $\int_0^3 \lfloor x \rfloor\,dx = (1)(0) + (1)(1) + (1)(2) + (0)(3) = 3 :$



(o)

$$\int_0^2 f(x)\ dx = \int_0^1 f(x)\ dx + \int_1^2 f(x)\ dx = \frac{\pi}{4} + \frac{3}{2} :$$



3. $\int_0^1 \sin \frac{\pi}{2} x^2 \, dx > \int_0^1 \sin \frac{\pi}{2} x^4 \, dx$.

4. (a) **False** (c) **False** (e) **True** (g) **True**

5. Both averages are $\frac{\pi}{4}$.

12. (a)

$$f(x) = \begin{cases} -1 & \text{if } x \text{ is rational ,} \\ 1 & \text{if } x \text{ is irrational.} \end{cases}$$

5.3

1. (a) $x^4 \Big|_0^1 = 1$ (c) $(x^3 + x)\Big|_{-1}^2 = 12$

   (e) $-x^{-1/2}\Big|_1^4 = \frac{1}{2}$ (g) $\ln x \Big|_1^2 = \ln 2$

   (i) $\sin x \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = 2$

2.

$$\Phi(x) = \begin{cases} \int_{-1}^x (-1) \, dt = -1 - x & \text{for } x < 0, \\ -1 + \int_0^x 1 \, dt = x - 1 & \text{for } x > 0 \end{cases}$$

which is to say, $\Phi(x) = |x| - 1$.

3. (a) $x = \sqrt{(n + \frac{1}{2})\pi}, \quad n = 0, 1, 2, ...$ and
   $x = -\sqrt{(|n| + \frac{1}{2})\pi}, \quad n = -1, -2, ...$

   (b) $\Phi(x) \uparrow$ on each closed interval of the form $\bar{I}_n = [x_{n-1}, x_n]$ with n *even*.

   (c) Similarly $\Phi(x) \downarrow$ on $\bar{I}_n = [x_{n-1}, x_n]$ with n *odd*.

   (d) Local maximum (*resp.* minimum) at each point in the list for (a) with n *even* (*resp.* *odd*).

5.4

1. (a) $x^3 + \frac{2}{3}x^{3/2} + C$ (c) $\frac{1}{6}$
   (e) $x - \arctan x + C$ (g) $-\frac{1}{3}\cos(3x+1)+C$
   (i) $\sqrt{x^2+1} + C$ (k) $\frac{1}{3}e^{3x-2} + C$
   (m) $\frac{x}{3}\sin 3x + \frac{1}{9}\cos 3x + C$ (o) $-\frac{1}{4}\sqrt{1-4x^2} + C$
   (q) $-\frac{1}{2}e^{-x^2} + C$ (s) $\frac{x^3}{3}\ln x - \frac{x^3}{9} + C$
   (u) $\frac{1}{x}\ln\frac{1}{x} - \frac{1}{x} + C$ (w) $\ln(\ln x) + C$

2. (a) $\frac{e^x}{2}(\sin x + \cos x) + C$
   (b) $-\frac{3}{13}e^{2x}\cos 3x + \frac{2}{13}e^{2x}\sin 3x + C$
   (c)

$$\int e^{ax}\sin bx\,dx = \frac{a}{a^2+b^2}e^{ax}\sin bx - \frac{b}{a^2+b^2}e^{ax}\cos bx + C;$$
$$\int e^{ax}\cos bx\,dx = \frac{b}{a^2+b^2}e^{ax}\sin bx + \frac{a}{a^2+b^2}e^{ax}\cos bx + C.$$

5.5

1. (a) $\frac{\sin^2 x}{2} + C$ (c) $\frac{1}{5}\cos^5 x - \frac{1}{3}\cos^3 x + C$
   (e) $\frac{3x}{8} - \frac{1}{4}\sin 2x + \frac{1}{32}\sin 4x + C$ (g) $2\ln|\sec x| + C$
   (i) $\ln|\sec x + \tan x| - \sin x + C$

2. (a) $\frac{1}{3}\tan^3 x + \tan x + C$ (c) $\sec x + C$
   (e) $\frac{1}{3}\tan^3 x + C$ (g) $\frac{1}{3}\sec^3 x + C$
   (i) $\frac{1}{4}\sec^4 x + C$

3. (a) $2\arcsin\frac{x}{2} + \frac{x}{2}\sqrt{4-x^2} + C$
   (c) $-2\ln\frac{x+\sqrt{4-x^2}}{x} + \sqrt{4-x^2} + C$
   (e) $\ln|\sec\theta + \tan\theta| + C$

4. (a) $\frac{x}{2}\sin 2x + \frac{1}{4}\cos 2x + C$
   (c) $\frac{x^3}{6} - \frac{x^2}{4}\sin 2x - \frac{x}{4}\cos 2x + \frac{1}{8}\sin 2x + C$
   (e) $\frac{x}{2}\tan 2x - \frac{1}{4}\ln|\sec 2x| + C$

(g) $x \arcsin x + \sqrt{1-x^2} + C$

(i) $\frac{1}{2}\arcsin^2 x + \frac{x}{2}\sqrt{1-x^2} - \frac{1}{4}\arcsin^2 x - \frac{x^2}{4} + C$

5. (a) $\frac{4}{7}\sin 3x \sin 4x + \frac{3}{7}\cos 3x \cos 4x + C$

(b) $\frac{b}{b^2-a^2}\sin ax \sin bx + \frac{a}{b^2-a^2}\cos ax \cos bx + C$

(c) $\frac{\cos(a+b)x}{2(a+b)} - \frac{\cos(a-b)x}{2(a-b)} + C$

(d) $\int \sin \pm bx \cos bx\, dx = \pm\frac{1}{2b}\sin^2 bx \; + C.$

5.6

1. (a) $\dfrac{2}{x-1} - \dfrac{1}{x-2}$ (c) $\dfrac{1}{x-1} - \dfrac{1}{x+1}$

(e) $-\dfrac{2x+1}{x^2+1} + \dfrac{2}{x-1}$ (g) $\dfrac{x+1}{x^2+1} - \dfrac{1}{x+1}$

(i) $\dfrac{2x+1}{(x^2+1)^2} + \dfrac{2}{x^2+1} + \dfrac{1}{x+1} - \dfrac{1}{x-1}$

2. (a) $\ln|x-1| - \ln|x+1| + C$

(c) $2\ln|x+3| + 3\ln|x-1| + C$

(e) $x^2 + x + \ln|x-3| + C$

(g) $2\ln|x+1| - 2\ln|x+2| - \frac{3}{x+2} + C$

(i) $2\ln|x+1| + \frac{1}{x+1} - 2\ln|x+2| + C$

(k) $2\ln|x+1| + \ln(x^2+2x+5) + \frac{1}{2}\arctan\left(\frac{x+1}{2}\right) + C$

3. (a) $\frac{1}{3} + \frac{1}{5}$ (c) $\frac{2}{3} + \frac{1}{3^2} + \frac{2}{3^3}$ (e) $\frac{2}{5} - \frac{1}{7}$

9. (f) i. $\dfrac{2}{x-1} - \dfrac{1}{x-2}$

iii. $\dfrac{1}{x-1} + \dfrac{1}{x+1}$

v. $\dfrac{1}{x-1} - \dfrac{5}{x-2} + \dfrac{5}{x-3}.$

5.7

1. (a) diverges (c) diverges (e) diverges

(g) diverges (i) π (k) diverges

(m) diverges (o) 0

2. (a) converges (c) converges (e) converges

3. (a) converges (c) converges (e) converges
   (g) diverges (i) converges

5.8

1. $A = \int_{-1}^{3} 2x\,dx - \int_{-1}^{3} (x^2 - 3)\,dx$

2. (a) (See Figure B.1.) $A = \frac{64}{3}$

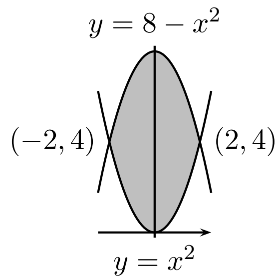


Figure B.1: Problem 2(a)

   (c) (See Figure B.2.) $A = \frac{81}{2}$

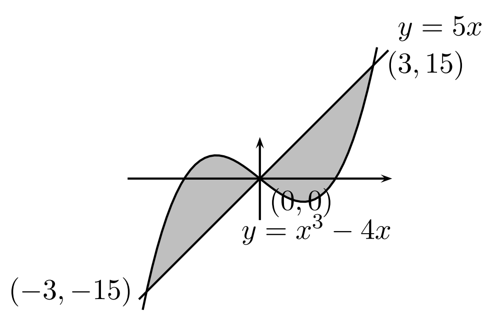


Figure B.2: Problem 2(c)

   (e) (See Figure B.3.) $A = \frac{584}{3} = 194\frac{1}{3}$.

3. (a) $A = \pi$ (see Figure B.4)

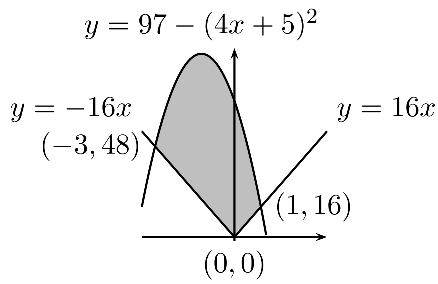


Figure B.3: Problem 2(e)

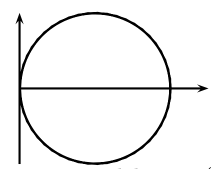


Figure B.4: Problem 3(a)

4. (a) see Figure B.5 and B.6.
   (b) $A = \pi/4$ for n odd and $\pi/2$ for n even.

5. (a) $V = \int_0^1 \pi(1 - x^{2/3})\,dx = \int_0^1 2\pi y^4\,dy = \frac{2\pi}{5}$ (See Figure B.7).
   (c) $V = \int_0^1 \pi(x - x^4)\,dx = \int_0^1 (y^{3/2} - y^3)\,dy = \frac{3\pi}{10}$ (See Figure B.8).
   (e) $V = \int_0^{1/\sqrt{2}} \pi(\arcsin y)^2\,dy + \int_{1/\sqrt{2}}^1 \pi(\arccos y)^2\,dy =$
   $2\pi \int_0^{\pi/4} x(\cos x - \sin x)\,dx = \pi\left(\frac{\pi}{\sqrt{2}} - 2\right)$. (See Figure B.9)

6. $V = \int_{-R\sqrt{3}}^{R\sqrt{3}} \pi[4R^2 - 1 - x^2]\,\triangle x = 2\pi\sqrt{3}R(3R^2 - 1)$

6.1

1. (a) $T_2^3 f(x) = \frac{1}{2} - \frac{x-2}{4} + \frac{(x-2)^2}{8} - \frac{(x-2)^3}{16}$
   (c) $T_{\frac{\pi}{2}}^3 f(x) = 1 - \frac{(x-\frac{\pi}{2})^2}{2!}$
   (e) $T_1^3 f(x) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3}$.

3. (a) $\left|f(1) - T_0^5 f(1)\right| = \left|\frac{f^{(6)}(s)}{6!}(1)^6\right| < \frac{1}{720} \approx 0.00139$.
   (b) $n \geq 7$

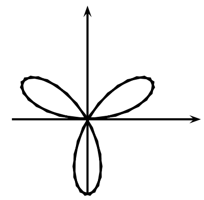


Figure B.5: $r = \sin 3\theta$

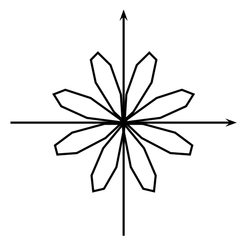


Figure B.6: $r = \sin 4\theta$

6.2

1. (a) converges (c) converges (e) diverges
   (g) converges for $p > 0$, diverges for $p \leq 0$
   (i) diverges (k) diverges
   (m) converges (o) converges

6. (c) $\sum \frac{1}{k^2}$ converges, but $\sum \frac{1}{k^{1/2}}$ diverges.

6.3

1. (a) converges unconditionally (c) converges unconditionally
   (e) converges conditionally (g) converges unconditionally
   (i) converges unconditionally

6.4

1. $[-1, 1]$ 3. $[-1, 1)$ 5. $[-1, 1]$

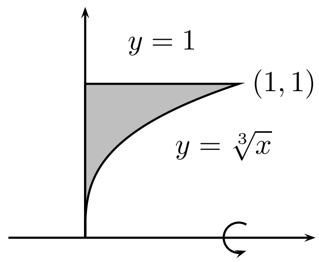


Figure B.7: Problem 5(a)

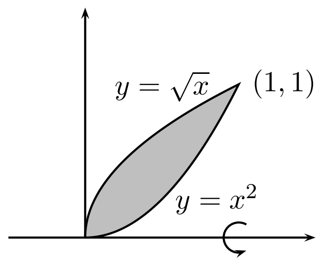


Figure B.8: Problem 5(c)

7. $(-\infty, \infty)$ 9. $(-\infty, \infty)$ 11. $(-1, 1]$

13. $\left(-\sqrt{2}, \sqrt{2}\right)$ 15. $\left[-\frac{1}{9}, \frac{1}{9}\right)$ 17. $[-2, 2)$

19. $[-1, 0]$ 21. $\left(\frac{2}{3}, \frac{4}{3}\right]$

22. (a) $\lim \left(\frac{(k+1)^{k+1}}{(k+1)!} \cdot \frac{k!}{k^k} \right) = e$

    (b) $R = \frac{1}{e}$.

    (c) Need to check convergence of $\displaystyle\sum_{k=1}^{\infty} \frac{(\pm 1)^k k^k e^{-k}}{k!}$.

6.5

1. $\frac{x^2}{1-x}$; $R = 1$, center $x = 0$

3. $\frac{x+1}{1-x}$; $R = 1$, center $x = 0$

5. $\frac{x^3}{1+x^2}$; $R = 1$, center $x = 0$

7. e^{x-1}; $R = \infty$, center $x = 0$

9. $\frac{1}{x-3}$; $R = 1$, center $x = 4$

11. $\frac{3x^3-2x^4}{(1-x)^2}$; $R = 1$, center $x = 0$

13. $-\ln|1-3x|$; $R = \frac{1}{3}$, center $x = 0$

15. $x \cos x$; $R = \infty$, center $x = 0$

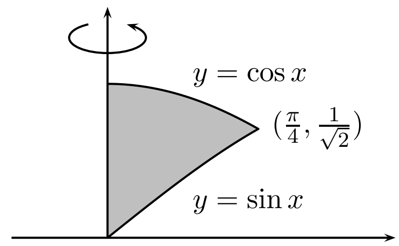


Figure B.9: Problem 5(e)

17. $\frac{1-2x^3}{1-x}$, $R = 1$, center $x = 0$

19. $-1 - \sum_{k=1}^{\infty} 2x^k$; $R = 1$, center $x = 0$.

21. $\sum_{k=0}^{\infty} x^{2k}$; $R = 1$, center $x = 0$

23. $\sum_{k=0}^{\infty} 2 \cdot (-1)^k x^{2k+1}$; $R = 1$, center $x = 0$

25. $(1 + \ln 2) + \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^k}{k \cdot 2^k}$; $R = 2$, center $x = 0$

27. $\sum_{k=0}^{\infty} \frac{(-3)^k}{k!} x^k$; $R = \infty$, center $x = 0$

29. $\sum_{k=0}^{\infty} \frac{2^k}{k!} x^{k+1}$; $R = \infty$, center $x = 0$

31. $\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+3}$; $R = \infty$, center $x = 0$

33. $e^{2x} \cos 3x = 1 + 2x - \frac{5}{2}x^2 - \frac{23}{3}x^3 - \frac{119}{24}x^4 + \ldots$

36. (a) $\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{4k+2}$

(b) $\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)! \cdot (4k+3)} x^{4k+3}$ for all x.

(c) $\frac{1}{3} - \frac{1}{42} + \frac{1}{1320} + \ldots$

6.6

1. (a) $-i$ (c) i (e) 13

(g) $\sqrt{13} + \frac{12}{\sqrt{13}}i$ (i) $\frac{1}{2} + \frac{\sqrt{3}}{2}i$ (k) $\frac{1}{2} + \frac{\sqrt{3}}{2}i$

(m) $\frac{e}{\sqrt{2}} - \frac{e}{\sqrt{2}}i$ (o) $e \cos 1 + (e \sin 1)i$

2. (a) $\sqrt{2}e^{i\pi/4}$ (c) $2e^{i\pi/6}$

3. (a) $z = \pm i$ (c) $z = 1, -\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$

4. (d)

$$\begin{aligned}\cos 5\theta &= \Re(\cos\theta + i\sin\theta)^5) \\ &= \cos^{5\theta} - 10\cos^3\theta\sin^2\theta + 5\cos\theta\sin^4\theta \\ \sin 5\theta &= \Im(\cos\theta + i\sin\theta)^5) \\ &= 5\cos^4\theta\sin\theta - 10\cos^2\theta\sin^3\theta + \sin^5\theta.\end{aligned}$$

5. (b)

$$\begin{aligned}\cos i &= \frac{e^2+1}{2e}; \\ \sin i &= -i\frac{e^2-1}{2e}; \\ \tan i &= -i\frac{e-e^{-1}}{e+e^{-1}}.\end{aligned}$$

Bibliography

[1] Niels Henrik Abel. Recherches sur la série $1 + \frac{m}{1}x + \frac{m(m-1)}{1 \cdot 2}x^2 + \frac{m(m-1)(m-2)}{1 \cdot 2 \cdot 3} + \ldots$. In L. Sylow and S. Lie, editors, *Œvres Complètes de Niels-Hendrik Abel*, volume 1, pages 219–250. Christiana, 1881.

[2] Archimedes. Measurement of a circle. In *The Works of Archimedes* [31], pages 91–98. Dover reprint, 2002.

[3] Archimedes. On the sphere and cylinder. In *The Works of Archimedes* [31], pages 1–90. Dover reprint, 2002.

[4] Margaret E. Baron. *The Origins of the Infinitesimal Calculus*. Pergamon, 1969. Dover reprint 1987, 2003.

[5] Garrett Birkhoff. *A Source Book in Classical Analysis*. Harvard Univ. Press, 1973.

[6] Bernard Bolzano. Purely analytic proof of the theorem that between any two values, which give results of opposite sign, there lies at least one real root of the equation. In *The Mathematical Works of Bernard Bolzano* [46], pages 250–277. Translation of *Rein analytischer Beweis des Lehrfasses, dass zwischen ye zwei Werthen, die ein entgegengefestes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege* (Prague, 1817).

[7] Bernard Bolzano. Theory of functions. In *The Mathematical Works of Bernard Bolzano* [46], pages 429–589. Translation of *Functionenlehre* (1830-35, unpublished).

[8] Umberto Bottazzini. *The Higher Calculus: A history of Real and Complex Analysis from Euler to Weierstrass*. Springer Verlag, 1986.

[9] Carl B. Boyer. *History of Analytic Geometry*. Yeshiva University, 1956. Dover reprint, 2004.

[10] Carl B. Boyer. *A History of Mathematics*. John Wiley & Sons, 1968.

[11] David M. Burton. *The History of Mathematics*. McGraw-Hill, 5th edition, 2003.

[12] Augustin-Louis Cauchy. *Cours d'analyse de l'école royale polytechnique*. Paris, 1821.

[13] Augustin-Louis Cauchy. Sur la développement des fonctions en séries, et sur l'integration des équations différentielles ou aux différences partielles. *Bulletin de la Societé Philomatique*, pages 49–54, 1822.

[14] J. M. Child. *The Early Mathematical Manuscripts of Leibniz*. Open Court Publ., 1920. Translations by J. M. Child of key works of Leibniz on calculus. Reprinted by Dover, 2005.

[15] I. Bernard Cohen and Anne Whitman. *Isaac Newton:The Principia (Mathematical Principles of Natural Philosophy), A new translation*. Univ. of California Press, 1999. Translation from the Latin by I. Bernard Cohen and Anne Whitman of *Philosophiae Naturalis Principia Mathematica* (1687, 1713, 1726).

[16] J. L. Coolidge. The number e. *American Mathematical Monthly*, 57:591–602, 1950. reprinted in [21, pp. 8-19].

[17] Lewis Carroll (Charles Dodgson). *Alice's Adventures in Wonderland*. Macmillan & Co., 1865. Numerous reissues.

[18] William Dunham. *The Calculus Gallery: Masterpieces from Newton to Lebesgue*. Princeton Univ. Press, 2005.

[19] Gerald A. Edgar, editor. *Classics on Fractals*. Westview Press, 2004.

[20] C. H. Edwards. *The Historical Development of the Calculus*. Springer-Verlag, 1979.

[21] Tom M. Apostol et. al., editor. *Selected Papers on Calculus*. Mathematical Association of America, 1969.

[22] Leonhard Euler. *Introduction to the Analysis of the Infinite*. Springer-Verlag, 1988. Translation by John D. Blanton of *Introductio in Analysin Infinitorum* (1748).

[23] Leonhard Euler. *Foundations of Differential Calculus.* Springer-Verlag, 2000. Translation by John D. Blanton of Chapters 1-9 of *Institutiones Calculi Differentialis* (1755).

[24] Giovanni Ferraro. *The Rise and Development of the Theory of Series Up To the Early 1820's.* Sources and Studies in the History of Mathematics and Physical Sciences. Springer, New York, 2008.

[25] David Fowler. *The Mathematics of Plato's Academy: A New Reconstruction.* Clarendon Press, Oxford, second edition, 1999.

[26] Judith V. Grabiner. *The Origins of Cauchy's Rigorous Calculus.* MIT Press, 1981. Dover reprint, 2005.

[27] I. Grattan-Guinness. *The Development of the Foundations of Mathematical Analysis from Euler to Riemann.* MIT Press, 1970.

[28] Niccolò Guicciardini. *Reading the Principia: The Debate on Newton's Mathematical Methods for Natural Philosophy from 1687 to 1736.* Cambridge University Press, 1999.

[29] Jacques Hadamard. Essai sur l'étude des fonctions données par leur développement de Taylor. *J. de Math.*, 8(4):101–186, 1892.

[30] E. Hairer and G. Wanner. *Analysis by its History.* Springer, 1995.

[31] Thomas L. Heath. *The Works of Archimedes.* Cambridge University Press, 1897, 1912. Dover reprint, 2002.

[32] Thomas L. Heath. *The Thirteen Books of Euclid's Elements.* Cambridge University Press, 1908. Dover reprint, 1956.

[33] Victor J. Katz. *A History of Mathematics, an introduction.* Addison-Wesley, second edition, 1998.

[34] Johannes Kepler. *Nova stereometria doliorum vinariorum.* Linz, 1615. A selection from this appears in [51, pp. 192-7].

[35] Karl Kiesswetter. A simple example of a function, which is everywhere continuous and nowhere differentiable. In Edgar [19], pages 340–9. Translation by Eric Olson of a 1966 article.

[36] Morris Kline. *Mathematical Thought from Ancient to Modern Times.* Oxford Univ. Press, 1972.

[37] Joseph Louis Lagrange. *Théorie des fonctions analytiques, contenant les principes du calcul différentiel, dégagés de toute considération d'infiniment petits, d'évanouissans, de limites ou de fluxions, et réduits à l'analyse algébraique des quantités finies.* l'imprimerie de la république, 1797. Second edition, Gauthier-Villars, Paris, 1813.

[38] Reinhard Laubenbacher and David Pengelley. *Mathematical Expeditions: Chronicles by the Explorers.* Undergraduate Texts in Mathematics: Readings in Mathematics. Springer, 1999.

[39] Gottfried Wilhelm Leibniz. *De quadratura arithmetica circuli ellipseos et hyperbolae cujus corollarium est trignonmetria sine tabulis.* Vandenhoeck & Ruprecht, 1993. Edited by E. Knobloch.

[40] Jesper Lützen. The foundations of analysis in the 19th century. In Hans Niels Jahnke, editor, *A History of Analysis*, chapter 6, pages 155–195. American Mathematical Society, 2003.

[41] Isaac Newton. Mathematical principles of natural philosophy. In *Isaac Newton: The Principia* [15], pages 401–944. Translation from the Latin by I. Bernard Cohen and Anne Whitman of *Philosophiae Naturalis Principia Mathematica* (1687, 1713, 1726).

[42] Fred W. Perkins. An elementary example of a continuous non-differentiable function. *American Mathematical Monthly*, 34:476–478, 1927. Reprinted in [21, pp. 137-9].

[43] Daniel Reem. New proofs of basic theorems in calculus. *Math Arxiv*, 0709.4492v1, 2007.

[44] Bernhard Riemann. Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe (on the representability of a function by means of a trignonmetric series). In Weber [53], pages 227–264. An English translation of part of this appears in [5, pp. 16-23].

[45] Ranjan Roy. The discovery of the series formula for pi by Leibniz, Gregory and Nilakantha. *Mathematics Magazine*, 63(5):291–306, 1990.

[46] Steve Russ. *The Mathematical Works of Bernard Bolzano.* Oxford Univ. Press, 2004.

[47] George F. Simmons. *Calculus Gems: Brief Lives and Memorable Mathematics.* McGraw-Hill, 1992. reissued by MAA, 2007.

[48] Donald B. Small and John M. Hosack. *Calculus, An Integrated Approach*. McGraaw-Hill, 1990.

[49] David E. Smith. *A Source Book in Mathematics*. McGraw-Hill, 1929. Dover reprint, 1959.

[50] Michael Spivak. *Calculus*. Publish or Perish, Inc., second edition, 1980.

[51] Dirk J. Struik. *A Source Book in Mathematics, 1200-1800*. Harvard Univ. Press, 1969.

[52] Otto Toeplitz. *The Calculus, A Genetic Approach*. Univ. of Chicago Press, 1963. Translation by Luise Lange of *Die Entwicklung der Infinitesimalrechnung*, unpublished manuscript from 1926-39. Reissued by MAA, 2007.

[53] Heinrich Weber, editor. *Gesammelte Mathematische Werke und Wissentschaftlicher Nachlass (Collected Mathematical Works of Bernhard Riemann)*. Dover, 1953.

[54] Karl Weierstrass. Über continuerliche Functionen eines reellen Arguments, die für keinen Werth des letzteren einen bestimmten Differentialquotienten besitzen (on continuous functions of a real argument that do not have a well-defined differential quotient). In Edgar [19], pages 2–9. Translated from the German by B. Sawhill.

[55] Derek T. Whiteside. *The Mathematical Works of Isaac Newton*. Johnson Reprint Corporation, 1964.

[56] Derek T. Whiteside. *The Mathematical Papers of Isaac Newton*. Cambridge University Press, 1967-81. 8 vols.

Index