Math 117/118:
Honours Calculus

John C. Bowman
University of Alberta
Edmonton, Canada

May 19, 2017

Contents

Preface

These notes were developed for a first-year honours-level mathematics course on differential and integral calculus at the University of Alberta. The author would like to thank the many students who took Math 117/118 from September 2000–April 2003 and September 2005–April 2007 for their help in developing these notes. Particular thanks goes to Mande Leung for typesetting the original version of these notes, to Daniel Harrison for his careful proofreading, and to Andy Hammerlindl and Tom Prince for coauthoring the high-level graphics language `Asymptote` (freely available at `http://asymptote.sourceforge.net`) that was used to draw the mathematical figures in this text. The code to lift TeX characters to three dimensions and embed them as surfaces in `PDF` files was developed in collaboration with Orest Shardt.

Chapter 1

Real Numbers

1.A Elementary Concepts from Set Theory

Definition: A *set* is a collection of *distinct* objects.

- Here are some examples of sets:
  $\{1, 2, 3\}$,
  $\{1, 2\}$,
  $\{1\}$,
  $\{\text{book}, \text{pen}\}$,
  $\mathbb{N} = \{1, 2, 3, \ldots\}$, the set of natural (counting) numbers,
  $\emptyset = \{\}$, the empty set.

Remark: Not all sets can be enumerated like this, as a (finite or infinite) list of elements. The set of real numbers is one such example.[1]

Remark: If we *can* write the elements of a set in a list, the order in which we list them is not important.

Definition: We say that a set A is a *subset* of a set B if every element of A is also an element of B. We write $A \subset B$.[2]

Definition: We say that a set A *contains* a set B if every element of B is also an element of A. We write $A \supset B$. Note that this definition implies that $B \subset A$.

---

[1] See the excellent article on countability, "How do I love thee? Let me count the ways!" by L. Marcoux, http://www.pims.math.ca/pi/issue1/page10-14.pdf, 2000.

[2] Some authors write this as $A \subseteq B$ and reserve the notation $A \subset B$ for the case where A is a subset of B but is not identical to B, that is, where A is a proper subset of B. In our notation, if we want to emphasize that A must be a proper subset of B, we explicitly write $A \subsetneq B$.

Definition: We say that two sets A and B are equal if $A \subset B$ and $B \subset A$, that is, if every element in A is also in B and vice-versa, so that A and B contain exactly the same elements. We write $A = B$.

- $\{1, 2\} = \{2, 1\}$.

Definition: The set containing all elements of A and all elements of B (but no additional elements) is called the *union* of A and B and is denoted $A \cup B$.

Definition: The set containing exactly those elements common to both A and B is called the *intersection* of A and B and is denoted $A \cap B$.

These definitions are illustrated in Figure 1.1.

- $\{1\} \cup \{2\} = \{1, 2\}$.
- $\{1, 2, 3\} \cap \{1, 4\} = \{1\}$.
- $\{1, 2\} \cup \{2\} = \{1, 2\}$.

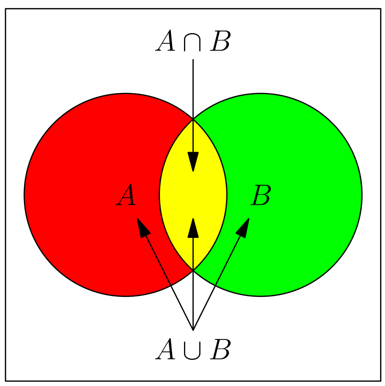


Figure 1.1: Venn Diagram

1.B Hierarchy of Sets of Numbers

We will find it useful to consider the following sets ($\in$ means *is an element of*):

$\emptyset = \{\}$ the empty set,

$\mathbb{N} = \{1, 2, 3, \ldots\}$, the set of natural (counting) numbers,

$\mathbb{Z} = \{-n : n \in \mathbb{N}\} \cup \{0\} \cup \mathbb{N}$, the set of integers,

$\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z},\ q \neq 0\}$, the set of rational numbers,

$\mathbb{R}$, the set of all real numbers.

Notice that $\emptyset \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Q. Why do we need the set $\mathbb{R}$ of real numbers to develop calculus? Why can't we just use the set $\mathbb{Q}$ of rational numbers? One might try to argue, for example, that every number representable on a (finite-precision) digital computer is rational. If a subset of $\mathbb{Q}$ is good enough for computers, shouldn't it be good enough for mathematicians, too?

To answer this question, it will be helpful to recall *Pythagoras' Theorem*, which states that the square of the length c of the hypotenuse of a right-angle triangle equals the sum of the squares of the lengths a and b of the other two sides. A simple geometric proof of this important result is illustrated in Figure 1.2. Four identical

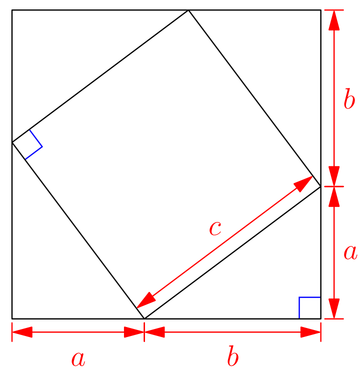
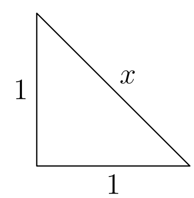


Figure 1.2: Pythagoras' Theorem

copies of the triangle, each with area $ab/2$, are placed around a square of side c, so as to form a larger square with side $a+b$. The area c^2 of the inner square is then just the area $(a+b)^2 = a^2 + 2ab + b^2$ of the large square minus the total area $2ab$ of the four triangles. That is, $c^2 = a^2 + b^2$.

Consider now the following problem. Suppose you draw a right-angle triangle having two sides of length one.



The Greek mathematicians of antiquity noticed that the length of the hypotenuse of such a triangle cannot possibly be a rational number; that is, it cannot be expressed as the ratio of two integers. Let us denote the length of the hypotenuse by x. From Pythagoras' Theorem, we know that $x^2 = 1^2 + 1^2 = 2$. Suppose that we could indeed write $x = P/Q$, where P and Q are integers (with $Q \neq 0$). By cancelling out any common integer factors greater than one, it would then always be possible to find new integers p and q that are *relatively prime* (have no common factors) such that $x = p/q$. Then

$$2 = x^2 = \frac{p^2}{q^2} \Rightarrow p^2 = 2q^2 \Rightarrow p^2 \text{ is even.}$$

If p were an odd integer, say $2n + 1$, then $p^2 = (2n + 1)^2 = 4n^2 + 4n + 1$ could not be even. Thus, p must be even: that is, $p = 2n$ for some integer n. Then

$$(2n)^2 = 2q^2 \Rightarrow 4n^2 = 2q^2 \Rightarrow 2n^2 = q^2.$$

This last result says that q^2 (and hence q) is also even, so now we know that both p and q are divisible by 2. But this contradicts the fact that p and q are relatively prime! Hence our original assumption that $x = P/Q$ must be false; that is, x cannot be represented as a rational number.

Remark: This style of mathematical proof is known as a *proof by contradiction*. By assuming that there are integers p and q such that $(p/q)^2 = 2$ we have produced two contradictory statements: p and q are relatively prime and p and q are both even.

Remark: If A and B are two statements, the notation $A{\Rightarrow}B$ says that if A holds, then B must also hold; that is, "A only if B." The notation $A{\Leftarrow}B$ says that if B holds, then A must also hold; that is, "A if B." If A and B are equivalent to each other, we write $A \iff B$, which means "A if and only if B."

Thus, the length of the hypotenuse of a right-angle triangle with unit sides cannot be expressed as a rational number. Mathematicians have invented a new number system, the real numbers, precisely to circumvent this kind of deficiency with the rational numbers $\mathbb{Q}$. The real numbers, denoted by $\mathbb{R}$, include all rational numbers plus the curious "missing" *irrational* numbers (like $\sqrt{2}$). In particular, the length of any line segment is contained in the set of real numbers. This means that there are no "holes" in the real line. Mathematicians express this fact by saying that the real numbers are *complete*.

Another important property of real numbers is that they can be written in a prescribed order on a horizontal number line, in such a way that every nonzero number is either to the right of the position occupied by the real number 0 (so that its negative is to the left of 0), or to the left of 0 (so that its negative is to the right of 0), and such that the sum and product of two numbers to the right of zero will also appear to the right of zero. Mathematicians express this particular property of the set of real numbers by saying that it can be *ordered*.

Remark: It is easy to see that the decimal expansion of a rational number must end in a repeating pattern (which could be all zeros, in which case the rational number can be represented exactly as a decimal number with a finite number of digits). When we divide the integer p by the natural number q, the remainder can only take on one of q different values, namely $0, 1, \ldots (q-1)$. If the number can be represented exactly with finitely many digits, then the decimal expansion will end with the repeating pattern $000\ldots$ (which we represent using the notation $\overline{0}$). Otherwise, we can never obtain the remainder 0, and only $q-1$ values of the remainder are possible. Upon doing q steps of long division, we will therefore encounter a repeated remainder, by the *Pigeon-Hole Principle*.[3] At the second occurrence of the repeated remainder, the pattern of digits in the quotient will then begin to repeat itself. For $q > 1$, there will never be more than $q-1$ digits in this pattern.

For example, when computing $1/7$ by long division, the pattern of quotient digits will start repeating at the second occurrence of the remainder 1. In this example, the maximum possible number of digits in the pattern, $q-1=6$, is actually achieved.

Problem 1.1: Show that the *converse* of the above remark holds; that is, if the decimal expansion of a number eventually ends in a repeating pattern of digits, the number must be rational.

Problem 1.2: Show that every real number may be approximated by a rational number as accurately as desired. This shows that the rationals densely cover the real line. We say that the rationals are *dense* in $\mathbb{R}$.

Problem 1.3: Prove that $\sqrt{3}$ is an irrational number.

Problem 1.4: Prove that $\sqrt[3]{2}$ is an irrational number.

Suppose that there existed integers p and q such that $p^3 = 2q^3$. Without loss of generality we may assume that p and q are not both even (otherwise we could cancel out the common factor of 2). We note that p^3 is even.

Express $p = 2n + r$ where $r = 0$ or 1. Then $p^3 = 8n^3 + 12n^2r + 6nr^2 + r^3$. This is even only if $r = 0$, that is, if p is even. (Alternatively, consider the prime factorization of p. Since 2 is prime, the only way it can be a factor of p^2 is if it is also a factor of p.)

Hence $8n^3 = 2q^3$, or $4n^3 = q^3$, so that q^3 is even. Replacing p by q in the above argument, we see that q is also even. This contradicts the fact that p and q are not both even.

---

[3] The *Pigeon-Hole Principle* [Fomin *et al.* 1996, pp. 31–37] (also known as *Dirichlet's Box Principle*) states that if you try to stuff more than n pigeons into n holes, at least one hole must contain two (or more) pigeons!

1.C Algebraic Properties of the Real Numbers

[Spivak 1994, pp. 3–10]

We now list the algebraic properties of the real numbers that we will use in our development of calculus.

(P1) If a, b, and c are any real numbers, then

$$a + (b + c) = (a + b) + c. \qquad \text{(associative)}$$

(P2) There is a real number 0 (the *additive identity*) such that for any real number a,

$$a + 0 = 0 + a = a. \qquad \text{(identity)}$$

Problem 1.5: Show that the additive identity is *unique*, that is, if $a + \theta = \theta + a = a$ for all a, then $\theta = 0$. Hint: set $a = \theta$ in one pair of equalities, set $a = 0$ in the other.

(P3) Every real number a has an *additive inverse* $-a$ such that

$$a + (-a) = (-a) + a = 0. \qquad \text{(inverse)}$$

Problem 1.6: Show that postulates (P1–P3) imply that every number has a unique additive inverse. That is, if $a + b = 0$, show that $b = -a$.

Definition: We define $a - b \doteq a + (-b)$. (We use the symbol $\doteq$ to emphasize a definition, although the notation := is more common.)

Problem 1.7: If $a - b = 0$, show that $a = b$.

(P4) If a and b are real numbers, then

$$a + b = b + a. \qquad \text{(commutative)}$$

Remark: Not all operations have this property. Can you give an example of an noncommutative operation?

(P5) If a, b, and c are any real numbers, then

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c. \qquad \textit{(}\text{associative}\textit{)}$$

(P6) There is a real number $1 \neq 0$ (the *multiplicative identity*) such that if a is any real number,

$$a \cdot 1 = 1 \cdot a = a. \qquad \textit{(}\text{identity}\textit{)}$$

(P7) If a, b, and c are any real numbers, then

$$a \cdot (b + c) = a \cdot b + a \cdot c. \qquad \textit{(}\text{distributive}\textit{)}$$

Remark: $a \cdot 0 = 0$ for all real a.

Proof:

$$\begin{aligned} a + a \cdot 0 &= a \cdot 1 + a \cdot 0 \\ &= a \cdot (1 + 0) \\ &= a \cdot 1 \\ &= a. \\ \therefore \quad a \cdot 0 &= 0. \end{aligned}$$

Note: the symbol $\therefore$ means *therefore*.

(P8) For any real number $a \neq 0$, there is a real number a^{-1} such that

$$a \cdot a^{-1} = a^{-1} \cdot a = 1. \qquad \textit{(}\text{inverse}\textit{)}$$

Q. Why do we restrict $a \neq 0$ here?

Problem 1.8: Show that both the multiplicative identity 1 and the multiplicative inverse a^{-1} of any real number a is unique.

(P9) If a and b are real numbers, then

$$a \cdot b = b \cdot a. \qquad \textit{(}\text{commutative}\textit{)}$$

Definition: If $a - b > 0$, we write $a > b$. Similarly, if $a - b < 0$, we write $a < b$.

(P10) Given two real numbers a and b, exactly one of the following relationships holds:

$$a < b, \quad a = b, \quad a > b. \qquad \textit{(}\text{Trichotomy Law}\textit{)}$$

(P11)

$$a > 0 \text{ and } b > 0 \quad \Rightarrow \quad a + b > 0. \qquad \textit{(}\text{closure under } +\textit{)}$$

(P12)

$$a > 0 \text{ and } b > 0 \quad \Rightarrow \quad a \cdot b > 0. \qquad \textit{(}\text{closure under } \cdot\textit{)}$$

Definition: If $a < b$ or $a = b$ we write $a \leq b$. If $a > b$ or $a = b$ we write $a \geq b$.

Q. Is it correct to write $1 \leq 2$? Why or why not?

Q. Let $x = 1$, $y = 2$. Is it correct to write $x \leq y$?

Remark: All the elementary rules of algebra and inequalities follow from these twelve properties.

- To see that $-ab = (-a)b$, we use the distributive property:

  $(-a)b + ab = (-a + a) \cdot b = 0 \cdot b = 0.$

- Likewise, we see that $(-a)(-b) = ab$

  $(-a)(-b) - ab = (-a) \cdot (-b + b) = (-a) \cdot 0 = 0,$
  so $(-a)(-b) = ab.$

- If $a < 0$ and $b < 0$, then

  $-a > 0$ and $-b > 0$
  $\Rightarrow ab = (-a)(-b) > 0.$

Remark: By setting $a = b$ in the above example and in (P12), we see that the square of any nonzero number is positive.

- If $a > b$ and $b > c$, then

  $a - b > 0$ and $b - c > 0 \;\Rightarrow\; a - c > 0$ by (P11)

  $$\Rightarrow a > c. \qquad \textit{(}\text{transitive}\textit{)}$$

- If $a > b$ and $c > 0$, then

  $a - b > 0$ and $c > 0 \Rightarrow ac - bc > 0$ by (P12) and (P7)
  i.e. $ac > bc$.

- If $a > b$ and $c < 0$, then

  $a - b > 0$ and $-c > 0 \Rightarrow -ac + bc > 0$ by (P12)
  i.e. $ac < bc$.

- If $a > b$, $c \in \mathbb{R}$, then

  $(a + c) - (b + c) = a - b > 0$
  $\Rightarrow a + c > b + c$.

- $ab > 0 \Rightarrow \begin{cases} a > 0 \text{ and } b > 0, \\ \quad\text{or} \\ a < 0 \text{ and } b < 0. \end{cases}$

  Proof: $a = 0$ or $b = 0 \Rightarrow ab = 0$ contradicts $ab > 0$.
  Also $a > 0$, $b < 0 \Rightarrow a > 0$, $-b > 0 \Rightarrow -ab > 0$ contradicts $ab > 0$.
  Likewise, $a < 0$, $b > 0$ contradicts $ab > 0$.

Problem 1.9: If $a < b$ and $c < d$, show that $a + c < b + d$.

Problem 1.10: If $0 < a < b$ and $0 < c < d$, show that $ac < bd$.

Definition: If $a < x$ and $x < b$, we write $a < x < b$ and say x is between a and b.

Lemma 1.1 (Midpoint Lemma):

$$a < b \quad \Rightarrow \quad a < \frac{a+b}{2} < b.$$

Proof:

$$a < b \quad \Rightarrow \quad a + a < a + b < b + b$$

$$\Rightarrow \quad a = \frac{a+a}{2} < \frac{a+b}{2} < \frac{b+b}{2} = b.$$

Remark: This *lemma* (small theorem) establishes that there is no least positive number. Moreover, between any two distinct numbers there exists another one.

Q. What about $1 - 0.\bar{9}$?

1.D Absolute Value

[Muldowney 1990, pp. 11–13]

The fact that for any nonzero real number either $x > 0$ or $-x > 0$ makes it convenient to define an *absolute value* function:

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

Properties: Let x and y be any real numbers.

(A1) $|x| \geq 0$.

(A2) $|x| = 0 \iff x = 0$.

(A3) $|-x| = |x|$.

(A4) $|xy| = |x|\,|y|$.

(A5) If $c \geq 0$, then

$$|x| \leq c \iff -c \leq x \leq c.$$

Proof:

$$\begin{aligned} |x| \leq c \quad &\iff \quad 0 \leq x \leq c \quad \text{or} \quad 0 < -x \leq c \\ &\iff \quad -c \leq x \leq c. \end{aligned}$$

(A6) $-|x| \leq x \leq |x|$.

Proof: Apply (A5) with $c = |x|$.

(A7)

$$\Big|\,|x| - |y|\,\Big| \leq |x \pm y| \leq |x| + |y|\,. \qquad \textit{(}\text{Triangle Inequality}\textit{)}$$

Proof:

$$\text{RHS:} \quad \textbf{(A6)} \Rightarrow \begin{cases} -|x| \leq x \leq |x| \\ -|y| \leq y \leq |y| \end{cases}$$

$$\begin{aligned} &\Rightarrow -(|x| + |y|) \leq x + y \leq |x| + |y| = c \\ \textbf{(A5)} &\Rightarrow |x + y| \leq |x| + |y|\,. \\ \text{Let } y \to -y: &\quad |x - y| \leq |x| + |y|\,. \\ \text{Thus} &\quad |x \pm y| \leq |x| + |y|\,. \end{aligned}$$

$$\text{LHS: } |x| = |(x+y) - y| \le |x+y| + |y|$$

$$\begin{aligned}
&\Rightarrow |x| - |y| \le |x+y| \\
x \leftrightarrow y: \ & |y| - |x| \le |y+x| = |x+y| \\
&\Rightarrow -|x+y| \le |x| - |y| \\
\therefore \textbf{(A5)} \Rightarrow & \Big| |x| - |y| \Big| \le |x+y| . \\
\text{Let } y \to -y: \ & \Big| |x| - |y| \Big| \le |x-y| .
\end{aligned}$$

1.E Induction

[Muldowney 1990, pp. 2–7]

Suppose that in a certain city located on the west coast of Canada, the weather office makes a long-term forecast consisting of two statements:

(A) If it rains on any given day, then it will also rain on the following day.

(B) It will rain today.

What would we conclude from these two statements? We would conclude that it will rain every single day from now on!

Or, consider a secret passed along an infinite line of people, $P_1 P_2 \ldots P_n P_{n+1} \ldots$, each of whom enjoys gossiping. If we know for every $n \in \mathbb{N}$ that P_n will always pass on a secret to P_{n+1}, then the mere act of telling a secret to the first person in line will result in everyone in the line eventually knowing the secret!

These amusing examples encapsulate the axiom of *Mathematical Induction*:

If a subset $\mathcal{S} \subset \mathbb{N}$ satisfies

(i) $1 \in \mathcal{S}$,

(ii) $k \in \mathcal{S} \Rightarrow k+1 \in \mathcal{S}$,

then $\mathcal{S} = \mathbb{N}$.

For example, suppose we wish to find the sum of the first n natural numbers. For small values of n, we could just compute the total of these n numbers directly. But for large values of n, this task could become quite time consuming! The great mathematician and physicist Carl Friedrich Gauss (1777–1855) at age 10 noticed that the rate of increase of the terms in the sum

$$1 + 2 + \ldots + n$$

could be exactly compensated by first writing the sum backwards, as

$$n + (n-1) + \ldots + 1,$$

and then averaging the two equal expressions term-by-term to obtain a sum of n identical terms:

$$\underbrace{\frac{n+1}{2} + \frac{n+1}{2} + \ldots + \frac{n+1}{2}}_{n \text{ terms}} = n\left(\frac{n+1}{2}\right).$$

We will use mathematical induction to verify Gauss' claim that

$$1 + 2 + \ldots + n \equiv \sum_{i=1}^{n} i = \frac{n(n+1)}{2}. \tag{1.1}$$

Let $\mathcal{S}$ be the set of numbers n for which Eq. (1.1) holds.

Step 1: Check $1 \in \mathcal{S}$:

$$1 = \frac{1(1+1)}{2} = 1.$$

Step 2: Suppose $k \in \mathcal{S}$, i.e.

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2}.$$

Then

$$\begin{aligned}
\sum_{i=1}^{k+1} i &= \left(\sum_{i=1}^{k} i\right) + (k+1) \\
&= \frac{k(k+1)}{2} + (k+1) \\
&= (k+1)\left(\frac{k}{2} + 1\right) \\
&= \frac{(k+1)(k+2)}{2}.
\end{aligned}$$

Hence $k+1 \in \mathcal{S}$.
That is, $k \in \mathcal{S} \Rightarrow k+1 \in \mathcal{S}$.

By the Axiom of Mathematical Induction, we know that $\mathcal{S} = \mathbb{N}$.
In other words,

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}, \quad \forall n \in \mathbb{N}.$$

Here, the symbol $\forall$ means *for all*.

- Prove that for all natural numbers n,

$$\sum_{i=1}^{n} i^3 = \left(\sum_{i=1}^{n} i\right)^2.$$

We have just seen that

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

Hence what we really want to show is that

$$\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}. \tag{1.2}$$

Step 1: We see for $n = 1$ that $1 = 1^2(1+1)^2/4$.

Step 2: Suppose

$$\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4} \doteq S_n.$$

Then

$$\begin{aligned}\sum_{i=1}^{n+1} i^3 &= \left(\sum_{i=1}^{n} i^3\right) + (n+1)^3 \\ &= \frac{n^2(n+1)^2}{4} + (n+1)^3 = \frac{(n+1)^2}{4}(n^2+4n+4) \\ &= \frac{(n+1)^2(n+2)^2}{4} = S_{n+1}.\end{aligned}$$

Hence by induction, Eq. (1.2) holds.

- If $0 < a < b$, show that

$$0 < a^n < b^n \tag{1.3}$$

for all $n \in \mathbb{N}$.

Step 1: For $n = 1$ we know that $0 < a < b$.

Step 2: Assume $0 < a^k < b^k$. On multiplying this inequality by $a > 0$ and the inequality $a < b$ by $b^k > 0$, we obtain

$$0 < a^{k+1} < b^k \cdot a < b^k \cdot b = b^{k+1},$$

from which we see that the case $n = k+1$ also holds.

$\therefore$ by induction, Eq. (1.3) holds for all n.

- (*Bernoulli Inequality*) If $a \geq -1$ then

$$(1+a)^n \geq 1 + na \quad \forall n \in \mathbb{N}.$$

Step 1: When $n = 1$, we see that

$$1 + a \geq 1 + a.$$

Step 2: Assume that the case $n = k$ holds: $(1 + a)^k \geq 1 + ka$. Then

$$\begin{aligned}(1 + a)^{k+1} \geq (1 + ka)(1 + a) &= 1 + ka + a + \underbrace{ka^2}_{\geq 0} \\ &\geq 1 + ka + a = 1 + (k + 1)a,\end{aligned}$$

so that the case $n = k + 1$ also holds.

- All students are geniuses!

  We claim that all students in any group of n students must be geniuses, for each $n = 1, 2, \ldots$.

  > Assume that the case $n = k$ holds. Given a group of $k + 1$ students, remove one of the students from the group. We know that each of the remaining k students are geniuses. Now swap one of these geniuses with the removed student. Since every student in this new group of k students are also geniuses, we deduce that all $k + 1$ students are geniuses.

  By induction, the claim holds.

Problem 1.11: Is there an error in the above "proof"? If so, where is the flaw?

- All girls have the same hair colour.

  We claim that all girls in any group of n girls have the same hair colour, for each $n = 1, 2, \ldots$.

  Step 1: When $n = 1$, there is only one girl in the group, so all girls within the group certainly have the same hair colour.

  Step 2: Assume that the case $n = k$ holds. Given a group of $k + 1$ girls, remove one of them from the group. By assumption, the remaining k girls all have the same hair colour. Now swap one of these girls with the girl we removed. Since the girls in this new group of k girls also have the same hair colour, we now know that all $k + 1$ girls have the same hair colour!

  By induction, the claim holds.

Problem 1.12: Is there an error in the above "proof"? If so, where is the flaw?

Remark: For simplicity, instead of specifying "the case $n = k$" and "the case $n = k+1$" in our induction arguments, we can simply say "for some particular n" and for "the case $n+1$". This frees up the variable k for other purposes.

Problem 1.13: Use induction to prove that $22^n - 15$ is a multiple of 7 for every natural number n.

Step 1: We see for $n = 1$ that $22 - 15 = 7$ is a multiple of 7.

Step 2: Assume for some particular n that $22^n - 15$ is a multiple of 7, say $7m$. We need only show that $22^{n+1} - 15$ is also a multiple of 7:

$$22^{n+1} - 15 = 22^n \cdot 22 - 15 = (7m + 15) \cdot 22 - 15 = 7m \cdot 22 + 15 \cdot 21 = 7(m \cdot 22 + 15 \cdot 3),$$

which is indeed a multiple of 7. By mathematical induction, we see that $22^n - 15$ is multiple of 7 for every $n \in \mathbb{N}$.

- Use induction to show that one can extract 2^n distinct subsets from any set of n elements. For example:

set	subsets
$\{\ \}$	$\{\ \}$
$\{a\}$	$\{\ \}\ \{a\}$
$\{a,b\}$	$\{\ \}\ \{a\}\ \{b\}\ \{a,b\}$
$\{a,b,c\}$	$\{\ \}\ \{a\}\ \{b\}\ \{c\}\ \{a,b\}\ \{a,c\}\ \{b,c\}\ \{a,b,c\}$

Step 1: Let $n = 1$. There are $2 = 2^1$ subsets of any such *singleton* set, namely the empty set and the set itself.

Step 2: Suppose the claim holds for some particular n.

Given any set $\mathcal{S}_{n+1}$ with $n+1$ elements, denote the collection of its subsets by $\mathcal{S}'_{n+1}$. Now remove one element from $\mathcal{S}_{n+1}$, leaving a set $\mathcal{S}_n$ containing n elements. Denote the collection of subsets of $\mathcal{S}_n$ by $\mathcal{S}'_n$. Notice that $\mathcal{S}'_n \subset \mathcal{S}'_{n+1}$. The remaining members of $\mathcal{S}'_{n+1}$ are obtained by adding the removed element to each of the sets in $\mathcal{S}'_n$.

Hence the number of sets in $\mathcal{S}'_{n+1}$ is exactly twice the number of sets in $\mathcal{S}'_n$. Given that there are 2^n members in $\mathcal{S}'_n$, we deduce that $\mathcal{S}'_{n+1}$ has 2^{n+1} members.

By induction, we see for *every* n that exactly 2^n distinct subsets can be formed from a set containing n elements.

Q. Can you think of a more direct way to establish this result?

Summation Notation

Recall

$$\sum_{k=1}^{k=n} k = 1 + 2 + \ldots + n = \frac{n(n+1)}{2}.$$

Q. What is $\sum_{k=0}^{k=n} k$?

A.

$$\sum_{k=0}^{k=n} k = 0 + \sum_{k=1}^{k=n} k = 0 + \frac{n(n+1)}{2} = \frac{n(n+1)}{2}.$$

Q. How about $\sum_{k=1}^{k=n+1} k$?

A.

$$\sum_{k=1}^{k=n+1} k = \left(\sum_{k=1}^{k=n} k\right) + (n+1) = \frac{n(n+1)}{2} + n + 1 = \frac{(n+1)(n+2)}{2}.$$

Q. How about $\sum_{k=1}^{k=n}(k+1)$?

A.

Method 1:

$$\sum_{k=1}^{k=n}(k+1) = \sum_{k=1}^{k=n} k + \sum_{k=1}^{k=n} 1 = \frac{n(n+1)}{2} + n = \frac{n(n+3)}{2}.$$

Method 2: First, let $k' = k+1$:

$$\sum_{k=1}^{k=n}(k+1) = \sum_{k'=2}^{k'=n+1} k'.$$

Next, it is convenient to replace the symbol k' with k (since it is only a dummy index anyway):

$$\sum_{k'=2}^{k'=n+1} k' = \sum_{k=2}^{k=n+1} k = \left(\sum_{k=1}^{k=n+1} k\right) - 1 = \frac{(n+1)(n+2)}{2} - 1 = \frac{n(n+3)}{2}.$$

In general,

$$\boxed{\sum_{k=L}^{k=U} a_{k+m} = \sum_{k=L+m}^{k=U+m} a_k.}$$

Verify this by writing out both sides explicitly.

Problem 1.14: For any real numbers a_1, a_2, ..., a_n, b_1, b_2, ..., b_n, and c prove that

$$\sum_{k=1}^{n} c(a_k + b_k) = c\sum_{k=1}^{n} a_k + c\sum_{k=1}^{n} b_k.$$

- *Telescoping sum*:

$$\begin{aligned}\sum_{k=1}^{n}(a_{k+1} - a_k) &= \sum_{k=1}^{n} a_{k+1} - \sum_{k=1}^{n} a_k \\ &= \sum_{k=2}^{n+1} a_k - \sum_{k=1}^{n} a_k \\ &= \cancel{\sum_{k=2}^{n}} a_k + a_{n+1} - \left(a_1 + \cancel{\sum_{k=2}^{n}} a_k\right) \\ &= a_{n+1} - a_1.\end{aligned}$$

1.F Binomial Theorem

[Muldowney 1990, pp. 8–11]

Definition:

$$\begin{aligned}&n! = 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n \quad \text{if } n \in \mathbb{N}, \\ &0! = 1.\end{aligned}$$

Equivalently, $0! = 1$ and $(k+1)! = (k+1)k!$ for $k = 0, 1, 2, \ldots$.

- $0! = 1, \;\; 1! = 1, \;\; 2! = 2, \;\; 3! = 6, \;\; 4! = 24.$

Definition: We introduce the *binomial coefficient*

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \begin{cases} 1 & \text{if } k = 0, \\ \dfrac{n(n-1)\ldots(n-k+1)}{1\cdot 2\cdot \ldots \cdot k} & \text{if } 1 \leq k \leq n. \end{cases}$$

- We find
  $\binom{3}{0} = 1,$
  $\binom{3}{1} = \frac{3}{1} = 3,$
  $\binom{3}{2} = \frac{3\cdot 2}{1\cdot 2} = 3,$
  $\binom{3}{3} = \frac{3\cdot 2\cdot 1}{1\cdot 2\cdot 3} = 1.$
- Also, $\binom{7}{3} = \frac{7\cdot 6\cdot 5}{1\cdot 2\cdot 3} = 35.$

Remark: $\binom{n}{k} = \binom{n}{n-k}$.

Remark: $\binom{n}{0} = \binom{n}{n} = 1$.

Remark: $\binom{n}{1} = \binom{n}{n-1} = n$.

Remark: If $n, k \in \mathbb{N}$, with $1 \leq k \leq n$, then

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}. \qquad \textit{(}\text{Pascal's Triangle Law}\textit{)}$$

Proof:

$$\begin{aligned} \binom{n}{k-1} + \binom{n}{k} &= \frac{n!}{(k-1)!\,(n-k+1)!} + \frac{n!}{k!\,(n-k)!} \\ &= \frac{n!}{(k-1)!\,(n-k)!}\left[\frac{1}{n-k+1} + \frac{1}{k}\right] \\ &= \frac{n!}{(k-1)!\,(n-k)!}\left[\frac{\cancel{k} + (n - \cancel{k} + 1)}{(n-k+1)k}\right] \\ &= \frac{(n+1)!}{k!\,(n-k+1)!} = \binom{n+1}{k}. \end{aligned}$$

$n\backslash k$	0 1 2 3 4		sum
0	1	1	1
1	1 1	1 1	2
2	1 2 1	or 1 2 1	4
3	1 3 3 1	1 3 3 1	8
4	1 4 6 4 1	1 4 6 4 1	16

Problem 1.15: Show that $\binom{n}{k}$ is an integer for all integers k and n satisfying $0 \le k \le n$. Hint: use induction on n and Pascal's Triangle Law. Alternatively, one can make use of the fact that $\binom{n}{k}$ is the number of distinct ways of choosing k objects from a set of n objects.

Claim: $\displaystyle\sum_{k=0}^{n} \binom{n}{k} = 2^n, \quad \forall n \in \mathbb{N} \cup \{0\}.$

Proof (by induction):

Step 1: Case $n = 0$:

$$\binom{0}{0} = 1 = 2^0.$$

Step 2: Suppose the claim holds for some particular n:

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Then

$$\begin{aligned}
\sum_{k=0}^{n+1} \binom{n+1}{k} &= \binom{n+1}{0} + \sum_{k=1}^{n} \binom{n+1}{k} + \binom{n+1}{n+1} \\
&= 1 + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] + 1 \\
&= 1 + \sum_{k=0}^{n-1} \binom{n}{k} + \sum_{k=1}^{n} \binom{n}{k} + 1.
\end{aligned}$$

Now

$$\begin{aligned}
\sum_{k=0}^{n-1} \binom{n}{k} &= \sum_{k=0}^{n} \binom{n}{k} - \binom{n}{n} = 2^n - 1, \\
\sum_{k=1}^{n} \binom{n}{k} &= \sum_{k=0}^{n} \binom{n}{k} - \binom{n}{0} = 2^n - 1.
\end{aligned}$$

Thus

$$\begin{aligned}\sum_{k=0}^{n+1}\binom{n+1}{k} &= 1+(2^n-1)+(2^n-1)+1\\ &= 2\cdot 2^n = 2^{n+1}.\end{aligned}$$

$\therefore$ the claim holds for $n+1$ as well.

Theorem 1.1 (Binomial Theorem): *For all* $n \in \mathbb{N}$,

$$\begin{aligned}(a+b)^n &= \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k\\ &= \binom{n}{0}a^n+\binom{n}{1}a^{n-1}b+\binom{n}{2}a^{n-2}b^2+\ldots+\binom{n}{n-1}ab^{n-1}+\binom{n}{n}b^n.\end{aligned} \tag{1.4}$$

Proof (by induction):

Step 1: Case $n=1$:

$$(a+b)^1=\binom{1}{0}a+\binom{1}{1}b=a+b.$$

Step 2: Suppose Eq. (1.4) holds for some n. Then

$$\begin{aligned}(a+b)^{n+1} &= (a+b)(a+b)^n\\ &= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k\\ &= \sum_{k=0}^{n}\binom{n}{k}a^{n+1-k}b^k+\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1}\\ &= \sum_{k=0}^{n}\binom{n}{k}a^{n+1-k}b^k+\sum_{k=1}^{n+1}\binom{n}{k-1}a^{n-(k-1)}b^k\\ &= \binom{n}{0}a^{n+1}+\sum_{k=1}^{n}\binom{n}{k}a^{n+1-k}b^k+\sum_{k=1}^{n}\binom{n}{k-1}a^{n-k+1}b^k+\binom{n}{n}b^{n+1}\\ &= a^{n+1}+\sum_{k=1}^{n}\left[\binom{n}{k}+\binom{n}{k-1}\right]a^{n+1-k}b^k+b^{n+1}\\ &= a^{n+1}+\sum_{k=1}^{n}\binom{n+1}{k}a^{n+1-k}b^k+b^{n+1}\\ &= \sum_{k=0}^{n+1}\binom{n+1}{k}a^{n+1-k}b^k.\end{aligned}$$

Thus, by induction, Eq. (1.4) holds for all $n \in \mathbb{N}$.

Remark: Alternative form of Binomial Theorem:

$$(a+b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2}a^{n-2}b^2 + \ldots + nab^{n-1} + b^n.$$

Remark: When $a = 1$ and $b = x$, we find

$$\begin{aligned}(1+x)^n &= 1 + nx + \frac{n(n-1)}{2}x^2 + \ldots + nx^{n-1} + x^n \\ &\geq 1 + nx \quad \text{if} \quad x \geq 0 \quad \text{and} \quad n \geq 1.\end{aligned}$$

In fact, Bernoulli's Inequality shows this is true even for $x \geq -1$.

Remark: Let $a = 1$. Then

$$\sum_{k=0}^{n} \binom{n}{k} b^k = (1+b)^n.$$

- Set $b = 1$:

$$\sum_{k=0}^{n} \binom{n}{k} = (1+1)^n = 2^n.$$

- Set $b = -1$:

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k = (1-1)^n = 0.$$

- Set $b = \frac{1}{2}$:

$$\sum_{k=0}^{n} \binom{n}{k} \frac{1}{2^k} = \left(1 + \frac{1}{2}\right)^n = \frac{3^n}{2^n}.$$

- Set $b = -\frac{1}{2}$:

$$\sum_{k=0}^{n} \binom{n}{k} \left(-\frac{1}{2}\right)^k = \left(1 - \frac{1}{2}\right)^n = \frac{1}{2^n}.$$

- Set $b = x - 1$:

$$\sum_{k=0}^{n} \binom{n}{k} (x-1)^k = (1+x-1)^n = x^n.$$

Problem 1.16: Let x be a real number. Consider the real numbers a_1, a_2, …, a_n defined by

$$a_1 = x,$$
$$a_{k+1} = \left(\frac{x-k}{k+1}\right) a_k, \quad k = 1, 2, 3 \dots .$$

(a) If x is a natural number, use induction on k to prove that

$$a_k = \begin{cases} \dbinom{x}{k} & \text{if } 1 \le k \le x, \\ 0 & \text{if } k > x. \end{cases}$$

We are given that $a_1 = x = \binom{x}{1}$. Suppose the statement holds for a particular value, k. If $k < x$ then

$$a_{k+1} = \left(\frac{x-k}{k+1}\right)\binom{x}{k} = \left(\frac{x-k}{k+1}\right)\frac{x!}{k!\,(x-k)!} = \frac{x!}{(k+1)!\,(x-k-1)!} = \binom{x}{k+1}.$$

If $k = x$ then $a_{k+1} = \left(\frac{x-k}{k+1}\right) a_k = 0$ and if $k > x$ then $a_{k+1} = \left(\frac{x-k}{k+1}\right) \cdot 0 = 0$. By induction, the desired result holds.

(b) In view of part (a), we can use a_k to provide a sensible definition for $\dbinom{x}{k}$ when x is not a natural number. Compute $\dbinom{\frac{1}{2}}{2}$, $\dbinom{\frac{1}{2}}{3}$, and $\dbinom{-1}{2005}$.

Given $a_1 = x = \frac{1}{2}$, we find

$$\binom{\frac{1}{2}}{2} = a_2 = \left(\frac{\frac{1}{2}-1}{2}\right) a_1 = -\frac{1}{8},$$

and

$$\binom{\frac{1}{2}}{3} = a_3 = \left(\frac{\frac{1}{2}-2}{3}\right) a_2 = -\frac{1}{2} a_2 = \frac{1}{16},$$

When $x = -1$ we have $a_1 = -1$ and $a_{k+1} = -a_k$, so that $a_k = (-1)^k$. Thus

$$\binom{-1}{2005} = (-1)^{2005} = -1.$$

1.G Open and Closed Intervals

[Muldowney 1990, pp. 13–14]

Let $a, b \in \mathbb{R}$ and $a < b$. There are 4 types of intervals:

$$\begin{aligned}
[a, b] &= \{x : a \leq x \leq b\}, \leftarrow \textit{closed} \text{ (contains both endpoints)} \\
(a, b) &= \{x : a < x < b\}, \leftarrow \quad \textit{open} \text{ (excludes both endpoints)} \\
[a, b) &= \{x : a \leq x < b\}, \\
(a, b] &= \{x : a < x \leq b\}.
\end{aligned}$$

It will be convenient to define also:

$$\begin{aligned}
&(-\infty, \infty) = \mathbb{R}, \\
&[a, \infty) = \{x : x \geq a\}, \\
&(a, \infty) = \{x : x > a\}, \\
&(-\infty, a] = \{x : x \leq a\}, \\
&(-\infty, a) = \{x : x < a\}.
\end{aligned}$$

However, these are not (finite) intervals.

1.H Lower and Upper Bounds

[Muldowney 1990, pp. 14–15]

Definition: Given $\mathcal{S} \subset \mathbb{R}$, we say that a real number b is an *upper bound* of $\mathcal{S}$ if

$$x \leq b \quad \text{for each} \quad x \in \mathcal{S}.$$

Q. Do all sets $\mathcal{S}$ have an upper bound?

Definition: If $\mathcal{S}$ has an upper bound we say $\mathcal{S}$ is *bounded above*. Otherwise we say $\mathcal{S}$ is *unbounded above*.

Remark: An upper bound of $\mathcal{S}$ may, or may not, be an element of $\mathcal{S}$.

Definition: Given $\mathcal{S} \subset \mathbb{R}$ we say that a real number a is a *lower bound* of $\mathcal{S}$ if

$$x \geq a \quad \text{for each} \quad x \in \mathcal{S}.$$

Definition: If $\mathcal{S}$ has a lower bound we say that $\mathcal{S}$ is *bounded below*. Otherwise we say that $\mathcal{S}$ is *unbounded below*.

Definition: If $\mathcal{S}$ is bounded above and below we say $\mathcal{S}$ is *bounded*. Otherwise, we say that $\mathcal{S}$ is *unbounded*. That is,

$$\mathcal{S} \subset \mathbb{R} \ \text{ is bounded } \iff \mathcal{S} \subset [a, b] \ \text{ for some } \ a, b \in \mathbb{R}.$$

Q. Consider the intervals $[0, 1]$ and $[0, 1)$. Do these sets have (i) an upper bound, (ii) a lower bound, (iii) an upper bound in the set itself, (iv) a lower bound in the set itself?

1.I Supremum and Infimum

[Muldowney 1990, pp. 15]

Definition: Let $\mathcal{S} \subset \mathbb{R}$. Suppose there exists a real number b such that

(i) $x \le b$ for each $x \in \mathcal{S}$ (b is an upper bound for $\mathcal{S}$),

(ii) If u is an upper bound of $\mathcal{S}$, then $b \le u$.

Then b is called the *least upper bound*, or *supremum*, of $\mathcal{S}$. We write

$$b = \text{l.u.b. } \mathcal{S} \quad \text{or} \quad b = \sup \mathcal{S}.$$

Definition: If $b = \sup \mathcal{S}$ and $b \in \mathcal{S}$, we say b is the *maximum* of $\mathcal{S}$. We write $b = \max S$.

Remark: A finite set of elements $\{a_1, a_2, \ldots, a_n\}$ always has a maximum element $\max(a_1, a_2, \ldots, a_n)$. Note that $\max(a_1, a_2, \ldots, a_n) \ge a_i$ for $i = 1, 2, \ldots, n$.

- Note that $[0, 1]$ has maximum element 1, but $[0, 1)$ has no maximum element.

Definition: Let $\mathcal{S} \subset \mathbb{R}$. Suppose there exists a real number a such that

(i) $x \ge a$ for each $x \in \mathcal{S}$ (a is a lower bound for $\mathcal{S}$),

(ii) If ℓ is a lower bound of $\mathcal{S}$, then $a \ge \ell$.

Then a is called the *greatest lower bound*, or *infimum*, of $\mathcal{S}$. We write

$$a = \text{g.l.b. } \mathcal{S} \quad \text{or} \quad a = \inf \mathcal{S}.$$

Definition: If $a = \inf \mathcal{S}$ and $a \in \mathcal{S}$, we say a is the *minimum* of $\mathcal{S}$. We write $a = \min S$.

Remark: A finite set of elements $\{a_1, a_2, \ldots, a_n\}$ always has a minimum element $\min(a_1, a_2, \ldots, a_n)$. Note that $\min(a_1, a_2, \ldots, a_n) \leq a_i$ for $i = 1, 2, \ldots, n$.

- Consider $S = \left\{ \dfrac{p}{q} : p^2 \leq 2q^2, p \in \mathbb{Z}, q \in \mathbb{N} \right\}$. The least upper bound of S is the real number $\sqrt{2}$, so S has a supremum in $\mathbb{R}$. However, $\sqrt{2} \notin \mathbb{Q}$, so the supremum of S is not itself in S; that is, S has no maximum element.

1.J Completeness Axiom

[Muldowney 1990, pp. 16]

The *completeness axiom* states that every *nonempty* subset of $\mathbb{R}$ with an upper bound has a *least* upper bound in $\mathbb{R}$.

- $\left\{ \dfrac{p}{q} : p^2 \leq 2q^2, p \in \mathbb{Z}, q \in \mathbb{N} \right\}$ has a least upper bound in $\mathbb{R}$.
- $[0, 1]$ has the supremum 1.
- $[0, 1)$ has the supremum 1.

Remark: $\emptyset = \{\}$ has no supremum. Any real number is an upper bound of the empty set, so the empty set cannot have a *least* upper bound.

Lemma 1.2 (Archimedean Property): *No real number is an upper bound for* $\mathbb{N}$.

Note: here $\mathbb{N}$ is the subset of $\mathbb{R}$ defined inductively by

(A) $1 \in \mathbb{N}$,

(B) $k \in \mathbb{N} \Rightarrow k + 1 \in \mathbb{N}$.

Proof (by contradiction):

Suppose that $\mathbb{N}$ had an upper bound. Then

$$\mathbb{N} \subset \mathbb{R},\ \mathbb{N} \neq \emptyset \Rightarrow \exists\, b = \sup \mathbb{N},$$

where b is some real number. Here the symbol $\exists$ means *there exists*.

By definition, $b = \sup \mathbb{N}$ means

(i) $b \geq k \quad \forall k \in \mathbb{N}$,

(ii) $b - 1 \not\geq k \quad \forall k \in \mathbb{N}$. That is, $b - 1 < k$ for some $k \in \mathbb{N} \Rightarrow b < k + 1$.

But $k \in \mathbb{N} \Rightarrow k + 1 \in \mathbb{N}$, so (ii) contradicts (i)!

Chapter 2

Sequences

[Muldowney 1990, Chapter 2]
[Spivak 1994, Chapter 22]

2.A Limit of a Sequence

Definition: A (real-valued) *function* f is a rule that associates a real number $f(x)$ to every number x in some subset $\mathcal{D} \subset \mathbb{R}$. The set $\mathcal{D}$ is called the *domain* of f.

Definition: The *range* $f(\mathcal{D})$ of f is the set $\{f(x) : x \in \mathcal{D}\}$.

- $f(x) = x^2$ on domain $\mathcal{D} = [0, 2)$:

  $f(\mathcal{D}) = \{x^2 : x \in [0, 2)\} = [0, 4)$.

Definition: A *sequence* is a function on the domain $\mathbb{N}$. The value of a function f at $n \in \mathbb{N}$ is often denoted by a_n,

$$a_n = f(n).$$

The consecutive function values are often written in a list:

$$\{a_n\}_{n=1}^{\infty} = \{a_1, a_2, \ldots\} \qquad \leftarrow \text{Repeated values are allowed.}$$

- $a_n = f(n) = n^2$,
  $\{a_n\}_{n=1}^{\infty} = \{1, 4, 9, 16 \ldots\}$.
- $\left\{\dfrac{(-1)^n}{n}\right\}_{n=1}^{\infty} = \left\{-1, \dfrac{1}{2}, -\dfrac{1}{3}, \dfrac{1}{4}, \ldots\right\}$.
  Notice that as n gets large, the terms of this sequence get closer and closer to zero. We say that they *converge* to 0. However, a_n is not equal to 0 for any $n \in \mathbb{N}$.

We can formalize this observation with the following concept:

Definition: The sequence $\{a_n\}_{n=1}^{\infty}$ is *convergent* with limit L if, for each $\epsilon > 0$, there exist a number N such that

$$n > N \Rightarrow |a_n - L| < \epsilon.$$

We abbreviate this as: $\lim_{n\to\infty} a_n = L$.

If no such number L exists, we say $\{a_n\}_{n=1}^{\infty}$ *diverges*.

Remark: The statement $\lim_{n\to\infty} a_n = L$ means that $|a_n - L|$ can be made as small as we please, simply by choosing n large enough.

Remark: Equivalently, as illustrated in Fig. 2.1, $\lim_{n\to\infty} a_n = L$ means that any open interval about L contains all but a finite number of terms of $\{a_n\}_{n=1}^{\infty}$.

Remark: If a sequence $\{a_n\}_{n=1}^{\infty}$ converges to L, the previous remark implies that every open interval $(L - \epsilon, L + \epsilon)$ will contain an infinite number of terms of the sequence (there cannot be only a finite number of terms inside the interval since a sequence has infinitely many terms and only finitely many of them are allowed to lie outside the interval).

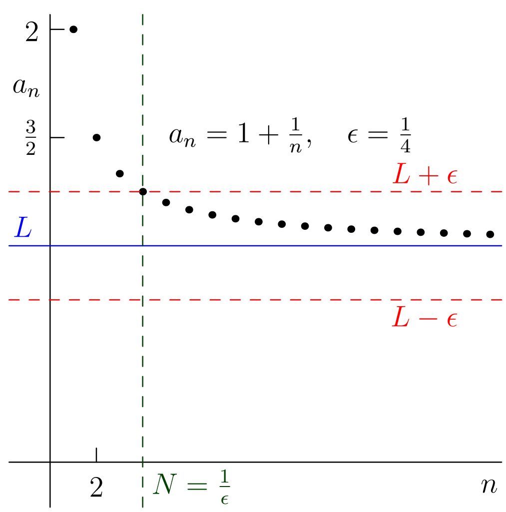


Figure 2.1: Limit of a sequence

- Let $a_n = 1, \quad \forall n \in \mathbb{N}$
  i.e. $\{1, 1, 1, \ldots\}$.
  Let $\epsilon > 0$. Choose $N = 1$.

$$n > 1 \Rightarrow |a_n - 1| = |1 - 1| = 0 < \epsilon.$$

That is, $L = 1$. Write $\lim_{n\to\infty} a_n = 1$.

Remark: Here N does not depend on ϵ, but normally it will.

- $a_n = \dfrac{(-1)^n}{n}$.

  $\lim_{n\to\infty} a_n = 0$ since $|a_n - 0| = \left|\dfrac{(-1)^n}{n}\right| = \dfrac{1}{n} < \dfrac{1}{N}$ if $n > N$.

  So, given $\epsilon > 0$, we may force $|a_n - 0| < \epsilon$ for $n > N$ simply by picking $N \geq \dfrac{1}{\epsilon}$:

$$n > N \Rightarrow |a_n - 0| < \frac{1}{N} \leq \epsilon.$$

Proposition 2.1 (Uniqueness of Limits): *If* $\lim_{n\to\infty} a_n = L_1$ *and* $\lim_{n\to\infty} a_n = L_2$*, then* $L_1 = L_2$.

Proof: Given $\epsilon > 0$, $\exists N_1, N_2 \in \mathbb{N}$ such that

$$n > N_1 \Rightarrow |a_n - L_1| < \epsilon,$$

$$n > N_2 \Rightarrow |a_n - L_2| < \epsilon.$$

Let $N = \max(N_1, N_2)$. Then

$$\begin{aligned} n > N &\Rightarrow n > N_1 \ \text{ and } \ n > N_2 \\ &\Rightarrow |a_n - L_1| < \epsilon \ \text{ and } \ |a_n - L_2| < \epsilon \\ &\Rightarrow |L_1 - L_2| = |L_1 - a_n + a_n - L_2| \leq |L_1 - a_n| + |a_n - L_2| < \epsilon + \epsilon = 2\epsilon. \end{aligned}$$

That is, given any number $a = 2\epsilon > 0$, then $|L_1 - L_2| < a$,
i.e. $|L_1 - L_2|$ is smaller than any positive number!

But we have already established from Lemma 1.1 that there is no smallest positive number, and since an absolute value can never be negative, the only choice left is

$$|L_1 - L_2| = 0 \Rightarrow L_1 = L_2.$$

Problem 2.1: Suppose $a_n \geq 0$ and $\{a_n\}_{n=1}^\infty$ is convergent with limit a. Show that $\{\sqrt{a_n}\}_{n=1}^\infty$ converges with limit $\sqrt{a}$. Hints: Do the case $a = 0$ separately. When $a > 0$, note that

$$\sqrt{a_n} - \sqrt{a} = \frac{a_n - a}{\sqrt{a_n} + \sqrt{a}}.$$

Problem 2.2: Suppose that $\{a_n\}$ and $\{b_n\}$ are convergent sequences such that $a_n < b_n$ for all $n \in \mathbb{N}$. Use a proof by contradiction to show that $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$. Can we conclude $\lim_{n\to\infty} a_n < \lim_{n\to\infty} b_n$?

Problem 2.3 (*Squeeze Principle*): Suppose $x_n \leq z_n \leq y_n$ for all $n \in \mathbb{N}$. If the sequences $\{x_n\}$ and $\{y_n\}$ both converge to the same number c, show that $\{z_n\}$ is also convergent and has limit c.

Definition: A sequence is *bounded* if there exists a number B such that

$$|a_n| \le B \quad \forall n \in \mathbb{N}.$$

Theorem 2.1 (Convergent $\Rightarrow$ Bounded): *A convergent sequence is bounded.*

Proof: Suppose $\lim\limits_{n\to\infty} a_n = L$.
Let $\epsilon = 1$.
Then $\exists N \ni$ (the symbol $\ni$ means "such that")

$$\begin{aligned} n > N \Rightarrow |a_n - L| &< 1 \\ &\Rightarrow |a_n| - |L| \le \Big| |a_n| - |L| \Big| \le |a_n - L| < 1 \\ &\Rightarrow |a_n| < 1 + |L|, \quad \forall n > N. \end{aligned}$$

Hence $|a_n| \le \max\{|a_1|, |a_2|, \ldots, |a_N|, 1 + |L|\} \doteq B$ for all $n \in \mathbb{N}$.

Remark: A bounded sequence need not be convergent.

- $\{(-1)^n\}$ is bounded since $|(-1)^n| = 1 \le B$, if we take $B = 1$. However, the sequence is not convergent:
  Suppose $\lim\limits_{n\to\infty} a_n = L$ where $a_n = (-1)^n$.
  Given $\epsilon = 1$, then for n sufficiently large, $|a_{2n} - L| = |1 - L| < 1$ and $|a_{2n+1} - L| = |-1 - L| = |1 + L| < 1$

  $$\Rightarrow 2 = |1 - L + L + 1| \le |1 - L| + |L + 1| < 1 + 1,$$

  i.e. $2 < 2$, a contradiction.

- $\lim\limits_{n\to\infty} \left(\sqrt{n+1} - \sqrt{n}\right) = 0.$
  Given $\epsilon > 0$, we can make

  $$\begin{aligned} \left|\sqrt{n+1} - \sqrt{n} - 0\right| &= \left|\sqrt{n+1} - \sqrt{n}\right| \left|\frac{\sqrt{n+1} + \sqrt{n}}{\sqrt{n+1} + \sqrt{n}}\right| \\ &= \frac{|n + 1 - n|}{\left|\sqrt{n+1} + \sqrt{n}\right|} \le \frac{1}{\sqrt{n}} < \epsilon \end{aligned}$$

  if $n > N$, as long as $\dfrac{1}{\sqrt{N}} \le \epsilon$, i.e. $N \ge \dfrac{1}{\epsilon^2}$.

Remark: By Theorem 2.1, we see that $\{\sqrt{n+1} - \sqrt{n}\}_{n=1}^{\infty}$ is bounded.

Corollary 2.1.1 (Unbounded $\Rightarrow$ Divergent): An unbounded sequence is divergent.

Remark: We say that Corollary 2.1.1 is the *contrapositive* of Theorem 2.1.

- Consider $\{n\}_{n=1}^{\infty}$. We know from the Archimedean Property that this sequence is unbounded, and hence divergent.

Problem 2.4: Suppose that $\{a_n\}_{n=1}^{\infty}$ is a bounded sequence and $\{b_n\}_{n=1}^{\infty}$ is a sequence that converges to 0. Prove that $\lim\limits_{n\to\infty} a_n b_n = 0$.

We are given that $|a_n| < B$ for some positive real number B. Given any $\epsilon > 0$, we know from the fact that $\{b_n\}_{n=1}^{\infty}$ converges to 0 that there exists a number N such that

$$n > N \Rightarrow |b_n| < \frac{\epsilon}{B}.$$

Then

$$n > N \Rightarrow |a_n b_n| < B\frac{\epsilon}{B} = \epsilon.$$

That is, $\lim\limits_{n\to\infty} a_n b_n = 0$.

Theorem 2.2 (Properties of Limits): *Let $\{a_n\}$ and $\{b_n\}$ be convergent sequences. Let $L = \lim\limits_{n\to\infty} a_n$ and $M = \lim\limits_{n\to\infty} b_n$. Then*

(a) $\lim\limits_{n\to\infty} (a_n + b_n) = L + M$;

(b) $\lim\limits_{n\to\infty} a_n b_n = LM$;

(c) $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{L}{M}$ *if* $M \neq 0$.

Proof of **(a)**: We want to show, given $\epsilon > 0$, that

$$|a_n + b_n - L - M| < \epsilon \tag{2.1}$$

for all sufficiently large n. Since the Triangle Inequality tells us that

$$|a_n + b_n - L - M| \le |a_n - L| + |b_n - M|,$$

it is enough to show that

$$|a_n - L| + |b_n - M| < \epsilon \tag{2.2}$$

for all sufficiently large n.

We know $\exists N_1, N_2 \ni$

$$n > N_1 \Rightarrow |a_n - L| < \frac{\epsilon}{2},$$

$$n > N_2 \Rightarrow |b_n - M| < \frac{\epsilon}{2}.$$

When $n > \max(N_1, N_2)$ then Eq. (2.2) will hold, which in turn implies Eq. (2.1). Hence

$$\lim_{n\to\infty}(a_n + b_n) = L + M.$$

Proof of **(b)**: Note that

$$\begin{aligned}|a_nb_n - LM| &= |a_nb_n - Lb_n + Lb_n - LM| \\ &\le |a_nb_n - Lb_n| + |Lb_n - LM| \\ &\le |a_n - L|\,|b_n| + |b_n - M|\,|L|\,.\end{aligned}$$

Theorem 2.1 $\Rightarrow \{b_n\}$ is bounded $\Rightarrow \exists B > 0 \ni |b_n| \le B \quad \forall n$.

$$\therefore |a_nb_n - LM| \le |a_n - L|\,B + |b_n - M|\,|L|\,.$$

Given $\epsilon > 0$, let $\epsilon_0 = \dfrac{\epsilon}{B + |L|} > 0$. There exists N_1, N_2 such that

$$n > N_1 \Rightarrow |a_n - L| < \epsilon_0,$$

$$n > N_2 \Rightarrow |b_n - M| < \epsilon_0.$$

So $n > \max(N_1, N_2) \Rightarrow$

$$|a_nb_n - LM| \le \epsilon_0 B + \epsilon_0\,|L| = \epsilon_0(B + |L|) = \epsilon.$$

That is,

$$\lim_{n\to\infty} a_nb_n = LM.$$

Proof of **(c)**: We only need to prove the special case

$$\lim_{n\to\infty} b_n = M \Rightarrow \lim_{n\to\infty}\frac{1}{b_n} = \frac{1}{M} \quad \text{if} \quad M \ne 0, \tag{2.3}$$

for **(c)** would then follow from **(b)**:

$$\lim_{n\to\infty}\frac{a_n}{b_n} = \lim_{n\to\infty}\left(a_n \cdot \frac{1}{b_n}\right) = \lim_{n\to\infty} a_n \lim_{n\to\infty}\frac{1}{b_n} = \frac{L}{M} \quad \text{if} \quad M \ne 0.$$

To prove Eq. (2.3), consider

$$\left|\frac{1}{b_n} - \frac{1}{M}\right| = \left|\frac{M - b_n}{b_nM}\right| = \frac{1}{|b_n|}\cdot\frac{|b_n - M|}{|M|}.$$

We know that there exists a number N_1 such that

$$\begin{aligned}n > N_1 &\Rightarrow |b_n - M| < \frac{|M|}{2} \\ &\Rightarrow |M| - |b_n| \le |b_n - M| < \frac{|M|}{2} \\ &\Rightarrow 0 < \frac{|M|}{2} < |b_n| \Rightarrow \frac{1}{|b_n|} < \frac{2}{|M|}.\end{aligned}$$

Thus

$$n > N_1 \Rightarrow \left| \frac{1}{b_n} - \frac{1}{M} \right| < \frac{2}{|M|^2} |b_n - M|.$$

Now given $\epsilon > 0$, there exists a number $N_2 \ni$

$$n > N_2 \Rightarrow |b_n - M| < \frac{|M|^2}{2}\epsilon \text{ since } M \neq 0.$$

Hence $n > \max(N_1, N_2) \Rightarrow$

$$\left| \frac{1}{b_n} - \frac{1}{M} \right| < \frac{2}{|M|^2} \frac{|M|^2}{2}\epsilon = \epsilon.$$

Thus

$$\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{M}.$$

- $$\lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = \lim_{n\to\infty} 1 + \lim_{n\to\infty} \frac{1}{n}.$$

- $$\lim_{n\to\infty} \frac{n}{n+1} = \lim_{n\to\infty} \frac{1}{1+\frac{1}{n}} = 1.$$

- $$\begin{aligned} \lim_{n\to\infty} \frac{2n^2+1}{3n^2+100n+2} &= \lim_{n\to\infty} \frac{2+\frac{1}{n^2}}{3+\frac{100}{n}+\frac{2}{n^2}} \\ &= \frac{\lim_{n\to\infty}(2+\frac{1}{n^2})}{\lim_{n\to\infty}(3+\frac{100}{n}+\frac{2}{n^2})} = \frac{2}{3}. \end{aligned}$$

- $$\lim_{n\to\infty} \frac{a_m n^m + a_{m-1}n^{m-1} + \ldots + a_0}{b_m n^m + b_{m-1}n^{m-1} + \ldots + b_0} =$$

  $$\lim_{n\to\infty} \frac{a_m + a_{m-1}\frac{1}{n} + a_{m-2}\frac{1}{n^2} + \ldots + \frac{a_0}{n^m}}{b_m + b_{m-1}\frac{1}{n} + b_{m-2}\frac{1}{n^2} + \ldots + \frac{b_0}{n^m}} = \frac{a_m}{b_m}.$$

- $$\lim_{n\to\infty} \frac{n^2+1}{n^3+3} = \lim_{n\to\infty} \frac{\frac{1}{n}+\frac{1}{n^3}}{1+\frac{3}{n^3}} = \frac{\lim_{n\to\infty}(\frac{1}{n}+\frac{1}{n^3})}{\lim_{n\to\infty}(1+\frac{3}{n^3})} = \frac{0}{1} = 0.$$

- $$\lim_{n\to\infty} \frac{n^3+3}{n^2+1} = \lim_{n\to\infty} \frac{1+\frac{3}{n^3}}{\frac{1}{n}+\frac{1}{n^3}} \neq \frac{\lim_{n\to\infty}(1+\frac{3}{n^3}) \to 1}{\lim_{n\to\infty}(\frac{1}{n}+\frac{1}{n^3}) \to 0}.$$

Corollary 2.2.1 (Case $L \neq 0$, $M = 0$): Let $\{a_n\}$ and $\{b_n\}$ be convergent sequences. If $\lim\limits_{n\to\infty} a_n \neq 0$ and $\lim\limits_{n\to\infty} b_n = 0$, then $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n}$ does not exist.

Proof: Suppose $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = K$.

Then Theorem 2.2 (b) $\Rightarrow \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n \cdot \lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = 0 \cdot K = 0$.

This contradicts the fact that $\lim\limits_{n\to\infty} a_n \neq 0$.

- Returning to the previous example, we see that
$$\lim_{n\to\infty} \frac{1+\frac{3}{n^3}}{\frac{1}{n}+\frac{1}{n^3}} \nexists.$$

- $$\lim_{n\to\infty} \frac{n^4+8n^3+6n+1}{1000n^3+3n^2+2} = \lim_{n\to\infty} \frac{1+\frac{8}{n}+\frac{6}{n^3}+\frac{1}{n^4}}{\frac{1000}{n}+\frac{3}{n^2}+\frac{2}{n^4}} \nexists.$$

  In general, if $a_m \neq 0$ and $b_k \neq 0$, then
$$\lim_{n\to\infty} \frac{a_m n^m + \ldots + a_0}{b_k n^k + \ldots + b_0} = \begin{cases} 0 & \text{if } m < k, \\ \frac{a_m}{b_m} & \text{if } m = k, \\ \nexists & \text{if } m > k. \end{cases}$$

Remark: We cannot use Theorem 2.2 to say
$$1 = \lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty} \frac{1}{n} \cdot \lim_{n\to\infty}(n+1) = 0 \cdot \lim_{n\to\infty}(n+1)$$
because $\lim_{n\to\infty}(n+1)$ does not exist. Whenever we use a theorem, we must be very careful to check that the conditions of the theorem apply!

Problem 2.5: Determine which of the following limits exist. For those that exist, compute the limit. Show your calculations.

(a)
$$\lim_{n\to\infty} \frac{-7n^4-n^2+6}{2n^4+n+12}$$
$$= \lim_{n\to\infty} \frac{-7-\frac{1}{n^2}+\frac{6}{n^4}}{2+\frac{1}{n^3}+\frac{12}{n^4}} = \frac{\lim\limits_{n\to\infty}\left(-7-\dfrac{1}{n^2}+\dfrac{6}{n^4}\right)}{\lim\limits_{n\to\infty}\left(2+\dfrac{1}{n^3}+\dfrac{12}{n^4}\right)} = -\frac{7}{2}.$$

(b)

$$\lim_{n\to\infty} \frac{3n^3 + \frac{1}{n}}{2n^4}$$

$$= \lim_{n\to\infty} \frac{\frac{3}{n^4} + \frac{1}{n^5}}{2} = 0.$$

(c)

$$\lim_{n\to\infty} \frac{3n^3 + \frac{1}{n}}{2n^2}$$

The limit does not exist by Corollary 2.2.1 since

$$\lim_{n\to\infty} \left(3 + \frac{1}{n^4}\right) = 3 \neq 0$$

and

$$\lim_{n\to\infty} \frac{2}{n} = 0.$$

2.B Monotone Sequences

Definition: A sequence $\{a_n\}_{n=1}^{\infty}$ is *increasing* if

$$a_1 \leq a_2 \leq a_3 \leq \dots, \quad \text{i.e.} \quad a_n \leq a_{n+1} \quad \forall n \in \mathbb{N}$$

and *decreasing* if

$$a_1 \geq a_2 \geq a_3 \geq \dots, \quad \text{i.e.} \quad a_n \geq a_{n+1} \quad \forall n \in \mathbb{N}.$$

Q. By the above definition, is it possible for a sequence to be both increasing *and* decreasing?

Definition: A sequence is *monotone* if it is either (i) an increasing sequence or (ii) a decreasing sequence.

Definition: A sequence $\{a_n\}_{n=1}^{\infty}$ is *strictly increasing* (*strictly decreasing*) if

$$a_1 < a_2 < a_3 < \dots$$
$$(a_1 > a_2 > a_3 > \dots).$$

- $\{n\}$, the sequence of all natural numbers, is strictly increasing.

- $\{2n\}$, $\{2n+1\}$, $\{n^2\}$, and $\{100, 101, \ldots\}$ are all strictly increasing sequences.

Recall convergent $\Rightarrow$ bounded. When does bounded $\Rightarrow$ convergent? The next theorem addresses this question in the special case of monotone sequences.

Theorem 2.3 (Monotone Sequences: Convergent $\iff$ Bounded): *Let $\{a_n\}$ be a monotone sequence. Then $\{a_n\}$ is convergent $\iff$ $\{a_n\}$ is bounded.*

Proof:

"$\Rightarrow$" Let $\{a_n\}$ be convergent. Then $\{a_n\}$ is bounded by Theorem 2.1.

"$\Leftarrow$" Suppose $\{a_n\}$ is increasing and bounded. Let $L = \sup\{a_n : n = 1, 2, \ldots\}$ (why does this always exist?) We show that $\lim_{n\to\infty} a_n = L$. Given $\epsilon > 0$, then $L - \epsilon$ is not an upper bound of $\{a_n : n = 1, 2, \ldots\}$ by the definition of a supremum. That is, $\exists$ element $a_N \ni$

$$L - \epsilon < a_N.$$

Now $n > N \Rightarrow a_n \geq a_N$ (why?). Thus

$$\begin{aligned} n > N &\Rightarrow L - \epsilon < a_N \leq a_n \leq L \Rightarrow L - \epsilon < a_n < L + \epsilon \\ &\Rightarrow |a_n - L| < \epsilon. \end{aligned}$$

Hence $\lim_{n\to\infty} a_n = L$.

The proof for the case where $\{a_n\}$ is decreasing and bounded is similar.

Problem 2.6: Complete the above proof for the case of a decreasing bounded sequence $\{a_n\}$. Suggestion: consider the sequence $\{-a_n\}$.

Remark: We see from the proof of Theorem 2.3 that an increasing bounded sequence converges to its supremum, whereas a decreasing bounded sequence converges to its infimum.

- $\left\{1 - \dfrac{1}{n} : n = 1, 2, \ldots\right\}$ is increasing and bounded.

  $\therefore \lim_{n\to\infty}\left(1 - \dfrac{1}{n}\right) \exists.$

  Of course, we already knew from Theorem 2.2 (a) that

$$\lim_{n\to\infty}\left(1 - \frac{1}{n}\right) = \lim_{n\to\infty} 1 - \lim_{n\to\infty}\frac{1}{n} = 1 - 0 = 1.$$

- If

$$a_n = \left(1 + \frac{1}{n}\right)^n,$$

show that $\{a_n\}_{n=1}^{\infty}$ is a convergent sequence.

In the Binomial Theorem $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$, set $x = 1$ and $y = 1/n$. Then

$$\begin{aligned}
a_n = \left(1+\frac{1}{n}\right)^n &= 1 + \frac{n}{1!}\left(\frac{1}{n}\right)^1 + \frac{n(n-1)}{2!}\left(\frac{1}{n}\right)^2 + \frac{n(n-1)(n-2)}{3!}\left(\frac{1}{n}\right)^3 \\
&\quad + \ldots + \frac{n(n-1)\ldots 3\cdot 2}{(n-1)!}\left(\frac{1}{n}\right)^{n-1} + \frac{n!}{n!}\left(\frac{1}{n}\right)^n \\
&= 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) \\
&\quad + \ldots + \frac{1}{(n-1)!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{n-2}{n}\right) \\
&\quad + \frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{n-1}{n}\right).
\end{aligned}$$

This expression for a_n can be used to establish two key properties.

Claim: $\{a_n\}$ is **(strictly) increasing.**

$$\begin{aligned}
a_{n+1} &= 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n+1}\right) + \frac{1}{3!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right) \\
&\quad + \ldots + \frac{1}{n!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\ldots\left(1-\frac{n-1}{n+1}\right) \\
&\quad + \underbrace{\frac{1}{(n+1)!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\ldots\left(1-\frac{n}{n+1}\right)}_{\text{positive}} > a_n.
\end{aligned}$$

Claim: $\{a_n\}$ is **bounded.**

$$\begin{aligned}
2 \le a_n &\le 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \left(\frac{1}{4!} + \ldots + \frac{1}{n!}\right) \\
&\le 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \left(\frac{1}{2^4} + \ldots + \frac{1}{2^n}\right) \quad \text{since } n! \ge 2^n \text{ for } n \ge 4 \text{ (induction)} \\
&\le 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \left(\frac{1}{2^4} + \ldots + \frac{1}{2^n}\right) \\
&\le 1 + \frac{1-(\frac{1}{2})^{n+1}}{1-\frac{1}{2}} < 1 + \frac{1}{\frac{1}{2}} = 3.
\end{aligned}$$

Thus, by Theorem 2.3, $\{a_n\}$ converges. We define

$$e \doteq \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n.$$

The above argument shows that $2 \le e \le 3$. In fact $e \approx 2.718281828459\ldots$.

2.C Subsequences

Definition: Given a sequence $\{a_n\}_{n=1}^{\infty}$ and a strictly increasing sequence of natural numbers $\{n_k\}_{k=1}^{\infty}$, we can form the *subsequence* $\{a_{n_k}\}_{k=1}^{\infty}$ of $\{a_n\}_{n=1}^{\infty}$.

- $\{(2k-1)^2\}_{k=1}^{\infty} = \{1, 9, 25, \ldots\}$ is a subsequence of $\{n^2\}_{n=1}^{\infty} = \{1, 4, 9, 16, 25, \ldots\}$.

Remark: Note that $n_k \ge k$ for all k. This is easily proven by induction: $n_1 \ge 1$ and if $n_k \ge k$ then $n_{k+1} > n_k \ge k$, so that $n_{k+1} \ge k+1$.

- $\{\frac{1}{100}, \frac{1}{101}, \ldots\}$ is a subsequence of $\{\frac{1}{n}\}_{n=1}^{\infty}$.
  So is $\{\frac{1}{100}, \frac{1}{102}, \ldots\}$.
  So is $\{\frac{1}{n}\}_{n=1}^{\infty}$ itself (here $n_k = k$).

Theorem 2.4 (Convergent $\iff$ All Subsequences Convergent): *A sequence $\{a_n\}_{n=1}^{\infty}$ is convergent with limit L $\iff$* **each** *subsequence $\{a_{n_k}\}_{k=1}^{\infty}$ of $\{a_n\}_{n=1}^{\infty}$ is convergent with limit L.*

Proof:

"$\Rightarrow$" Suppose $\{a_n\}$ is convergent with limit L. Given $\epsilon > 0$, $\exists N \ni$

$$n > N \Rightarrow |a_n - L| < \epsilon. \tag{2.4}$$

Let $\{a_{n_k}\}$ be a subsequence of $\{a_n\}$. Then

$$k > N \Rightarrow n_k \ge k > N.$$

Considering only the indices $n = n_k$ in Eq. (2.4) $\Rightarrow |a_{n_k} - L| < \epsilon$.

$$\therefore \lim_{k\to\infty} a_{n_k} = L.$$

"$\Leftarrow$" Suppose each subsequence of $\{a_n\}$ is convergent with limit L. Note that $\{a_n\}$ is a subsequence of itself. Hence $\{a_n\}$ is convergent with limit L.

- $\{(-1)^n\}$ is not convergent since $\begin{array}{l}\{(-1)^{2n}\} = \{1, 1, 1, \ldots\} \to 1 \\ \{(-1)^{2n+1}\} = \{-1, -1, -1, \ldots\} \to -1\end{array}$
  and $-1 \ne 1$.

In the examples to follow, we will make use of a lemma.

Lemma 2.1: *For all $n \in \mathbb{N}$,*

(a) $0 \leq c < 1 \Rightarrow c^n \leq c < 1$,

(b) $c > 1 \Rightarrow c^n \geq c > 1$.

Proof:

(a) Holds for $n = 1$; $c^n < 1 \Rightarrow c^{n+1} = cc^n \leq c \cdot 1 < 1$.

(b) Holds for $n = 1$; $c^n > 1 \Rightarrow c^{n+1} = cc^n > c \cdot 1 > 1$.

- $\lim_{n\to\infty} c^n = \begin{cases} 0 & \text{if } \ 0 \leq c < 1, \\ 1 & \text{if } \ c = 1, \\ \nexists & \text{if } \ c > 1 \text{ (divergent; in fact, unbounded (exercise))} \end{cases}$.

Case $0 \leq c < 1$: We have $0 \leq c^{n+1} \leq c^n \leq c < 1, \quad \forall n \geq 1$.

$\Rightarrow \{c^n\}_{n=1}^{\infty}$ is a decreasing, bounded sequence.

$\therefore\ L = \lim_{n\to\infty} c^n\ \exists$ and $0 \leq L \leq c < 1$.

$\{c^{n+1}\}_{n=1}^{\infty}$ is a subsequence of $\{c^n\}_{n=1}^{\infty}$, so $\lim_{n\to\infty} c^{n+1} = L$. But $c^{n+1} = c\,c^n$, so Theorem 2.2 (b) implies

$$L = \lim_{n\to\infty} c^{n+1} = c \lim_{n\to\infty} c^n = c\,L$$

$$\Rightarrow L(1-c) = 0 \Rightarrow L = 0 \quad \text{(why?)}$$

.

Case $c = 1$: $\lim_{n\to\infty} c^n = \lim_{n\to\infty} 1 = 1$.

Case $c > 1$: If $\{c^n\}$ were convergent, we could let $L = \lim_{n\to\infty} c^n$. Then

$$L = \lim_{n\to\infty} c^{n+1} = c \lim_{n\to\infty} c^n = c\,L \Rightarrow L(1-c) = 0 \Rightarrow L = 0.$$

This would contradict $c^n > 1$ for all $n \in \mathbb{N}$, which requires that $L \geq 1$.

Hence $\{c^n\}$ is divergent.

- $\lim_{n\to\infty} \frac{1}{2^n} = \lim_{n\to\infty} \left(\frac{1}{2}\right)^n = 0$.

Definition: If $c > 0$, let $c^{1/n}$ denote the n^{th} positive *root* of c (which we will soon see always exists). This real root is unique: consider

$$x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \ldots + xy^{n-2} + y^{n-1})$$

and observe that, if x and y are both positive, $x^n - y^n = 0 \Rightarrow x = y$. The above factorization may be established using a Telescoping sum:

$$\begin{aligned}(x-y)\sum_{k=0}^{n-1} x^{n-1-k}y^k &= \sum_{k=0}^{n-1} x^{n-k}y^k - \sum_{k=0}^{n-1} x^{n-1-k}y^{k+1} = \sum_{k=0}^{n-1} x^{n-k}y^k - \sum_{k=1}^{n} x^{n-k}y^k \\ &= x^n - y^n.\end{aligned}$$

An immediate corollary of Lemma 2.1 is

Lemma 2.2: *For all $n \in \mathbb{N}$,*

(a) $0 \le c < 1 \Rightarrow c \le c^{1/n} < 1$,

(b) $c > 1 \Rightarrow c \ge c^{1/n} > 1$.

- We now establish that $\lim\limits_{n\to\infty} c^{1/n} = 1$ for any $c > 0$.

Case $c > 1$: Note that $1 < c^{1/n} \Rightarrow c < cc^{1/n} = c^{(n+1)/n}$.

Let $a_n = c^{1/n}$. Then

$$\begin{aligned}&a_{n+1}^{n+1} = \left(c^{1/(n+1)}\right)^{n+1} = c < c^{(n+1)/n} = a_n^{n+1} \\ &\Rightarrow \left(\frac{a_{n+1}}{a_n}\right)^{n+1} < 1 \Rightarrow \frac{a_{n+1}}{a_n} < 1 \\ &\Rightarrow \{a_n\} \text{ decreasing and bounded: } 1 < a_{n+1} < a_n \le a_1 = c.\end{aligned}$$

Hence $\lim\limits_{n\to\infty} c^{1/n} \exists = L$. But then $\lim\limits_{n\to\infty} c^{1/(2n)} = L$.

Note that $c^{1/n} > 1$ for all $n \in \mathbb{N}$ implies that $L \ge 1$.

But $\lim\limits_{n\to\infty} c^{1/(2n)} \lim\limits_{n\to\infty} c^{1/(2n)} = \lim\limits_{n\to\infty} c^{1/n}$

$$\begin{aligned}&\Rightarrow L \cdot L = L \\ &\Rightarrow L(L-1) = 0.\end{aligned}$$

Hence $L \ne 0 \Rightarrow L = 1$.

Case $c = 1$: $\lim\limits_{n\to\infty} c^{1/n} = \lim\limits_{n\to\infty} 1 = 1$.

Case $0 < c < 1$**:** Let $k = c^{-1} \Rightarrow c = k^{-1}$

$$\Rightarrow c^{1/n} = k^{-1/n} = \frac{1}{k^{1/n}}.$$

Then $k > 1 \Rightarrow \lim_{n\to\infty} k^{1/n} = 1$.

Theorem 2.2 (c) $\Rightarrow \lim_{n\to\infty} c^{1/n} = \frac{1}{\lim_{n\to\infty} k^{1/n}} = \frac{1}{1} = 1$.

- The sequence $a_n = \frac{n}{2^n}$ converges.

$$\frac{a_{n+1}}{a_n} = \left(\frac{n+1}{2^{n+1}}\right)\left(\frac{2^n}{n}\right) = \frac{n+1}{2n}$$

$\Rightarrow 0 < a_{n+1} \le a_n$ since $n+1 \le 2n \; \forall n \in \mathbb{N}$.
$\therefore \; \{a_n\}$ is convergent. Its limit can be found from the observation that

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \left(\frac{n+1}{2n}\right) a_n = \frac{1}{2} \lim_{n\to\infty} a_n.$$

Hence $\lim_{n\to\infty} a_n = 0$.

- Consider the sequence

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{A}{x_n}\right), \qquad A > 0, x_1 > 0.$$

Suppose $p = \lim_{n\to\infty} x_n \exists$. Then $p = \frac{1}{2}p + \frac{A}{2p} \Rightarrow p^2 = A$. This sequence can therefore be used as an algorithm for computing square roots. To show that it actually converges, consider

$$x_n > 0 \Rightarrow x_{n+1} > 0 \quad \forall n.$$

$$\therefore x_1 > 0 \Rightarrow x_n > 0 \quad \forall n.$$

Consider

$$x_{n+1} - x_n = -\frac{1}{2}x_n + \frac{A}{2x_n} = \frac{A - x_n^2}{2x_n}.$$

Now, for all $n \in \mathbb{N}$,

$$x_{n+1}^2 - A = \frac{1}{4}\left(x_n^2 + 2A + \frac{A^2}{x_n^2}\right) - A = \frac{1}{4}\left(x_n - \frac{A}{x_n}\right)^2 \ge 0.$$

$$\therefore x_n \ge \sqrt{A}, \quad n = 2, 3, \ldots$$

$$\Rightarrow x_{n+1} - x_n \le 0, \quad n = 2, 3, \ldots$$

Since $\{x_{n+1}\}_{n=1}^{\infty}$ is decreasing and bounded below by $\sqrt{A} > 0$, we know that $\{x_{n+1}\}_{n=1}^{\infty}$, and hence $\{x_n\}_{n=1}^{\infty}$, is convergent.

Problem 2.7: Consider the sequence $\{a_n\}_{n=1}^{\infty}$ defined inductively by $a_1 = 0$, and $a_{n+1} = \sqrt{a_n + 2}$ for $n \geq 1$.

(a) Prove that $\{a_n\}_{n=1}^{\infty}$ is increasing.
Consider

$$a_{n+1}^2 - a_n^2 = a_n + 2 - a_n^2 = -(a_n^2 - a_n - 2) = -(a_n - 2)(a_n + 1).$$

Note that $a_n \geq 0$ for all n. For $a_n \in [-1, 2]$ we see that $-(a_n - 2)(a_n + 1) \geq 0$. The initial element $a_1 = 0$ belongs to the interval $[0, 2]$. Therefore, if we can show that subsequent values of a_n are less than or equal to 2, then $a_n \in [0, 2]$ for all n, so that

$$0 \leq a_{n+1}^2 - a_n^2 = (a_{n+1} - a_n)(a_{n+1} + a_n).$$

It then follows that $a_{n+1} - a_n \geq 0$; that is, $\{a_n\}_{n=1}^{\infty}$ is an increasing sequence.

We now use induction to establish our claim that

$$a_n \leq 2 \quad \forall n.$$

Step 1:

$$a_1 = 0 \leq 2.$$

Step 2: Suppose that the claim holds for some integer n. Then

$$a_{n+1}^2 = a_n + 2 \leq 2 + 2 = 4.$$

Hence $a_{n+1} \leq 2$. By induction, the claim holds for all natural numbers n.

(b) Prove that $\{a_n\}_{n=1}^{\infty}$ converges.

Since $\{a_n\}_{n=1}^{\infty}$ is an increasing sequence, bounded below by $a_1 = 0$ and above by 2, we deduce that $\{a_n\}_{n=1}^{\infty}$ converges.

(c) Find $\lim\limits_{n\to\infty} a_n$. Justify each step in your argument.

From part (a) we know that $L = \lim\limits_{n\to\infty} a_n$ exists. Since $a_n \geq 0$ for all n we know that $L \geq 0$ (otherwise, if $L < 0$, a contradiction would result upon considering the case $\epsilon = -L$ in the definition of a limit). We know that each subsequence of $\{a_n\}_{n=1}^{\infty}$ is convergent to the same limit L. From the properties of limits we know that

$$L^2 = \lim_{n\to\infty} a_{n+1}^2 = \lim_{n\to\infty} (a_n + 2) = L + 2.$$

Of the possible solutions to $L^2 - L - 2 = (L - 2)(L + 1) = 0$, which are $L = 2$ and $L = -1$, only the solution $L = 2$ satisfies $L \geq 0$. Hence

$$\lim_{n\to\infty} a_n = 2.$$

Problem 2.8 (Sequence Limit Ratio Test)): Let $\{a_n\}_{n=1}^{\infty}$ be a sequence such that $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = r$, where $r \in [0, 1)$. Let $s = (1 + r)/2$, so that $0 \le r < s < 1$.

(a) Show that there exists a number N such that $n \ge N \Rightarrow \left|\frac{a_{n+1}}{a_n}\right| \le s$. Hint: Consider $\epsilon = s - r > 0$.

(b) Use part (a) and induction to show for $n \ge N$ that

$$0 \le |a_n|\, s^N \le |a_N|\, s^n.$$

(c) Prove that $\lim_{n\to\infty} a_n = 0$.

(d) Apply this result to prove for any $x \in \mathbb{R}$ that

$$\lim_{n\to\infty} \frac{x^n}{n!} = 0,$$

(e) If $r > 1$, show that $\{a_n\}$ is divergent. Hint: consider the sequence $\{1/a_n\}$.

(f) If $r = 1$, give examples to illustrate that $\{a_n\}$ may be either convergent or divergent.

Problem 2.9: Let r be a real number. Consider the sequence $\{S_n\}_{n=0}^{\infty}$, where S_n is the *partial sum* of the *geometric series*

$$S_n = \sum_{k=0}^{n} r^k.$$

For what values of r does $S = \lim_{n\to\infty} S_n$ exist? Compute S (when the limit exists) in terms of r. When the limit exists, we say that the *infinite series*

$$\sum_{k=0}^{\infty} r^k$$

converges and has limit S. Hint: Consider the telescoping sum $rS_n - S_n$.

Problem 2.10: Compute the sum

$$\sum_{k=1}^{\infty} 9 \times 10^{-k}.$$

2.D Bolzano–Weierstrass Theorem

The converse of Theorem 2.1 (convergent $\Rightarrow$ bounded) does not necessarily hold for a nonmonotonic sequence. Nevertheless, Theorem 2.1 does have the following partial converse:

Theorem 2.5 (Bolzano–Weierstrass Theorem): *A bounded sequence has a convergent subsequence.*

Proof: Let $\{a_n\}_{n=1}^{\infty}$ be a bounded sequence. Then there exists real numbers A and B such that

$$a_n \in [A, B] \quad n = 1, 2, \ldots$$

Split $[A, B]$ into 2 subintervals $[A, \frac{A+B}{2}]$ and $[\frac{A+B}{2}, B]$. At least one of the 2 subintervals must contain infinitely many members of the sequence $\{a_n\}_{n=1}^{\infty}$ (why?). Denote this interval by $[A_1, B_1]$ and let $a_{n_1} \in [A_1, B_1]$ be one such member.

Set $i = 1$. Having constructed the interval $[A_i, B_i]$, consider the two subintervals $[A_i, \frac{1}{2}(A_i + B_i)]$ and $[\frac{1}{2}(A_i + B_i), B_i]$, each of length $\frac{1}{2}(B_i - A_i)$. Again at least one of the intervals, call it $[A_{i+1}, B_{i+1}]$, contains infinitely many elements of $\{a_n\}_{n=1}^{\infty}$. Let $a_{n_{i+1}} \in [A_{i+1}, B_{i+1}]$ be one such element.

Proceeding inductively, we define a sequence of nested intervals $[A_i, B_i]$ such that for each $i = 1, 2, \ldots$,

$$a_{n_{i+1}} \in [A_{i+1}, B_{i+1}] \subset [A_i, B_i],$$

i.e.

$$A_i \le A_{i+1} \le a_{n_{i+1}} \le B_{i+1} \le B_i, \quad \forall i = 1, 2, \ldots.$$

Note that $\{A_i\}_{i=1}^{\infty}$ is a bounded increasing sequence $\Rightarrow \lim\limits_{i\to\infty} A_i \exists = L$. Likewise, $\{B_i\}_{i=1}^{\infty}$ is a bounded decreasing sequence $\Rightarrow \lim\limits_{i\to\infty} B_i \exists$. Moreover,

$$\begin{aligned} \lim_{i\to\infty} B_i - \lim_{i\to\infty} A_i &= \lim_{i\to\infty} (B_i - A_i) = \lim_{i\to\infty} \frac{B-A}{2^i} = (B-A) \lim_{i\to\infty} \frac{1}{2^i} = 0 \\ &\Rightarrow \lim_{i\to\infty} B_i = \lim_{i\to\infty} A_i = L. \end{aligned}$$

Note that $A_i \le a_{n_i} \le B_i \; \forall i \in \mathbb{N}$.

Since $\{A_i\}_{i=1}^{\infty}$ and $\{B_i\}_{i=1}^{\infty}$ have the same limit L, the Squeeze Principle tells us $\lim\limits_{i\to\infty} a_{n_i} \exists = L$.

We have thus constructed a convergent subsequence $\{a_{n_i}\}_{i=1}^{\infty}$ of $\{a_n\}_{n=1}^{\infty}$.

2.E Cauchy Criterion

[Muldowney 1990, pp. 38]

Q. To prove convergence by the ϵ, N definition, we need to first know the limit L. For a monotonic sequence we saw we could prove convergence without knowing L: all we have to establish is that the sequence is bounded. But what can we do in the case of nonmonotonic sequences when we don't know which value of L we should use in the limit definition?

A. Use the Cauchy Criterion.

Definition: The sequence $\{a_n\}$ is a *Cauchy Sequence* if for all $\epsilon > 0$, there exists a number N such that
$$m, n > N \Rightarrow |a_m - a_n| < \epsilon.$$
N.B. This must hold not just for $m = n+1$ and n, but also for $m = n+2$, $m = n+3$, ..., deep into the "tail" of the sequence.

- Let $a_n = \dfrac{1}{n}$. Given $\epsilon > 0$, the Triangle Inequality implies
$$|a_m - a_n| = \left|\frac{1}{m} - \frac{1}{n}\right| \leq \left|\frac{1}{m}\right| + \left|\frac{1}{n}\right| < \frac{2}{N} = \epsilon$$
whenever $m, n > N$ provided we choose $N = \dfrac{2}{\epsilon}$.
Hence $\{\frac{1}{n}\}$ is a Cauchy Sequence. Notice that $\{\frac{1}{n}\}$ is convergent.

Remark: Without loss of generality we can always take $m \geq n > N$.

Remark: If for every N we can find even a single pair of values n, m, both larger than N, that violates $|a_n - a_m| < \epsilon$, then the sequence is not a Cauchy Sequence.

- $\{(-1)^n\}$ is not a Cauchy Sequence. Notice $\{(-1)^n\}$ is divergent.
Pick $m = n + 1$: $|(-1)^{n+1} - (-1)^n| = 2 \not< \epsilon$, had we been given $\epsilon = 1$ (say).
What we have observed in these two cases was formulated by Cauchy into the next theorem.

Theorem 2.6 (Cauchy Criterion): $\{a_n\}$ *is convergent* $\iff$ $\{a_n\}$ *is a Cauchy sequence.*

Proof:

"$\Rightarrow$" Suppose $\{a_n\}$ is convergent. Let $L = \lim\limits_{n\to\infty} a_n$.

Given $\epsilon > 0$, $\exists N \ni$
$$n > N \Rightarrow |a_n - L| < \frac{\epsilon}{2}.$$
Also,
$$m > N \Rightarrow |a_m - L| < \frac{\epsilon}{2}.$$
Therefore
$$m, n > N \Rightarrow |a_m - a_n| \leq |a_m - L| + |L - a_n| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Thus $\{a_n\}$ is a Cauchy sequence.

"$\Leftarrow$" Let $\{a_n\}$ be a Cauchy sequence.

Step 1: $\{a_n\}$ is bounded.

There exists a number N such that

$$m, n > N \Rightarrow |a_n - a_m| < 1 \quad \text{(in particular)}.$$

E.g. Let $m = N + 1$. Then $n > N \Rightarrow$

$$\begin{aligned} |a_n| - |a_{N+1}| \le |a_n - a_{N+1}| < 1 \\ \Rightarrow |a_n| \le 1 + |a_{N+1}|, \end{aligned}$$

so $|a_n| \le B = \max\{|a_1|, \dots, |a_N|, |a_{N+1}| + 1\}$ for all $n \in \mathbb{N}$.

Step 2: $\{a_n\}_{n=1}^{\infty}$ has a convergent subsequence $\{a_{n_k}\}_{k=1}^{\infty}$.

This follows from Step 1 and the Bolzano–Weierstrass Theorem.

Step 3: $\{a_n\}$ converges to the limit $L \doteq \lim\limits_{k\to\infty} a_{n_k}$.

We are given that $\{a_n\}$ is a Cauchy sequence.

$\therefore$ Given $\epsilon > 0$, $\exists N \ni$

$$m, n > N \Rightarrow |a_n - a_m| < \frac{\epsilon}{2}.$$

Also, from Step 2 we know that there exists a number K, which we may take to be at least as large as N, such that

$$k > K \Rightarrow |a_{n_k} - L| < \frac{\epsilon}{2}.$$

We then take $m = n_k$ and use the fact that subsequence indices satisfy $n_k \ge k$:

$$n > N, n_k \ge k > K \ge N \Rightarrow |a_n - a_{n_k}| < \frac{\epsilon}{2}.$$

Thus

$$n > N \Rightarrow |a_n - L| \le |a_n - a_{n_k}| + |a_{n_k} - L| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence $\{a_n\}$ also converges (to the same limit L).

- $\{n\}$ is not a Cauchy sequence. To see this, just pick $m = n + 1$ and $\epsilon = 1$. From Theorem 2.6, we then know that $\{n\}$ diverges.

- $\left\{1 + \frac{(-1)^n}{n}\right\}$ is a Cauchy sequence.

  Given $\epsilon > 0$, we can make

$$\left|\left(1 + \frac{(-1)^n}{n}\right) - \left(1 + \frac{(-1)^m}{m}\right)\right| = \left|\frac{(-1)^n}{n} - \frac{(-1)^m}{m}\right| \le \frac{1}{n} + \frac{1}{m} < \frac{1}{N} + \frac{1}{N} = \frac{2}{N} = \epsilon,$$

  provided we make the natural numbers n and m both greater than $N = 2/\epsilon$. Hence this is a Cauchy sequence and it therefore converges.

- What about the sequence defined by $a_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$?

  Pick $m = 2n$:

$$\begin{aligned} |a_{2n} - a_n| &= \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} \\ &\geq \underbrace{\frac{1}{2n} + \frac{1}{2n} + \ldots + \frac{1}{2n}}_{n \text{ terms}} = \frac{n}{2n} = \frac{1}{2}. \end{aligned}$$

  Then, if we should be given $\epsilon = \frac{1}{2}$, we won't be able to satisfy $|a_m - a_n| < \epsilon$. So $\{a_n\}$ is not a Cauchy sequence. Hence, by Theorem 2.6 we see that $\{a_n\}$ diverges, i.e.

$$\sum_{k=1}^{\infty} \frac{1}{k} \doteq \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k} \nexists.$$

  This is known as the *harmonic series*. It diverges, but only very slowly.

Q. Using the above estimate, how many terms would you want to take to be sure that the sum is greater than 10?

Problem 2.11 (*Infinite limits*): Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of real numbers. If, given any natural number M, we can find a number N such that $n > N \Rightarrow a_n > M$, we say that

$$\lim_{n\to\infty} a_n = \infty.$$

(a) Suppose $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ are sequences of real numbers with $\lim_{n\to\infty} a_n = \infty$ and $\lim_{n\to\infty} b_n = \infty$. Show that $\lim_{n\to\infty} (a_n + b_n) = \infty$ and $\lim_{n\to\infty} a_n b_n = \infty$.

(b) Under the conditions of part (a), find examples that demonstrate $\lim_{n\to\infty} (a_n - b_n)$ and $\lim_{n\to\infty} a_n/b_n$ may exist as a real number, may have an infinite limit, or may fail to exist at all.

(c) Show that

$$\lim_{n\to\infty} a_n = \infty \Rightarrow \lim_{n\to\infty} \frac{1}{a_n} = 0.$$

(d) Does the converse to (c) hold? That is does,

$$\lim_{n\to\infty} \frac{1}{a_n} = 0 \Rightarrow \lim_{n\to\infty} a_n = \infty?$$

Problem 2.12 (*Limit Superior and Limit Inferior*): Let $\{a_n\}_{n=1}^{\infty}$ be a bounded sequence. Consider the sequence $\{s_n\}_{n=1}^{\infty}$ defined by $s_n = \sup\{a_n, a_{n+1}, a_{n+2}, \ldots\}$ for $n \in \mathbb{N}$.

(a) Prove that $\{s_n\}_{n=1}^{\infty}$ is a bounded sequence.

(b) Prove that $\{s_n\}_{n=1}^{\infty}$ is a monotone sequence. Is $\{s_n\}_{n=1}^{\infty}$ an increasing or a decreasing sequence?

(c) Prove that $\{s_n\}_{n=1}^{\infty}$ is convergent.

Note: The limit of the sequence $\{s_n\}_{n=1}^{\infty}$ is known as the *limit superior* of the sequence $\{a_n\}_{n=1}^{\infty}$ and is written $\limsup\limits_{n\to\infty} a_n$. This is just the supremum of the values in the tail of the sequence. In a similar manner, one can define the *limit inferior*: $\liminf\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} i_n$, where $i_n = \inf\{a_n, a_{n+1}, a_{n+2}, \ldots\}$. Note that $\liminf\limits_{n\to\infty} a_n \leq \limsup\limits_{n\to\infty} a_n$. In fact, a sequence $\{a_n\}_{n=1}^{\infty}$ converges $\iff$ $\{a_n\}_{n=1}^{\infty}$ is bounded and $\liminf\limits_{n\to\infty} a_n = \limsup\limits_{n\to\infty} a_n$. For example, the bounded sequence $\{\sin n\}_{n=1}^{\infty}$ does not converge because $\liminf\limits_{n\to\infty} \sin n = -1$ and $\limsup\limits_{n\to\infty} \sin n = 1$.

Chapter 3

Functions

[Spivak 1994, Chapter 3]

3.A Examples of Functions

In the previous chapter, we defined a function f as a rule that associates a number y to each number x in its domain. An equivalent definition is [Spivak 1994, p. 47]:

Definition: A *function* is a collection of pairs of numbers (x, y) such that if (x, y_1) and (x, y_2) are in the collection, then $y_1 = y_2$. That is,

$$x_1 = x_2 \Rightarrow f(x_1) = f(x_2).$$

This can be restated as the *vertical line test*: an set of ordered pairs (x, y) is a function if every vertical line intersects their graph at most once.

Definition: If a function f has domain A and range B, we write $f : A \to B$.

Definition: *Constant functions* are functions of the form $f(x) = c$, where c is a constant.

Definition: *Polynomials* are functions of the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0.$$

When $a_n \neq 0$, we say that the *degree* of f is n and write $\deg f = n$. While a nonzero constant function has degree 0, it turns out to be convenient to define the degree of the zero function $f(x) = 0$ to be $-\infty$.

Note that a polynomial $f(x)$ with only even-degree terms (all the odd-degree coefficients are zero) satisfies the property $f(-x) = f(x)$, while a polynomial $f(x)$ with only odd-degree terms satisfies $f(-x) = -f(x)$. We generalize this notion with the following definition.

Definition: A function f is said to be *even* if $f(-x) = f(x)$ for every x in its domain.

Definition: A function f is said to be *odd* if $f(-x) = -f(x)$ for every x in its domain.

Problem 3.1: Show that every function $f : \mathbb{R} \to \mathbb{R}$ can be written as a sum of an even function and an odd function.

Definition: *Rational functions* are functions of the form $f(x) = \dfrac{P(x)}{Q(x)}$, where $P(x)$ and $Q(x)$ are polynomials. They are defined on the sets of all x for which $Q(x) \neq 0$.

- $\dfrac{x^3 + 3x^2 + 1}{x^2 + 1}$ and $\dfrac{1}{x}$ are both rational functions.

Composition Once we have defined a few elementary functions, we can create new functions by combining them together using the arithmetic operations $+$, $-$, $\cdot$, $\div$, or by introducing the *composition* operator $\circ$.

Definition: If $f : A \to B$ and $g : B \to C$ then we define $g \circ f : A \to C$ to be the function that takes $x \in A$ to $g(f(x)) \in C$.

- $$\begin{array}{ll} f(x) = x^2 + 1 & f : \mathbb{R} \to [1, \infty), \\ g(x) = 2\sqrt{x} & g : [1, \infty) \to [2, \infty), \\ g(f(x)) = 2\sqrt{x^2 + 1} & g \circ f : \mathbb{R} \to [2, \infty). \end{array}$$

  Note however that $f(g(x)) = 4x + 1$, so that $f \circ g : [0, \infty) \to [1, \infty)$.

- $$\begin{array}{ll} f(x) = x^2 + 1 & f : \mathbb{R} \to [1, \infty), \\ g(x) = \frac{1}{x} & g : [1, \infty) \to (0, 1], \\ g(f(x)) = \frac{1}{x^2+1} & g \circ f : \mathbb{R} \to (0, 1]. \end{array}$$

  One can also build new functions from old ones using *cases*, or *piecewise* definitions.

- $$f(x) = \begin{cases} 0 & x < 0, \\ \frac{1}{2} & x = 0, \\ 1 & x > 0. \end{cases}$$

- $$f(x) = |x| = \begin{cases} x & x \geq 0, \\ -x & x < 0. \end{cases}$$

  Notice that cases can sometimes introduce jumps in a function.

3.B Trigonometric Functions

Trigonometric functions are functions relating the shape of a right-angle triangle to one of its other angles.

Definition: If we label one of the non-right angles by θ, the length of the hypotenuse by hyp, and the lengths of the sides opposite and adjacent to x by opp and adj, respectively, then

$$\sin\theta = \frac{\text{opp}}{\text{hyp}},$$
$$\cos\theta = \frac{\text{adj}}{\text{hyp}},$$
$$\tan\theta = \frac{\text{opp}}{\text{adj}}.$$

Note here that since θ is one of the nonright angles of a right-angle triangle, these definitions apply only when $0 < \theta < 90^\circ$. Note also that $\tan\theta = \sin\theta/\cos\theta$. Sometimes it is convenient to work with the reciprocals of these functions:

$$\csc\theta = \frac{1}{\sin\theta},$$
$$\sec\theta = \frac{1}{\cos\theta},$$
$$\cot\theta = \frac{1}{\tan\theta}.$$

Pythagorean Identities: Recall from Pythagoras' Theorem that

$$(\text{opp})^2 + (\text{adj})^2 = (\text{hyp})^2.$$

Dividing by the square of the length of the hypotenuse, we see that

$$\sin^2\theta + \cos^2\theta = 1. \tag{3.1}$$

Other useful identities result from dividing both sides of this equation either by $\sin^2\theta$:

$$1 + \cot^2\theta = \csc^2\theta,$$

or by $\cos^2\theta$:

$$\tan^2\theta + 1 = \sec^2\theta.$$

Note that Eq. (3.1) implies both that $|\sin\theta| \leq 1$ and $|\cos\theta| \leq 1$.

Definition: We define the number π to be the area of a *unit circle* (a circle with radius 1).

Problem 3.2: Give a geometrical argument to show that $2 < \pi < 4$.

Definition: Instead of using degrees, in our development of calculus it will be more convenient to measure angles in terms of the area of the sector they subtend on the unit circle. Specifically, we define an angle measured in *radians* to be twice[1] the area of the sector that it subtends, as shown in Figure 3.1. For example, our definition of π says that a full unit circle (360°) has area π; the corresponding angle in radians would then be 2π. Thus, we can convert between radians and degrees with the formula

$$\pi \text{ radians} = 180^\circ.$$

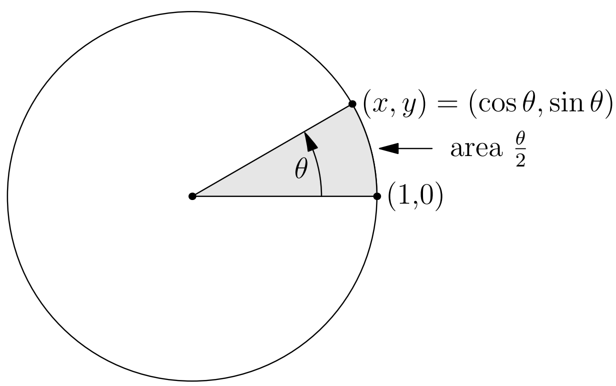


Figure 3.1: The unit circle

The coordinates x and y of a point P on the unit circle are related to θ as follows:

$$\cos\theta = \frac{\text{adj}}{\text{hyp}} = \frac{x}{1} = x,$$

$$\sin\theta = \frac{\text{opp}}{\text{hyp}} = \frac{y}{1} = y.$$

[1]The reason for introducing the factor of two in this definition is to make the angle x expressed in radians equal to the length of the arc it subtends on the unit circle, as we will see later using integral calculus, once we have developed the notion of the length of an arc. For example, the circumference of a full circle of unit radius will be found to be precisely 2π.

Complementary Angle Identities:

$$\cos\theta = \sin\left(\frac{\pi}{2} - \theta\right),$$

$$\cos\left(\frac{\pi}{2} - \theta\right) = \sin\theta.$$

Supplementary Angle Identities:

$$\sin(\pi - \theta) = \sin\theta,$$

$$\cos(\pi - \theta) = -\cos\theta.$$

Symmetries:

$$\sin(-\theta) = -\sin\theta,$$
$$\cos(-\theta) = \cos\theta,$$
$$\sin(\theta + 2\pi) = \sin\theta,$$
$$\cos(\theta + 2\pi) = \cos\theta.$$

Problem 3.3: We thus see that $\sin\theta$ is an odd periodic function of θ and $\cos\theta$ is an even periodic function of θ, both with period 2π. Use these facts to prove that $\tan\theta$ is an odd periodic function of θ with period π.

Special Values:

$$\begin{aligned}
\sin(0) &= \cos\left(\frac{\pi}{2}\right) = 0, \\
\sin\left(\frac{\pi}{2}\right) &= \cos(0) = 1, \\
\sin\left(\frac{\pi}{4}\right) &= \cos\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}, \\
\sin\left(\frac{\pi}{6}\right) &= \cos\left(\frac{\pi}{3}\right) = \frac{1}{2}, \\
\sin\left(\frac{\pi}{3}\right) &= \cos\left(\frac{\pi}{6}\right) = \frac{\sqrt{3}}{2}.
\end{aligned}$$

Addition Formulae:

Claim:

$$\cos(A - B) = \cos A\cos B + \sin A\sin B.$$

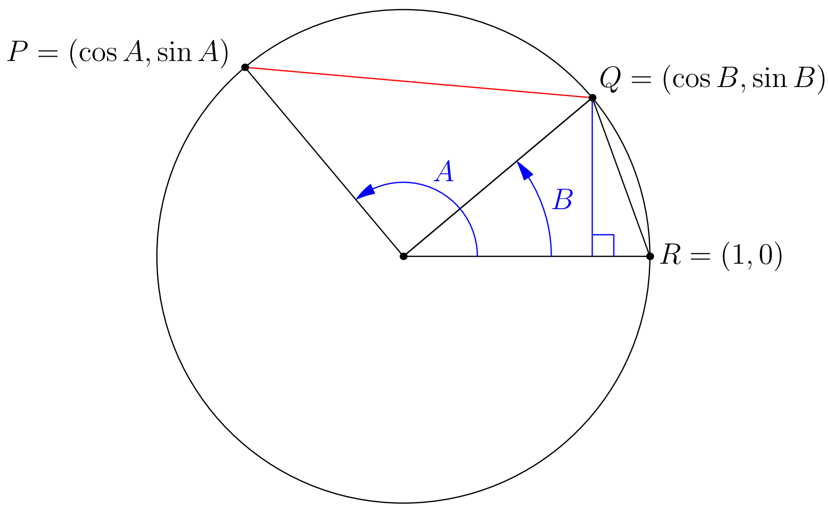


Figure 3.2: The unit circle with points $P = (\cos A, \sin A)$, $Q = (\cos B, \sin B)$, and $R = (1, 0)$

Proof: Consider the points $P = (\cos A, \sin A)$, $Q = (\cos B, \sin B)$, and $R = (1, 0)$ on the unit circle, as illustrated in Fig. 3.2. We can use Pythagoras' Theorem to obtain a formula for the length (squared) of a chord subtended by an angle:

$$\overline{QR}^2 = (1 - \cos B)^2 + \sin^2 B = 1 - 2\cos B + \cos^2 B + \sin^2 B = 2 - 2\cos B.$$

For example, since the angle subtended by $\overline{PQ}$ is $A - B$,

$$\overline{PQ}^2 = 2 - 2\cos(A - B).$$

Alternatively, we could compute $\overline{PQ}^2$ directly:

$$\begin{aligned}\overline{PQ}^2 &= (\cos A - \cos B)^2 + (\sin A - \sin B)^2 \\ &= \cos^2 A - 2\cos A\cos B + \cos^2 B + \sin^2 A - 2\sin A\sin B + \sin^2 B \\ &= 2 - 2(\cos A\cos B + \sin A\sin B).\end{aligned}$$

On comparing these two results, we conclude that

$$\cos(A - B) = \cos A\cos B + \sin A\sin B.$$

The claim thus holds.

Other addition formulae then follow easily:

$$\begin{aligned}\cos(A+B) &= \cos(A-(-B)) \\ &= \cos A\cos(-B) + \sin A\sin(-B) \\ &= \cos A\cos B - \sin A\sin B.\end{aligned}$$

$$\begin{aligned}\sin(A+B) &= \cos\left[\frac{\pi}{2} - (A+B)\right] \\ &= \cos\left[\left(\frac{\pi}{2} - A\right) - B\right] \\ &= \cos\left(\frac{\pi}{2} - A\right)\cos B + \sin\left(\frac{\pi}{2} - A\right)\sin B \\ &= \sin A\cos B + \cos A\sin B.\end{aligned}$$

$$\begin{aligned}\sin(A-B) &= \sin(A-(-B)) \\ &= \sin A\cos(-B) + \cos A\sin(-B) \\ &= \sin A\cos B - \cos A\sin B.\end{aligned}$$

$$\begin{aligned}\tan(A+B) = \frac{\sin(A+B)}{\cos(A+B)} &= \frac{\sin A\cos B + \cos A\sin B}{\cos A\cos B - \sin A\sin B} \\ &= \frac{(\sin A\cos B + \cos A\sin B)\cdot\frac{1}{\cos A\cos B}}{(\cos A\cos B - \sin A\sin B)\cdot\frac{1}{\cos A\cos B}} \\ &= \frac{\tan A + \tan B}{1 - \tan A\tan B}, \qquad \text{provided } A,\ B,\ A+B \text{ are not odd multiples of } \tfrac{\pi}{2}.\end{aligned}$$

Double-Angle Formulae:

$$\begin{aligned}\sin 2A &= \sin(A+A) \\ &= \sin A\cos A + \sin A\cos A \\ &= 2\sin A\cos A.\end{aligned}$$

$$\begin{aligned}\cos 2A &= \cos(A+A) \\ &= \cos A\cos A - \sin A\sin A \\ &= \cos^2 A - \sin^2 A \\ &= \cos^2 A - (1-\cos^2 A) \\ &= 2\cos^2 A - 1 \\ &= (1-\sin^2 A) - \sin^2 A \\ &= 1 - 2\sin^2 A.\end{aligned}$$

Also, if A is not an odd multiple of $\pi/4$ or $\pi/2$,

$$\begin{aligned}\tan 2A &= \tan(A+A)\\ &= \frac{\tan A + \tan A}{1-\tan A \tan A}\\ &= \frac{2\tan A}{1-\tan^2 A}.\end{aligned}$$

Inequalities: We have already seen that $|\sin x| \leq 1$ and $|\cos x| \leq 1$. Our development of trigonometric calculus will rely on the following additional key result:

$$\sin x \leq x \leq \tan x \quad \forall x \in \left[0, \frac{\pi}{2}\right).$$

We establish this result geometrically, referring to the arc of unit radius in Fig 3.3. The shaded area of the sector ABC subtended by the angle x (measured in radians) is $x/2$. Since $BE = \sin x$ and $DC = \tan x$, we deduce

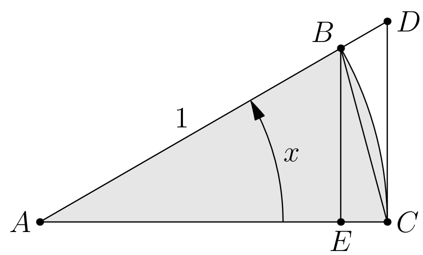


Figure 3.3: Geometric proof of $\sin x \leq x \leq \tan x$

$$\begin{aligned}&\text{Area}_{\triangle_{ABC}} \leq \text{Area}_{\text{Sector}_{ABC}} \leq \text{Area}_{\triangle_{ADC}}\\ \Rightarrow\ &\frac{1}{2}(1)\sin x \leq \frac{x}{2} \leq \frac{1}{2}(1)\tan x\\ &\Rightarrow \sin x \leq x \leq \tan x \quad \forall x \in \left[0, \frac{\pi}{2}\right).\end{aligned}$$

For $x \geq \pi/2$ we know that

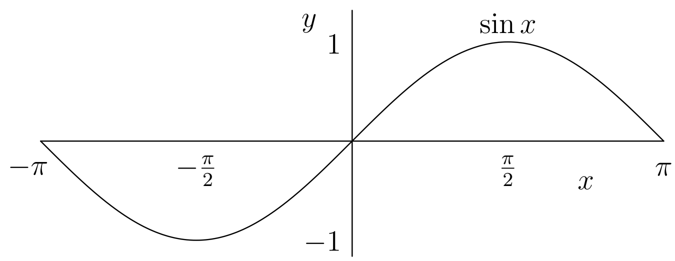
$$\sin x \leq 1 < \frac{\pi}{2} \leq x.$$

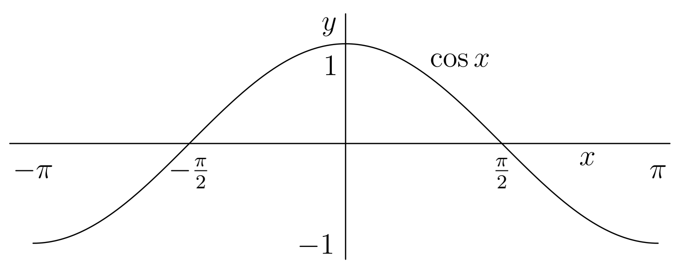
Hence $\sin x \leq x$ for all $x \geq 0$, from which we see that

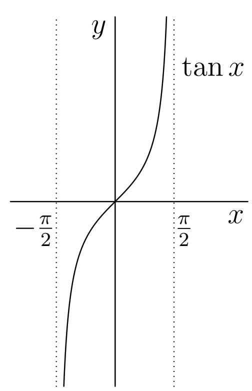
$$|\sin x| \leq |x| \quad \forall x \in \mathbb{R}.$$

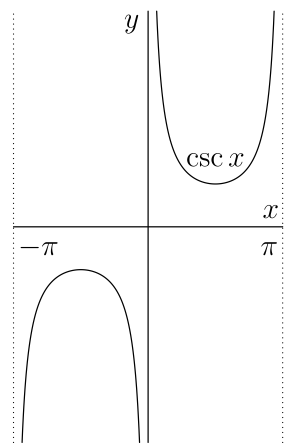
Problem 3.4: Verify that the graphs of the functions $y = \sin x$, $y = \cos x$, and $y = \tan x$ are periodic extensions of the illustrated graphs.

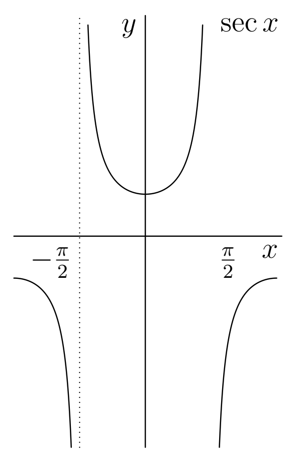
Problem 3.5: Verify that the graphs of the functions $y = \csc x = 1/\sin x$, $y = \sec x = 1/\cos x$, and $y = \cot x = 1/\tan x$ are periodic extensions of the illustrated graphs.

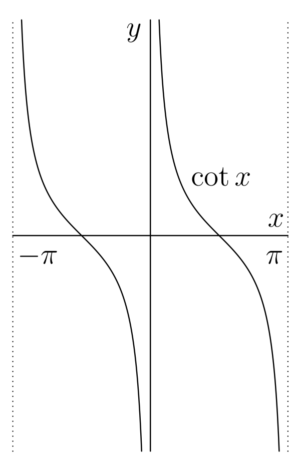







y
csc x
x
−π
π


y
sec x
−π/2
π/2
x


y
cot x
x
−π
π

3.C Limit of a Function

Consider the function $f(x) = \frac{1}{x}$ $(x \neq 0)$.

Notice that as x gets large, $f(x)$ gets closer to, but never quite reaches, 0, very much like the terms of the sequence $\{\frac{1}{n}\}$ as $n \to \infty$. In fact, at integer values of x, f evaluates to a member of the sequence $\{\frac{1}{n}\}$:

$$f(n) = \frac{1}{n} \longrightarrow 0 \quad \text{as } n \to \infty.$$

Unlike a sequence, f is defined also for nonintegral values of x. We therefore need to generalize our definition of a limit:

Definition: We say $\lim_{x\to\infty} f(x) = L$ if for every $\epsilon > 0$ we can find a real number $N \ni$

$$x > N \Rightarrow |f(x) - L| < \epsilon.$$

- Let $f(x) = 1/x$. Given any $\epsilon > 0$, we can make

$$|f(x) - 0| = \left|\frac{1}{x}\right| < \frac{1}{N} = \epsilon$$

for $x > N$ simply by picking $N = \frac{1}{\epsilon}$.

Hence $\lim_{x\to\infty} f(x) = 0$.

- Here is an interesting function:

$$g(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

This represents a "logical test" for rational numbers.

An even more interesting example is the function

$$f(x) = xg(x) = \begin{cases} x & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

Notice for all real numbers near $x = 0$ that $f(x)$ is very close to 0. That is, if δ is a small positive number, the value of $f(x)$ is very close to zero for all $x \in (-\delta, \delta)$. Given $\epsilon > 0$, we can always find a small region $(-\delta, \delta)$ about the origin such that

$$x \in (-\delta, \delta) \Rightarrow |f(x)| < \epsilon,$$
$$\text{i.e.} \quad |x| < \delta \Rightarrow |f(x)| < \epsilon.$$

For example, we could choose $\delta = \epsilon$ (or smaller). We express this fact with the statement $\lim_{x\to 0} f(x) = 0$.

Definition: We say $\lim_{x \to a} f(x) = L$ if for every $\epsilon > 0$ we can find a $\delta > 0$ such that

$$0 < |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon.$$

Remark: In the previous example we see that $a = 0$ and the limit L is 0. Notice in this case that $\lim_{x \to 0} f(x) = 0 = f(0)$. However, this is not true for all functions f. The value of a limit as $x \to a$ might be quite different from the value of the function at $x = a$. Sometimes the point a might not even be in the domain of the function, but the limit may still exist. This is why we restrict $0 < |x - a|$ (that is, $x \neq a$) in the above definition.

Remark: The value of f at a itself is irrelevant to the limit. We don't need to evaluate f at $x = a$ any more than we need to evaluate $1/n$ at $n = \infty$ to find its limit.

- Let

$$f(x) = \begin{cases} x & \text{if } x > 0, \\ -x & \text{if } x < 0. \end{cases}$$

When we say $\lim_{x \to 0} f(x) = 0$ we mean the following. Given $\epsilon > 0$, we can make

$$|f(x)| < \epsilon$$

for all x satisfying $0 < |x| < \delta$ just by choosing $\delta = \epsilon$. That is,

$$0 < |x| < \delta \Rightarrow |f(x)| = |x| < \delta = \epsilon.$$

- How about

$$f(x) = |x| = \begin{cases} x & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -x & \text{if } x < 0; \end{cases}$$

is $\lim_{x \to 0} f(x) = 0$? Yes, the value of f at $x = 0$ does not matter.

- Consider now

$$f(x) = \begin{cases} x & \text{if } x > 0, \\ 1 & \text{if } x = 0, \\ -x & \text{if } x < 0. \end{cases}$$

Is $\lim_{x \to 0} f(x) = 0$? Yes, the value of f at $x = 0$ does not matter.

- Let

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ \frac{1}{2} & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

This function is defined everywhere. Does

$$\lim_{x \to 0} f(x) \exists ?$$

No, given $\epsilon = \frac{1}{2}$, there are values of $x \neq 0$ in every interval $(-\delta, \delta)$ with very different values of f:

$$f\left(\frac{\delta}{2}\right) = 1,$$
$$f\left(-\frac{\delta}{2}\right) = 0.$$

Thus

$$\left| f\left(\frac{\delta}{2}\right) - L \right| < \frac{1}{2} \Rightarrow |1 - L| < \frac{1}{2},$$

$$\left| f\left(-\frac{\delta}{2}\right) - L \right| < \frac{1}{2} \Rightarrow |L| < \frac{1}{2}.$$

The first statement implies

$$-\frac{1}{2} < L - 1 < \frac{1}{2} \Rightarrow \frac{1}{2} < L < \frac{3}{2}$$

and this contradicts $|L| < \frac{1}{2}$. So we see that no such number L exists; that is, $\lim_{x \to 0} f(x) \nexists$.

- Let $f(x) = 7x - 3$. Show that $\lim\limits_{x \to 1} f(x) = 4$.

  Let $\epsilon > 0$. Our task is to produce a $\delta > 0$ such that

$$0 < |x - 1| < \delta \Rightarrow |f(x) - 4| < \epsilon.$$

Well, $|f(x) - 4| = |7x - 7| = 7\,|x - 1| < 7\delta$.
  How can we make $|f(x) - 4| < \epsilon$?
  No matter what ϵ we are given, we can easily choose $\delta = \epsilon/7$, so that $7\delta = \epsilon$.

Q. Suppose

$$f(x) = \begin{cases} 7x - 3 & x \neq 1, \\ 5 & x = 1. \end{cases}$$

What is $\lim\limits_{x \to 1} f(x)$?

A. The limit is still 4; the value of $f(x)$ at $x = 1$ is completely irrelevant. The function need not even be defined at $x = 1$.

Remark: $\lim\limits_{x\to a}$ describes the behaviour of a function near a, not at a.

- Let $f(x) = x^2$, $x \in \mathbb{R}$.
  Show $\lim\limits_{x\to 3} f(x) = 9$.

$$\begin{aligned}|x-3| < \delta \Rightarrow |f(x)-9| = \left|x^2-9\right| = |x-3|\,|x+3| &= |x-3|\,|x-3+6| \\ &< \delta(\delta+6) \quad \text{from the Triangle Inequality.}\end{aligned}$$

  We could solve the quadratic equation $\delta(\delta+6) = \epsilon$, but it is easier to restrict $\delta \le 1$ so that
$$\delta(\delta+6) \le \delta(1+6) = 7\delta \le \epsilon \quad \text{if} \quad \delta \le \frac{\epsilon}{7}.$$
  Note here that we must allow for the possibility that $\delta < \epsilon/7$ instead of just setting $\delta = \epsilon/7$, in order to satisfy our simplifying restriction that $\delta \le 1$.
  Hence
$$|x-3| < \min\left(1, \frac{\epsilon}{7}\right) \Rightarrow |f(x)-9| < \epsilon.$$

- Let $f(x) = \frac{1}{x}$, $x \ne 0$.
  Show $\lim\limits_{x\to 2} f(x) = \dfrac{1}{2}$.
  Given $\epsilon > 0$, try to find a $\delta \ni$
$$0 < |x-2| < \delta \Rightarrow \left|f(x) - \frac{1}{2}\right| < \epsilon.$$
  Note $\left|f(x) - \dfrac{1}{2}\right| = \left|\dfrac{1}{x} - \dfrac{1}{2}\right| = \left|\dfrac{2-x}{2x}\right|$ becomes very large near $x = 0$.
  Is this a problem? No, we are only interested in the behaviour of the function near $x = 2$.
  Let us restrict $\delta \le 1$, to keep the factor $2x$ in the denominator from getting really small (and hence the whole expression from getting really large). Then
$$|x-2| < 1 \Rightarrow -1 < x-2 < 1 \Rightarrow 1 \le x \le 3 \Rightarrow \frac{1}{x} \le 1.$$
  So
$$\left|f(x) - \frac{1}{2}\right| = \left|\frac{2-x}{2x}\right| \le \frac{1}{2}|x-2| < \frac{1}{2}\delta \le \epsilon,$$
  if we take $\delta = \min(1, 2\epsilon)$.

Problem 3.6: Suppose $\lim\limits_{x\to a} f(x) > 0$. Show that there exists a number $\delta > 0$ such that $f(x) > 0$ for all x satisfying $0 < |x-a| < \delta$. Can one necessarily conclude that $f(a) > 0$?

3.D Properties of Limits

Q. The definition of $\lim_{x \to a} f(x)$ is reminiscent of $\lim_{n \to \infty} a_n$, for which we proved several properties, including a Cauchy Criterion. Do there exist similar theorems for $\lim_{x \to a} f(x)$?

A. Yes.

Q. Do we have to prove these theorems all over again?

A. No, the following theorem provides the "bridge" we need to connect sequence limits to function limits.

Theorem 3.1 (Equivalence of Function and Sequence Limits): $\lim_{x \to a} f(x) = L \iff f$ *is defined near a and* **every** *sequence of points $\{x_n\}$ in the domain of f, with $x_n \neq a$ but $\lim_{n \to \infty} x_n = a$, satisfies $\lim_{n \to \infty} f(x_n) = L$.*

Proof:

"$\Rightarrow$" Suppose $\lim_{x \to a} f(x) = L$. Let $\{x_n\}$ be any sequence of points from the domain of f with $x_n \neq a$, but $\lim_{n \to \infty} x_n = a$.

That is, given $\epsilon > 0$, $\exists \delta > 0 \ni$

$$0 < |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon$$

and $\exists N \ni$

$$n > N \Rightarrow 0 < |x_n - a| < \delta.$$

Then $n > N \Rightarrow |f(x_n) - L| < \epsilon$,

$$\text{i.e. } \lim_{n \to \infty} f(x_n) = L.$$

"$\Leftarrow$" Suppose that $\lim_{n \to \infty} f(x_n) = L$ for each sequence $\{x_n\}$ in the domain of f with $x_n \neq a$, but $\lim_{n \to \infty} x_n = a$.

If $f(x)$ does not have limit L at a, then it is not true that:

$$\text{Given } \epsilon > 0, \exists \delta > 0 \text{ such that } 0 < |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon.$$

Hence there exists some ϵ, say ϵ_0, for which this fails, no matter which δ we try. This means that no matter how large we make $n \in \mathbb{N}$, there exists some x with $0 < |x - a| < \frac{1}{n}$ that violates our error requirement:

$$|f(x) - L| \not< \epsilon_0.$$

Denote this particular value of x by $x_n \neq a$. In this way, we construct for each integer $n = 1, 2, \ldots$ a member of a sequence $\{x_n\}$. As $n \to \infty$, this sequence converges to a, yet the function values $\{f(x_n)\}$ do not converge to L since $|f(x) - L| \geq \epsilon_0$. This contradicts the premise $\lim_{n\to\infty} f(x_n) = L$ for any such sequence $\{x_n\}$. Hence it must be that $\lim_{x\to a} f(x) = L$.

Corollary 3.1.1 (Properties of Function Limits): Suppose $L = \lim_{x\to a} f(x)$ and $M = \lim_{x\to a} g(x)$. Then

(a) $\lim_{x\to a}(f(x) + g(x)) = L + M$;

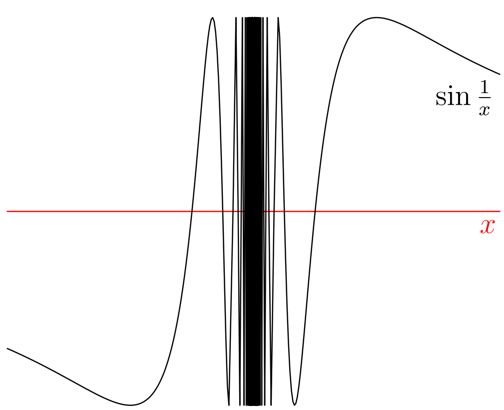
(b) $\lim_{x\to a} f(x)g(x) = LM$;

(c) $\lim_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{L}{M}$ if $M \neq 0$.

Corollary 3.1.2 (Cauchy Criterion for Functions): $\lim_{x\to a} f(x)$ exists $\iff$ for **every** $\epsilon > 0$, $\exists \delta > 0$ such that, whenever two numbers x and y are chosen from the set $(a - \delta, a) \cup (a, a + \delta)$, their functions values satisfy $|f(x) - f(y)| < \epsilon$.

- How about $\lim_{x\to 0} f(x)$, where

$$f(x) = \begin{cases} \sin\frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$



Note that there exist values $x \neq 0$ and other values $y \neq 0$ in any interval $(-\delta, \delta)$ with $f(x) = 1$ and $f(y) = -1$, so that $|f(x) - f(y)| = 2$. We cannot possibly satisfy the Cauchy Criterion for $\epsilon = 1$, so by Corollary 3.1.2, we see that $\lim_{x \to 0} f(x) \nexists$.

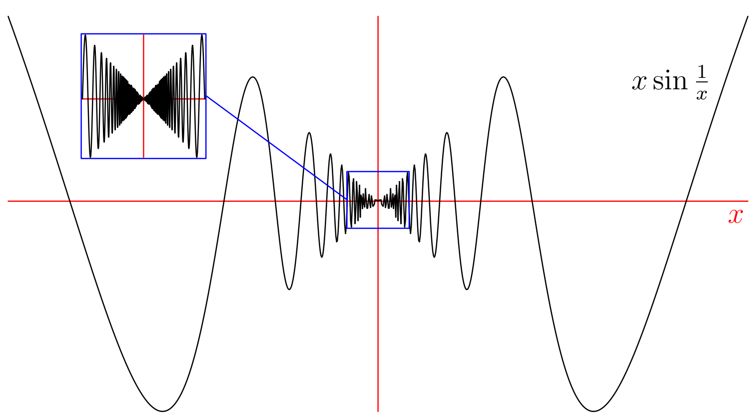
For example, when $\dfrac{1}{x} = \dfrac{\pi}{2} + 2n\pi$, where $n \in \mathbb{N}$, then $\sin\left(\dfrac{1}{x}\right) = 1$. This happens when $x = \left(\dfrac{\pi}{2} + 2n\pi\right)^{-1}$. There are infinitely many such values in $(0, \delta)$, one for every $n > \dfrac{1}{2\pi\delta} - \dfrac{1}{4}$.

Similarly, when $\dfrac{1}{y} = \dfrac{3\pi}{2} + 2n\pi$, where $n \in \mathbb{N}$, then $\sin\left(\dfrac{1}{y}\right) = -1$. This happens when $y = \left(\dfrac{3\pi}{2} + 2n\pi\right)^{-1}$. There are infinitely many such values in $(0, \delta)$, one for every $n > \dfrac{1}{2\pi\delta} - \dfrac{3}{4}$.

- However, for the function

$$h(x) = \begin{cases} x \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases}$$

$\lim_{x \to 0} h(x)$ exists and equals 0: given $\epsilon > 0$, we can make $|h(x) - 0| \leq |x| < \epsilon$ for all x satisfying $0 < |x| < \delta$ if we simply choose $\delta = \epsilon$.



Corollary 3.1.3 (Squeeze Principle for Functions): Suppose $f(x) \leq h(x) \leq g(x)$ when $x \in (a - \delta, a + \delta)$, for some $\delta > 0$. Then

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = L \Rightarrow \lim_{x \to a} h(x) = L.$$

- Consider

$$f(x) = -|x|,$$
$$h(x) = \begin{cases} x\sin\frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$
$$g(x) = |x|.$$

Since

$$f(x) \le h(x) \le g(x),$$

and

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} g(x) = 0,$$

we learn from Corollary 3.1.3 that

$$\lim_{x\to 0} h(x) = 0.$$

3.E Continuity

[Muldowney 1990, pp. 55–58]
[Spivak 1994, Chapter 6]

Definition: Let $D \subset \mathbb{R}$. A point c is an *interior point* of D if it belongs to some open interval (a, b) entirely contained in D: $c \in (a, b) \subset D$.

- $\frac{1}{10}, \frac{1}{2}, \frac{2}{3}, \frac{9}{10}$ are interior points of $[0, 1]$ but 0 and 1 are not.

- All points of $(0, 1)$ are interior points of $(0, 1)$.

Recall that the value of f at $x = a$ is completely irrelevant to the value of its limit as $x \to a$. Sometimes, however, these two values will happen to agree. In that case, we say that $f(x)$ is *continuous* at $x = a$.

Definition: A function f is *continuous* at an interior point a of its domain if

$$\lim_{x\to a} f(x) = f(a).$$

Remark: Otherwise, if

(a) the limit fails to exist, or

(b) the limit exists and equals some number $L \neq f(a)$,

the function is said to be *discontinuous*.

Remark: f is continuous at $a \iff$ for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Note that when $x = a$ we have $|f(a) - f(a)| = 0 < \epsilon$.

- $f(x) = x$ is continuous at every point a of its domain ($\mathbb{R}$) since $\lim\limits_{x\to a} x = a = f(a)$ for all $a \in \mathbb{R}$.

- $f(x) = x^2$ is continuous at all points a by Corollary 3.1.1, since

$$\lim_{x\to a} f(x) = \lim_{x\to a} x^2 = \lim_{x\to a} x \cdot \lim_{x\to a} x = a \cdot a = a^2 = f(a).$$

- By repeated use of Corollary 3.1.1, we see that a polynomial is continuous at all real numbers a.

Corollary 3.1.4 (Properties of Continuous Functions): Suppose f and g are continuous at a. Then $f + g$ and fg are continuous at a and f/g is continuous at a if $g(a) \neq 0$.

Corollary 3.1.5 (Continuity of Rational Functions): A rational function is continuous at all points of its domain.

- $f(x) = \dfrac{1}{x}$ is continuous at all $x \neq 0$.

- $f(x) = \dfrac{1}{x^2 + 1}$ is continuous everywhere.

- $f(x) = \dfrac{1}{x^2 - 1}$ is continuous on $(-\infty, -1) \cup (-1, 1) \cup (1, \infty)$.

Problem 3.7: Are any of the following functions continuous at 0?

$$f(x) = \begin{cases} \sin\frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

$$g(x) = \begin{cases} x\sin\frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

$$h(x) = \begin{cases} x\sin\frac{1}{x} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

Corollary 3.1.6 (Continuous Functions of Sequences): f is continuous at an interior point a of the domain of $f \iff$ each sequence $\{x_n\}$ in the domain of f with $\lim\limits_{n\to\infty} x_n = a$ satisfies $\lim\limits_{n\to\infty} f(x_n) = f(a)$.

- On p. 34 we showed for any sequence of non-negative numbers a_n that $a_n \to a \Rightarrow \sqrt{a_n} \to \sqrt{a}$. Hence $f(x) = \sqrt{x}$ is continuous at all $a > 0$. Note that $a = 0$ is not an interior point of the domain of f.

Corollary 3.1.7 (Composition of Continuous Functions): Suppose g is continuous at a and f is continuous at $g(a)$. Then $f \circ g$ is continuous at a.

Proof: Let $\{x_n\}$ be any sequence in the domain of g such that $\lim_{n\to\infty} x_n = a$. Then by Corollary 3.1.6, $\{g(x_n)\}$ is a sequence in the domain of f such that $\lim_{n\to\infty} g(x_n) = g(a)$ since g is continuous at a.

Likewise, since f is continuous at $g(a)$, $\lim_{n\to\infty} f(g(x_n)) = f(g(a))$, (again appealing to Corollary 3.1.6).

- Given any continuous function $g(x)$, is $|g(x)|$ continuous?

We have already seen that $\lim_{x\to 0} |x| = 0 = |0|$. This means that the function $f(x) = |x|$ is continuous at $x = 0$. In fact, $f(x) = |x|$ is continuous at all $x \in \mathbb{R}$ since at positive x points it behaves locally as the function x and at negative x it behaves locally as the function $-x$, both of which are continuous functions. We then know from Corollary 3.1.7 that $f \circ g = |g(x)|$ is continuous at all $x \in \mathbb{R}$.

- Corollary 3.1.7 can also be used to show that the function

$$\begin{cases} x \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

is continuous at all $a \in \mathbb{R}$, not just at $a = 0$. All we need to do is to prove that $\sin x$ is a continuous function at all $a \in \mathbb{R}$. On p. 61 we showed that

$$|\sin x| \leq |x| \quad \forall x \in \mathbb{R}. \tag{3.2}$$

Recall

$$\begin{aligned} \sin(A + B) &= \sin A \cos B + \cos A \sin B, \\ \sin(A - B) &= \sin A \cos B - \cos A \sin B. \end{aligned}$$

Then

$$\sin(A + B) - \sin(A - B) = 2 \cos A \sin B.$$

Let

$$\left.\begin{aligned} A + B &= x \\ A - B &= y \end{aligned}\right\} \text{ i.e. } \begin{aligned} A &= \tfrac{x+y}{2}, \\ B &= \tfrac{x-y}{2}. \end{aligned}$$

Thus

$$\sin x - \sin y = 2 \cos\left(\frac{x + y}{2}\right) \sin\left(\frac{x - y}{2}\right) \quad \forall x, y \in \mathbb{R}.$$

We use this result together with Eq. (3.2) to show that $\sin x$ is continuous at any $a \in \mathbb{R}$. Given $\epsilon > 0$, choose $\delta = \epsilon$. Then

$$\begin{aligned}|x-a| < \delta \Rightarrow |\sin x - \sin a| &= 2\left|\cos\left(\frac{x+a}{2}\right)\right|\left|\sin\left(\frac{x-a}{2}\right)\right| \\ &\le 2\left|\sin\left(\frac{x-a}{2}\right)\right| \\ &\le 2\left|\frac{x-a}{2}\right| = |x-a| < \delta = \epsilon.\end{aligned}$$

From Corollary 3.1.7, it then follows that $\cos x = \sin\left(\frac{\pi}{2} - x\right)$ is also continuous on $\mathbb{R}$ and that $\tan x$ is continuous on its domain.

3.F One-Sided Limits

Definition: We write $\lim\limits_{x\to a^+} f(x) = L$ if for each $\epsilon > 0$, $\exists \delta > 0 \ni$

$$\underbrace{0 < x - a < \delta}_{\text{i.e. } a<x<a+\delta} \Rightarrow |f(x) - L| < \epsilon.$$

- For the function
$$H(x) = \begin{cases} 1 & \text{if } x \ge 0, \\ 0 & \text{if } x < 0, \end{cases}$$
we see that $\lim\limits_{x\to 0^+} H(x) = 1$.

Definition: We write $\lim\limits_{x\to a^-} f(x) = L$ if for each $\epsilon > 0$, $\exists \delta > 0 \ni$

$$0 < a - x < \delta \Rightarrow |f(x) - L| < \epsilon.$$

- In the above example, we see that $\lim\limits_{x\to 0^-} H(x) = 0$.

Remark: $\lim\limits_{x\to a} f(x) = L \iff \lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^-} f(x) = L$.

Definition: A function f is *continuous from the right* at a if

$$\lim_{x\to a^+} f(x) = f(a).$$

Definition: A function f is *continuous from the left* at a if

$$\lim_{x \to a^-} f(x) = f(a).$$

- $f(x) = \sqrt{x}$ is continuous from the right at $x = 0$.

- In the above example, $H(x)$ is continuous from the right at all points, including 0, and from the left at all points except 0.

Remark: A function is continuous at an interior point a of its domain if and only if it is continuous both from the left and from the right at a.

Definition: A function f is said to be *continuous on* $[a, b]$ if f is continuous at each point in (a, b) and continuous from the right at a and from the left at b.

- $f(x) = \sqrt{x}$ is continuous on $[\,0, \infty)$.

Remark: Extending Corollary 3.1.6 to one-sided limits, we see that f is continuous on $[a, b] \iff \lim_{n\to\infty} f(x_n) = f(c)$ for each sequence $\{x_n\}$ such that $x_n \in [a, b]$ and $\lim_{n\to\infty} x_n = c$. That is, $\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right)$.

3.G Properties of Continuous Functions

[Muldowney 1990, pp. 59–68]
[Spivak 1994, Chapter 6]

We have seen that continuous functions are free of sudden jumps. This property may be exploited to help locate the roots of a continuous function. Suppose we want to know whether the continuous function $f(x) = x^3 + x^2 - 1$ has a root in $(0, 1)$. We might notice that $f(0)$ is negative and $f(1)$ is positive. Since f has no jumps, it would then seem plausible that there exists a number $c \in (0, 1)$ where $f(c) = 0$. The following theorem establishes that this is indeed the case.

Theorem 3.2 (Intermediate Value Theorem [IVT]): *Suppose*

(i) *f is continuous on [a,b],*

(ii) *$f(a) < 0 < f(b)$.*

Then there exists a number $c \in (a, b)$ such that $f(c) = 0$.

Proof: Let $a_1 = a$ and $b_1 = b$. Inductively we define, for $n = 1, 2, \ldots$,

$$a_{n+1} = \begin{cases} \dfrac{a_n + b_n}{2} & \text{if } f\left(\dfrac{a_n + b_n}{2}\right) \le 0, \\ a_n & \text{if } f\left(\dfrac{a_n + b_n}{2}\right) > 0 \end{cases}$$

and

$$b_{n+1} = \begin{cases} b_n & \text{if } f\left(\dfrac{a_n + b_n}{2}\right) \le 0, \\ \dfrac{a_n + b_n}{2} & \text{if } f\left(\dfrac{a_n + b_n}{2}\right) > 0. \end{cases}$$

We first use induction to show that $f(a_n) \le 0$ and $f(b_n) > 0$ for all n.

Step 1: We are given that this holds for $n = 1$.

Step 2: Suppose $f(a_n) \le 0$ and $f(b_n) > 0$. Then in both cases $f(a_{n+1}) \le 0$ and $f(b_{n+1}) > 0$.

Note that $a \le a_n \le a_{n+1} \le b_{n+1} \le b_n \le b \ \forall n$. Being bounded and monotone, $\{a_n\}$ and $\{b_n\}$ are hence convergent sequences. Also,

$$b_n - a_n = \frac{b - a}{2^{n-1}} \Rightarrow \lim_{n\to\infty} b_n = \lim_{n\to\infty} a_n = c \in [a, b].$$

$$\text{Now } f \text{ continuous} \Rightarrow \begin{cases} \lim\limits_{n\to\infty} f(a_n) = f\left(\lim\limits_{n\to\infty} a_n\right) = f(c), \\ \lim\limits_{n\to\infty} f(b_n) = f\left(\lim\limits_{n\to\infty} b_n\right) = f(c). \end{cases}$$

But

$$\begin{aligned} f(a_n) \le 0 &\Rightarrow f(c) = \lim_{n\to\infty} f(a_n) \le 0, \\ f(b_n) > 0 &\Rightarrow f(c) = \lim_{n\to\infty} f(b_n) \ge 0. \end{aligned}$$

$\therefore\ f(c) = 0$. Note that $\left\{ \begin{array}{l} f(a) < 0 \Rightarrow c \ne a \\ f(b) > 0 \Rightarrow c \ne b \end{array} \right\}$. Hence $c \in (a, b)$.

- Consider the continuous function $f(x) = x^n - C$, where $C > 0$ and $n \in \mathbb{N}$. Noting that $f(0) = -C < 0$ and $f(x) > 0$ for sufficiently large x, the equation $f(x) = 0$ is seen to have at least one positive root. Hence, we have established our claim on p. 45 that every number $C > 0$ always has a unique n^{th} positive root, so that one can define a function $f(x) = x^{1/n}$ on $[0, \infty)$. This example leads us to formulate the following corollary to Theorem 3.2.

Corollary 3.2.1 (Generalized Intermediate Value Theorem): Suppose

(i) f is continuous on [a,b],

(ii) $f(a) < y < f(b)$.

Then there exists a number $c \in (a, b)$ such that $f(c) = y$.

Proof: Apply Theorem 3.2 to the function $F(x) = f(x) - y$.

Problem 3.8: For $n \in \mathbb{N}$, prove that the function $f(x) = x^{1/n}$ is continuous on $[0, \infty)$. Hint: apply the factorization on p. 45 to $x - a = (x^{1/n})^n - (a^{1/n})^n$ for $a > 0$. Do the case $a = 0$ separately.

Definition: We say that a function is *bounded* on a set $\mathcal{S}$ if

$$\exists M \ni |f(x)| \leq M \quad \forall x \in \mathcal{S}.$$

- $f(x) = \dfrac{1}{x}$ is bounded on $[1, 2]$, but not on $(0, 1]$.

Theorem 3.3 (Boundedness of Continuous Functions on Closed Intervals): *If f is continuous on $[a, b]$ then f is bounded on $[a, b]$.*

Proof: We want to show that $|f(x)| \leq M \quad \forall x \in [a, b]$. We establish this by contradiction.

Suppose that no such number M exists. Then, for each $n = 1, 2, \ldots$, there exists $x_n \in [a, b] \ni$

$$|f(x_n)| > n. \tag{3.3}$$

The sequence $\{x_n\}$ is bounded (why?) and therefore has a subsequence $\{x_{n_k}\}$ that converges to some number $c \in [a, b]$ as $k \to \infty$. Now

$$f \text{ continuous on } [a, b] \Rightarrow \lim_{k\to\infty} f(x_{n_k}) = f(c).$$

This contradicts Eq. (3.3) (which implies that $\lim_{k\to\infty} |f(x_{n_k})| = \infty$).
Hence f is bounded on $[a, b]$.

Theorem 3.4 (Weierstrass Max/Min Theorem): *If f is continuous on $[a, b]$ then it achieves both a maximum and minimum value on $[a, b]$. That is, there exists numbers c and d in $[a, b]$ such that*

$$f(c) \leq f(x) \leq f(d) \quad \forall x \in [a, b].$$

Proof: Let $M = \sup\{f(x) : x \in [a, b]\}$ (why does this exist?).

For $n = 1, 2, \ldots$ we know that $M - 1/n$ is not an upper bound of $\{f(x) : x \in [a, b]\}$. That is, there exists a number $x_n \in [a, b]$ such that

$$M - \frac{1}{n} < f(x_n) \leq M.$$

By the Squeeze Principle, $\lim_{n\to\infty} f(x_n) = M$. Now, the bounded sequence $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\} \to d \in [a, b]$ and

$$f \text{ continuous on } [a, b] \ \Rightarrow \lim_{k\to\infty} f(x_{n_k}) = f(d).$$

But $\{f(x_{n_k})\}$ is a subsequence of the convergent sequence $\{f(x_n)\}$, so

$$\lim_{k\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(x_n) = M.$$

Hence $f(d) = M \geq f(x)\ \forall x \in [a, b]$. Similarly, $\exists c \ni f(c) = m \leq f(x)\ \forall x \in [a, b]$, where $m = \inf\{f(x) : x \in [a, b]\}$.

Problem 3.9: Complete the above proof to show that such a number c indeed exists.

Remark: Theorem 3.4 does not hold when the closed interval $[a, b]$ is replaced by the open interval (a, b). For example, $f(x) = 1/x$ on $(0, 1)$ has no maximum value. Does it have a minimum value?

Problem 3.10: Precisely what step goes wrong in the proof of Theorem 3.4 if we try to replace $[a, b]$ with (a, b) for functions f that are bounded on (a, b)? For example, consider the function $f(x) = 1/x$ on $(1, 2)$.

Corollary 3.4.1 (Image of a Continuous Function on a Closed Interval): If f is continuous on $[a, b]$ then $f([a, b])$ is either a closed interval or a point.

Proof: Theorem 3.4 $\Rightarrow$ f achieves its minimum value $f(c)$ and maximum value $f(d)$ at some points c and d in $[a, b]$, respectively. Corollary 3.2.1 $\Rightarrow$ f also achieves every value in between $f(c)$ and $f(d)$. Hence $f([a, b]) = [f(c), f(d)]$.

Problem 3.11: Determine which of the following limits exist as a finite number, which are ∞, which are $-\infty$, and which do not exist at all. Where possible, compute the limit.

(a)

$$\lim_{x\to 2} \frac{x^2 - 5x + 6}{x - 2}$$

$$= \lim_{x\to 2} \frac{(x-2)(x-3)}{x-2} = \lim_{x\to 2}(x - 3) = -1.$$

(b)

$$\lim_{x\to 2}\frac{x-2}{x^2-4}$$

$$=\lim_{x\to 2}\frac{x-2}{(x-2)(x+2)}=\lim_{x\to 2}\frac{1}{x+2}=\frac{1}{4}.$$

(c)

$$\lim_{x\to 2}\frac{x+2}{x^2-4}$$

The limit

$$\lim_{x\to 2}\frac{x+2}{(x-2)(x+2)}=\lim_{x\to 2}\frac{1}{x-2}.$$

does not exist.

(d)

$$\lim_{x\to 2}\frac{x+2}{(x-2)(x^2-4)}$$

The limit

$$=\lim_{x\to 2}\frac{x+2}{(x-2)^2(x+2)}=\lim_{x\to 2}\frac{1}{(x-2)^2}=\infty.$$

(e)

$$\lim_{x\to 8}\frac{x+\sqrt{x+1}}{x-\sqrt{x+1}}$$

Since the function under consideration is continuous, it may be evaluated directly by substitution to obtain the answer 11/5.

Problem 3.12: Let $f:\mathbb{R}\to\mathbb{R}$ be a continuous odd function. Prove that there exists a number $c\in[0,1)$ such that $f(c)=\frac{1}{2}f(1)$. Hint: Draw a picture.

Since $f(x)=-f(-x)$ we know that $f(0)=0$. If $f(1)=0$ then $c=0$. Otherwise either $f(1)>\frac{1}{2}f(1)>0$ or $f(1)<\frac{1}{2}f(1)<0$. In either case, since f is continuous, the generalized Intermediate Value Theorem tells us that there exists a $c\in(0,1)$ such that $f(c)=\frac{1}{2}f(1)$.

Problem 3.13: If f is continuous at a and $f(a)>0$, show that there exists a real number $\delta>0$ such that $f(x)>0$ for all $x\in(a-\delta,a+\delta)$. Hint: Draw a picture.

Let $\epsilon=f(a)>0$. We know that there exists a $\delta>0$ such that

$$x\in(a-\delta,a+\delta)\Rightarrow|f(x)-f(a)|<\epsilon\Rightarrow-\epsilon<f(x)-f(a)\Rightarrow 0<f(x).$$

Chapter 4

Differentiation

4.A The Derivative and Its Properties

[Muldowney 1990, pp. 69–124]
[Spivak 1994, Chapter 9]

Definition: Let a be an interior point of the domain of a function f. If

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}$$

exists, then f is said to be *differentiable* at a. The limit is denoted $f'(a)$ and is called the *derivative* of f at a. If f is differentiable at every point a of its domain, we say that f is *differentiable*.

Written in this way, we see that the derivative is the limit of the *slope*

$$m(x)=\frac{f(x)-f(a)}{x-a}$$

of a *secant line* joining the points $(a,f(a))$ and $(x,f(x))$, where $x\neq a$. The limit is taken as x gets closer to a; that is,

$$f'(a)=\lim_{x\to a}m(x).$$

Problem 4.1: Show that the equation of the secant line joining the points $(a,f(a))$ and $(x,f(x))$ is given by

$$y=f(a)+m(x)(x-a),$$

where $m(x)$ is defined above. As x approaches a, we obtain the equation for the *tangent* line to the graph of $f(x)$ that goes through the point a:

$$y=f(a)+f'(a)(x-a).$$

Problem 4.2: Show that

$$\lim_{x\to a} f(x) = L \iff \lim_{h\to 0} f(a+h) = L.$$

Hence, the definition of the derivative may be written equivalently as

$$f'(a) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}.$$

- Let $f(t)$ be the position of a particle on a curve at time t. The *average velocity* of the particle between time t and $t+h$ is the ratio of the distance travelled over the time interval h:

  $$\frac{\text{change in position}}{\text{change in time}} = \frac{f(t+h)-f(t)}{h} \quad (h \neq 0).$$

  The *instantaneous velocity* at t is calculated by taking the limit as $h \to 0$:

  $$\lim_{h\to 0} \frac{f(t+h)-f(t)}{h} = f'(t).$$

- If $f(x) = c$, where c is a constant, then

  $$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = \lim_{h\to 0} \frac{c-c}{h} = 0 \quad \forall a \in \mathbb{R}.$$

- The derivative of the *affine* function $f(x) = mx + b$, where m and b are constants, (the graph of which is a straight line) has the constant value m:

  $$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = \lim_{h\to 0} \frac{m(a+h)-ma}{h} = m.$$

  In the case where $b = 0$, the function $f(x) = mx$ is said to be *linear*. A function that is neither linear nor affine is said to be *nonlinear*.

Remark: The derivative is the natural generalization of the slope of linear and affine functions to nonlinear functions. In general, the value of the local (or instantaneous) slope of a nonlinear function will depend on the point at which it is evaluated.

- Consider the function $f(x) = x^2$. Then

$$\begin{aligned} f'(a) &= \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \\ &= \lim_{h\to 0} \frac{(a+h)^2 - a^2}{h} \\ &= \lim_{h\to 0} \frac{\cancel{a^2} + 2ha + h^2 - \cancel{a^2}}{h} \\ &= \lim_{h\to 0} (2a + h) = 2a. \end{aligned}$$

We see here that the value of the derivative of f at the point a depends on a. Note that

$$\begin{aligned} f'(a) < 0 &\text{ for } a < 0, \\ f'(a) = 0 &\text{ for } a = 0, \\ f'(a) > 0 &\text{ for } a > 0. \end{aligned}$$

It is convenient to think of the derivative as a function on its own, which in general will depend on exactly where we evaluate it. We emphasize this fact by writing the derivative in terms of a dummy argument such as a or x. In this case, we can express this functional relationship as $f'(a) = 2a$ for all a, or with equal validity, $f'(x) = 2x$ for all x.

- Let $f(x) = x^n$, where $n \in \mathbb{N}$. Then

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} \\ &= \lim_{h\to 0} \frac{\cancel{x^n} + nx^{n-1}h + \frac{n(n-1)}{2} x^{n-2}h^2 + \ldots + h^n - \cancel{x^n}}{h} \\ &= \lim_{h\to 0} \left[nx^{n-1} + \frac{n(n-1)}{2} x^{n-2} h + \ldots + h^{n-1} \right] \\ &= nx^{n-1} \end{aligned}$$

(see also [Muldowney 1990, p.72, proof II]).

- For the function $f(x) = x^{1/n}$ where $x > 0$ and $n \in \mathbb{N}$, we can use the factorization of $y - x = (y^{1/n})^n - (x^{1/n})^n$ on p. 45 to compute the derivative:

$$\begin{aligned}
f'(x) &= \lim_{y\to x} \frac{f(y)-f(x)}{y-x} \\
&= \lim_{y\to x} \frac{y^{1/n}-x^{1/n}}{y-x} \\
&= \lim_{y\to x} \frac{y^{1/n}-x^{1/n}}{(y^{1/n}-x^{1/n})(y^{(n-1)/n}+y^{(n-2)/n}x^{1/n}+\ldots+y^{1/n}x^{(n-2)/n}+x^{(n-1)/n})} \\
&= \frac{1}{\underbrace{\lim_{y\to x} y^{(n-1)/n}+\lim_{y\to x} y^{(n-2)/n}x^{1/n}+\ldots+\lim_{y\to x} x^{(n-1)/n}}_{n \text{ terms}}} \\
&= \frac{1}{nx^{(n-1)/n}} = \frac{1}{n}x^{\frac{1-n}{n}} \\
&= \frac{1}{n}x^{\frac{1}{n}-1}.
\end{aligned}$$

Problem 4.3: Use the following procedure to show that the derivative of $\sin x$ is $\cos x$.

(a) Use the inequality $\sin x \le x \le \tan x$ for $0 \le x < \pi/2$ to prove that

$$\cos x \le \frac{\sin x}{x} \le 1 \text{ for } 0 < |x| < \frac{\pi}{2}.$$

(b) Prove that

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

(c) Prove that

$$1 - \cos x \le \frac{|x|^2}{2} \quad \forall x \in \mathbb{R}.$$

Hint: Try replacing x by $2x$.

(d) Prove that

$$\lim_{x\to 0} \frac{1-\cos x}{x} = 0.$$

(e) Use the above results to prove that $\sin x$ is differentiable at any real number a and find its derivative. That is, show that

$$\lim_{h\to 0} \frac{\sin(a+h)-\sin a}{h}$$

exists and evaluate the limit.

Q. Are all functions differentiable?

A. No, consider

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases}$$

We see that

$$\lim_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^+} \frac{1 - 1}{x} = \lim_{x \to 0^+} \frac{0}{x} = 0,$$

but

$$\lim_{x \to 0^-} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^-} \frac{0 - 1}{x} \nexists.$$

So $\lim_{x \to 0} \frac{f(x) - 1}{x - 0} \nexists$. It appears, at the very least, that we must avoid jumps, as the following theorem points out.

Theorem 4.1 (Differentiable $\Rightarrow$ Continuous): *If f is differentiable at a then f is continuous at a.*

Proof: For $x \neq a$, we may write

$$f(x) = f(a) + \frac{f(x) - f(a)}{x - a}(x - a).$$

If $\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$ exists, then

$$\begin{aligned} \lim_{x \to a} f(x) &= \lim_{x \to a} f(a) + \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \cdot \lim_{x \to a}(x - a) \\ &= f(a) + f'(a) \cdot 0 \\ &= f(a), \end{aligned}$$

so f is continuous at a.

Q. Are all **continuous** functions differentiable?

A. No, consider $f(x) = |x|$:

$$\frac{f(x) - f(0)}{x - 0} = \frac{|x| - 0}{x - 0} = \frac{|x|}{x} = \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x < 0. \end{cases}$$

Hence $\lim_{x \to 0} \frac{f(x) - f(0)}{x - 0}$ does not exist; f is not differentiable at 0, even though f is continuous at 0.

Derivative Notation

Three equivalent notations for the derivative have evolved historically. Letting $y = f(x)$, $\Delta y = f(x+h) - f(x)$, and $\Delta x = (x+h) - x = h$, we may write

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

To help us remember this, we sometimes denote the derivative by dy/dx (Leibniz notation).

The operator notation Df (or $D_x f$, which reminds us that the derivative is with respect to x) is also occasionally used to emphasize that the derivative Df is a function derived from the original function, f.

Theorem 4.2 (Properties of Differentiation): *If f and g are both differentiable at a, then*

(a) $(f+g)'(a) = f'(a) + g'(a)$,

(b) $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$,

(c) $\left(\dfrac{f}{g}\right)'(a) = \dfrac{f'(a)g(a) - f(a)g'(a)}{[g(a)]^2}$ *if* $g(a) \neq 0$.

Proof: We are given that $f'(a) = \lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a}$ $\exists$ and $g'(a) = \lim\limits_{x \to a} \dfrac{g(x) - g(a)}{x - a}$ $\exists$.

(a)

$$\begin{aligned}
\lim_{x \to a} \frac{(f+g)(x) - (f+g)(a)}{x - a} &= \lim_{x \to a} \frac{f(x) + g(x) - f(a) - g(a)}{x - a} \\
&= \lim_{x \to a} \frac{f(x) - f(a)}{x - a} + \lim_{x \to a} \frac{g(x) - g(a)}{x - a} \\
&= f'(a) + g'(a).
\end{aligned}$$

(b)

$$\begin{aligned}
&\lim_{x \to a} \frac{(fg)(x) - (fg)(a)}{x - a} = \lim_{x \to a} \frac{f(x)g(x) - f(a)g(a)}{x - a} \\
&= \lim_{x \to a} \frac{f(x)g(x) - f(a)g(x) + f(a)g(x) - f(a)g(a)}{x - a} \\
&= \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \underbrace{\lim_{x \to a} g(x)}_{\exists = g(a) \text{ by Theorem } 4.1} + f(a) \lim_{x \to a} \frac{g(x) - g(a)}{x - a} \\
&= f'(a)g(a) + f(a)g'(a).
\end{aligned}$$

(c) Let $h(x) = \dfrac{1}{g(x)}$. Then

$$\begin{aligned} h'(a) &= \lim_{x\to a} \frac{h(x)-h(a)}{x-a} = \lim_{x\to a} \frac{\frac{1}{g(x)} - \frac{1}{g(a)}}{x-a} \\ &= \lim_{x\to a} \frac{\frac{g(a)-g(x)}{g(x)g(a)}}{x-a} \\ &= -\frac{1}{g^2(a)} \lim_{x\to a} \frac{g(x)-g(a)}{x-a} = -\frac{g'(a)}{g^2(a)}. \end{aligned}$$

Then from **(b)**,

$$\begin{aligned} \left(\frac{f}{g}\right)'(a) = (fh)'(a) &= f'(a)h(a) + f(a)h'(a) \\ &= \frac{f'(a)}{g(a)} - \frac{f(a)g'(a)}{g^2(a)} \\ &= \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}. \end{aligned}$$

Corollary 4.2.1: Any polynomial is differentiable on $\mathbb{R}$.

Corollary 4.2.2: A rational function is differentiable at every point of its domain.

Problem 4.4: Compute

(a)

$$\frac{d}{dx}(x\cos x)$$

$$= \cos x - x\sin x.$$

(b)

$$\lim_{x\to\infty} x\tan\left(\frac{1}{x}\right)$$

Hint: let $y = 1/x$.

$$= \lim_{y\to 0^+} \frac{\tan y}{y} = \lim_{y\to 0^+} \left(\frac{\sin y}{y}\right) \lim_{y\to 0^+} \left(\frac{1}{\cos y}\right) = 1 \cdot 1 = 1.$$

Problem 4.5: Show that the rule $dx^n/dx = nx^{n-1}$ can be extended to $n \in \mathbb{Z}$.

For $n = 0$, the derivative evaluates to $\lim_{h\to 0} \frac{1-1}{h} = 0$. For $n < 0$, we have

$$\frac{d}{dx}x^n = \frac{d}{dx}\frac{1}{x^{-n}} = \frac{0 \cdot x^{-n} - 1 \cdot \frac{d}{dx}x^{-n}}{(x^{-n})^2} = \frac{-(-n)x^{-n-1}}{(x^{-n})^2} = nx^{n-1}.$$

Problem 4.6: (a) If $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$, find at least two different functions g such that $g' = f$.

(b) If

$$f(x) = \frac{b_2}{x^2} + \frac{b_3}{x^3} + \ldots + \frac{b_m}{x^m},$$

find a function g for which $g' = f$.

(c) Is there a function

$$g(x) = a_n x^n + \ldots + a_0 + \frac{b_1}{x} + \ldots + \frac{b_m}{x^m},$$

such that $g'(x) = 1/x$?

Theorem 4.3 (Chain Rule): *Suppose $h = f \circ g$, i.e. $h(x) = f(g(x))$. Let a be an interior point of the domain of h and define $b = g(a)$. If $f'(b)$ and $g'(a)$ both exist, then h is differentiable at a and*

$$h'(a) = f'(b)g'(a).$$

That is, if $y = f(u)$ and $u = g(x)$, then

$$\left.\frac{dy}{dx}\right|_a = \left.\frac{dy}{du}\right|_b \left.\frac{du}{dx}\right|_a.$$

- Consider that $\frac{d}{dx}(x^2+1)^2 = \frac{d}{dx}f(g(x))$, where $u = g(x) = x^2 + 1$ and $f(u) = u^2$. We let $h(x) = f(g(x))$:

$$\begin{aligned} h'(x) &= f'(u)g'(x) \\ &= 2u \cdot 2x \\ &= 2(x^2+1) \cdot 2x = 4x^3 + 4x. \end{aligned}$$

As a check, we could also work out this derivative directly:

$$\frac{d}{dx}(x^2+1)^2 = \frac{d}{dx}(x^4 + 2x^2 + 1) = 4x^3 + 4x.$$

- The Chain Rule makes it easy to find

$$\begin{aligned}\frac{d}{dx}(x^3+1)^{100} &= 100(x^3+1)^{99}3x^2\\ &= 300x^2(x^3+1)^{99}.\end{aligned}$$

- Let$f(u) = u^{\frac{1}{n}} \Rightarrow f'(u) = \frac{1}{n}u^{\frac{1}{n}-1}$ and $g(x) = x^m \Rightarrow g'(x) = mx^{m-1}$.

  Then $h(x) = f(g(x)) = x^{\frac{m}{n}} \Rightarrow h'(x) = f'(u)g'(x)$ where $u = g(x)$. Thus

$$\begin{aligned}h'(x) &= f'(g(x))g'(x)\\ &= \frac{1}{n}(x^m)^{\frac{1}{n}-1}mx^{m-1}\\ &= \frac{m}{n}x^{\frac{m}{n}-\cancel{m}+\cancel{m}-1}.\end{aligned}$$

Hence $\frac{d}{dx}x^q = qx^{q-1}$ for all $q \in \mathbb{Q}$.

- Find $\frac{d}{dx}\frac{1}{g(x)}$ (cf. Theorem 4.2(c)).

  Let $f(x) = x^{-1}$, $f'(x) = -x^{-2}$, and $h(x) = \frac{1}{g(x)} = f(g(x))$. Then

$$\begin{aligned}h'(x) &= f'(g(x))g'(x)\\ &= -\frac{1}{[g(x)]^2}g'(x),\end{aligned}$$

We may express this using an alternative notation. Letting $y = \frac{1}{u}$ and $u = g(x)$, we find

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = -\frac{1}{u^2}g'(x) = -\frac{g'(x)}{g^2(x)}.$$

-

$$\begin{aligned}\frac{d}{dx}\sqrt{\frac{1}{1+x^3}} &= \frac{1}{2\sqrt{\frac{1}{1+x^3}}}\left[-\frac{1}{(1+x^3)^2}\cdot 3x^2\right]\\ &= -\frac{3x^2(1+x^3)^{\frac{1}{2}}}{2(1+x^3)^2} = -\frac{3}{2}x^2(1+x^3)^{-\frac{3}{2}}.\end{aligned}$$

  Here is an even easier way to find this derivative:

$$\begin{aligned}\frac{d}{dx}\sqrt{\frac{1}{1+x^3}} &= \frac{d}{dx}(1+x^3)^{-\frac{1}{2}}\\ &= -\frac{1}{2}(1+x^3)^{-\frac{3}{2}}\cdot 3x^2.\end{aligned}$$

- $$\frac{d}{dx}\sin(\sin(x)) = \cos(\sin(x))\cos(x).$$

- $$\frac{d}{dx}\sin(\sin(\sin(x))) = \cos(\sin(\sin(x)))\cos(\sin(x))\cos(x).$$

- The derivative of $\cos x$ and $\tan x$ can be calculated as follows:

$$\frac{d}{dx}\cos x = \frac{d}{dx}\sin\left(\frac{\pi}{2} - x\right) = \cos\left(\frac{\pi}{2} - x\right)(-1) = -\sin x,$$

$$\frac{d}{dx}\tan x = \frac{d}{dx}\frac{\sin x}{\cos x} = \frac{\cos x\cos x - \sin x(-\sin x)}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x.$$

To prove the Chain Rule it is not enough to argue

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta u\to 0}\frac{\Delta y}{\Delta u}\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}$$

because $\Delta u = g(x) - g(a)$ might be zero for values of x close to (but not equal to) a. However, we can easily fix up this argument as follows.

Proof (of Theorem 4.3):

Let $b = g(a)$ and define

$$m(u) = \begin{cases} \dfrac{f(u) - f(b)}{u - b} & \text{if } u \neq b, \\[2ex] f'(b) & \text{if } u = b. \end{cases}$$

Then

$$f'(b)\exists \Rightarrow \lim_{u\to b} m(u) = f'(b) = m(b) \Rightarrow m \text{ is continuous at } b$$

and

$$g'(a)\exists \Rightarrow g \text{ is continuous at } a \Rightarrow m\circ g \text{ is continuous at } a \text{ (Corollary 3.1.7)},$$
$$\Rightarrow \lim_{x\to a} m(g(x)) = m(g(a)) = m(b) = f'(b).$$

Note that

$$f(u) - f(b) = m(u)(u - b)\ \forall u.$$

Letting $u = g(x)$, we then find that

$$\begin{aligned} \lim_{x\to a}\frac{f(g(x)) - f(g(a))}{x - a} &= \lim_{x\to a} m(g(x))\lim_{x\to a}\frac{g(x) - g(a)}{x - a} \\ &= f'(b)g'(a). \end{aligned}$$

Problem 4.7: Find

$$\frac{d}{dx}\left(\frac{1}{\cos x}\right)$$

$$= \left(\frac{1}{\cos^2 x}\right)\sin x.$$

Problem 4.8: Let f be a differentiable function. Find the following derivatives

(a)

$$\frac{d}{dx} f(f(f(x)))$$

$$= f'(f(f(x)))f'(f(x))f'(x).$$

(b)

$$\frac{d}{dx}\left[\frac{f^3(x)+1}{f^2(x)}\right]$$

$$= \frac{d}{dx}\left[f(x) + \frac{1}{f^2(x)}\right] = \left[1 - \frac{2}{f^3(x)}\right] f'(x).$$

Problem 4.9: Let

$$f(x) = \begin{cases} x^2 \cos\left(\frac{1}{x}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Prove that f is differentiable for all $x \in \mathbb{R}$ and find $f'(x)$.

Since $\cos(x)$ is differentiable at all x and $1/x$ is differentiable on $(-\infty, 0) \cup (0, \infty)$, the composite function $\cos(1/x)$, and hence f, is differentiable on $(-\infty, 0) \cup (0, \infty)$. Moreover, f is also differentiable at $x = 0$, with derivative 0:

$$\lim_{x \to 0} \frac{x^2 \cos\left(\frac{1}{x}\right) - 0}{x - 0} = \lim_{x \to 0} x \cos\left(\frac{1}{x}\right) = 0$$

since, given $\epsilon > 0$, we can make

$$\left| x \cos\left(\frac{1}{x}\right) \right| \leq |x| < \delta = \epsilon$$

whenever $|x| < \delta$, simply by choosing $\delta = \epsilon$.

Hence f is differentiable on $\mathbb{R}$ and

$$f'(x) = \begin{cases} 2x \cos\left(\frac{1}{x}\right) + \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

4.B Maxima and Minima

Definition: f has a *global maximum* (*global minimum*) at c if

$$f(x) \leq f(c) \quad (f(x) \geq f(c))$$

$\forall x$ in the domain of f.

Definition: A function f has an *interior local maximum* (*interior local minimum*) at an interior point c of its domain if for some $\delta > 0$,

$$\begin{aligned} x \in (c-\delta, c+\delta) \Rightarrow f(x) &\leq f(c) \\ (f(x) &\geq f(c)). \end{aligned}$$

Definition: An *extremum* is either a maximum or a minimum.

Remark: A global extremum is always a *local extremum* (but not necessarily an *interior local extremum*).

Theorem 4.4 (Interior Local Extrema (Maxima/Minima)): *Suppose*

(i) *f has an* interior local extremum *at c,*

(ii) *$f'(c)$ exists.*

Then $f'(c) = 0$.

Proof: Without loss of generality (why?) we can consider the case where f has an interior local maximum, i.e. $\exists \delta > 0 \ni$

$$\begin{aligned} x \in (c-\delta, c+\delta) &\Rightarrow f(x) \leq f(c) \\ &\Rightarrow \frac{f(x)-f(c)}{x-c} \begin{cases} \geq 0 & \text{if } x \in (c-\delta, c), \\ \leq 0 & \text{if } x \in (c, c+\delta) \end{cases} \\ &\Rightarrow f'_L(c) \doteq \lim_{x \to c^-} \frac{f(x)-f(c)}{x-c} \geq 0, \\ &\quad\; f'_R(c) \doteq \lim_{x \to c^+} \frac{f(x)-f(c)}{x-c} \leq 0 \\ &\quad \Rightarrow f'(c) = \lim_{x \to c} \frac{f(x)-f(c)}{x-c} = 0. \end{aligned}$$

Remark: Theorem 4.4 establishes that the condition $f'(c) = 0$ is *necessary* for a differentiable function to have an interior local extremum. However, this condition alone is not *sufficient* to ensure that a differentiable function has an extremum at c; consider the behaviour of the function $f(x) = x^3$ near the point $c = 0$.

Remark: If a function is continuous on a closed interval, we know from Theorem 3.4 that it must achieve global maximum and minimum values somewhere in the interval. We know from Theorem 4.4 that if these extrema occur in the interior of the interval, the derivative of the function must either vanish there or else not exist. However, it is possible that the global maximum or minimum occurs at one of the endpoints of the interval; at these points, it is not at all necessary that the derivative vanish, even if it exists. It is also possible that an extremum occurs at a point where the derivative doesn't exist. For example, consider the fact that $f(x) = |x|$ has a minimum at $x = 0$.

Extrema can occur either at

(i) an end point,

(ii) a point where f' does not exist,

(iii) a point where $f' = 0$.

- Find the maxima and minima of

$$f(x) = 2x^3 - x^2 + 1 \text{ on } [0, 1].$$

Since f is continuous on $[0, 1]$ we know that it has a global maximum and a global minimum on $[0, 1]$. Note that $f'(x) = 6x^2 - 2x = 2x(3x - 1) = 0$ in $(0, 1)$ only at the point $x = 1/3$. Theorem 4.4 implies that the only possible global interior extremum (which is of course also a local interior extremum) is at the point $x = 1/3$. By comparing the function values $f(1/3) = 26/27$ with the endpoint function values $f(0) = 1$ and $f(1) = 2$ we see that f has an (interior) global minimum value of $26/27$ at $x = 1/3$ and an (endpoint) global maximum value of 2 at $x = 1$. Hence $26/27 \leq f(x) \leq 2$ for all $x \in [0, 1]$.

- Determine the rectangle having the largest area that can be inscribed inside a right-angle triangle of side lengths a, b, and $\sqrt{a^2 + b^2}$, if the sides of the rectangle are constrained to be parallel to the sides of length a and b.

Let the vertices of the triangle be $(0, 0)$, $(a, 0)$, $(0, b)$ and those of the rectangle be $(0, 0)$, $(0, x)$, (x, y), $(0, y)$, where $0 \leq x \leq a$. By similar triangles we see that

$$\frac{y}{a - x} = \frac{b}{a}.$$

The area A of the rectangle is given by

$$A(x) = xy = \frac{b}{a}x(a - x) = bx - \frac{b}{a}x^2,$$

so that

$$A'(x) = b - \frac{2b}{a}x = \frac{b}{a}(a - 2x).$$

Since A is continuous on the closed interval $[0, a]$ we know that A must achieve maximum and minimum values in $[0, a]$. Since $A'(x)$ exists everywhere in $(0, a)$, the only points we need to check are $x = a/2$, where $A'(x) = 0$, and the endpoints $x = 0$ and $x = a$; at least one of these must represent a maximum area and one must represent a minimum area. Since $A(a/2) = ab/4$ and $A(0) = A(a) = 0$ we see that the maximum area is $ab/4$ and the minimum area is 0. Thus, the largest rectangle that can be inscribed has side lengths $a/2$ and $b/2$.

Corollary 4.4.1 (Rolle's Theorem): Suppose

(i) f is continuous on $[a, b]$,

(ii) f' exists on (a, b),

(iii) $f(a) = f(b)$.

Then there exists a number $c \in (a, b)$ for which $f'(c) = 0$.

Proof:

Case I: $f(x) = f(a) = f(b)\ \forall x \in [a, b]$ (i.e. f is constant on $[a, b]$)
$\Rightarrow f'(c) = 0\ \forall c \in (a, b)$.

Case II: $f(x_0) > f(a) = f(b)$ for some $x_0 \in (a, b)$. Theorem 3.4 $\Rightarrow f$ achieves its maximum value $f(c)$ for some $c \in [a, b]$. But

$$f(c) \geq f(x_0) > f(a) = f(b) \Rightarrow c \in (a, b).$$

$\therefore f$ has an interior local maximum at c.
Theorem 4.4 $\Rightarrow f'(c) = 0$.

Case III (Exercise): $f(x_0) < f(a) = f(b)$ for some $x_0 \in (a, b)$.

- $f(x) = x^3 - x + 1$.

  $f(0) = 1,\ f(1) = 1 \Rightarrow \exists c \in (0, 1) \ni f'(c) = 0$.
  In this case we can actually find the point c. Since $f'(x) = 3x^2 - 1$, we can solve the equation $0 = f'(c) = 3c^2 - 1$ to deduce $c = \dfrac{1}{\sqrt{3}} \in (0, 1)$.

- Recall that $\sin n\pi = 0, \forall n \in \mathbb{N}$. Rolle's Theorem tells us that $\frac{d}{dx} \sin x = \cos x$ must *vanish* (become zero) at some point $x \in (n\pi, (n + 1)\pi)$. Indeed, we know that

$$\cos\left[\left(n + \frac{1}{2}\right)\pi\right] = \cos\left(\frac{2n + 1}{2}\pi\right) = 0 \quad \forall n \in \mathbb{N}.$$

- We can use Rolle's Theorem to show that the equation

$$f(x) = x^3 - 3x^2 + k = 0$$

never has 2 distinct roots in $[0,1]$, no matter what value we choose for the real number k. Suppose that there existed two numbers a and b in $[0,1]$, with $a \neq b$ and $f(a) = f(b) = 0$. Then Rolle's Theorem $\Rightarrow \exists c \in (a,b) \subset (0,1)$ such that $f'(c) = 0$. But $f'(x) = 3x^2 - 6x = 3x(x-2)$ has no roots in $(0,1)$; this is a contradiction.

Q. What happens when the condition $f(a) = f(b)$ is dropped from Rolle's Theorem? Can we still deduce something similar?

A. Yes, the next corollary addresses precisely this situation.

Corollary 4.4.2 (Mean Value Theorem [MVT]): Suppose

(i) f is continuous on $[a,b]$,

(ii) f' exists on (a,b).

Then there exists a number $c \in (a,b)$ for which

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Remark: Notice that when $f(a) = f(b)$, the Mean Value Theorem reduces to Rolle's Theorem.

Proof: Consider the function

$$\varphi(x) = f(x) - M(x - a),$$

where M is a constant. Notice that $\varphi(a) = f(a)$. We choose M so that $\varphi(b) = f(a)$ as well:

$$M = \frac{f(b) - f(a)}{b - a}.$$

Then φ satisfies all three conditions of Rolle's Theorem:

(i) φ is continuous on $[a,b]$,

(ii) φ' exists on (a,b),

(iii) $\varphi(a) = \varphi(b)$.

Hence $\exists c \in (a,b)$ such that

$$0 = \varphi'(c) = f'(c) - M = f'(c) - \frac{f(a) - f(b)}{b - a}.$$

Q. We know that when $f(x)$ is constant that $f'(x) = 0$. Does the converse hold?

A. No, a function may have zero slope somewhere without being constant (e.g. $f(x) = x^2$ at $x = 0$). However, if $f'(x) = 0$ for all $x \in [a, b]$, where $a \neq b$, we may then make use of the following result.

Corollary 4.4.3 (Zero Derivative on an Interval): Suppose $f'(x) = 0$ for every x in an interval I (of nonzero length). Then f is constant on I.

Proof: Let x, y be any two elements of I, with $x < y$. Since f is differentiable at each point of I, we know by Theorem 4.1 that f is continuous on I. From the MVT, we see that

$$\frac{f(x) - f(y)}{x - y} = f'(c) = 0$$

for some $c \in (x, y) \subset I$. Hence $f(x) = f(y)$. Thus, f is constant on I.

Corollary 4.4.4 (Equal Derivatives): Suppose $f'(x) = g'(x)$ for every x in an interval I (of nonzero length). Then $f(x) = g(x) + k$ for all $x \in I$, where k is a constant.

Proof: Let $F(x) = f(x) - g(x)$. Then by Corollary 4.4.3,

$$F'(x) = 0 \Rightarrow F(x) = k$$

for some constant k.

Problem 4.10: Let f be a differentiable function on (a, b) and M be a real number such that $|f'(x)| \leq M$ for all $x \in (a, b)$. Let ϵ be a fixed positive number.

(a) Prove that

$$|f(x) - f(y)| \leq M\,|x - y| \quad \forall x, y \in (a, b).$$

Let x and y be two numbers chosen from (a, b) with $x < y$ (without loss of generality). Since f is continuous on $[x, y] \subset (a, b)$ and differentiable on (x, y), we can make use of the Mean Value Theorem to obtain the desired result:

$$\left|\frac{f(x) - f(y)}{x - y}\right| = |f'(c)| \leq M.$$

(b) Find a $\delta > 0$ such that

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon \quad \forall x, y \in (a, b).$$

Note here that a single δ must work for *all* x and y in the interval (a, b). This *uniform continuity* is a stronger condition than continuity at each point $x \in (a, b)$, as the latter allows δ to depend on x in addition to ϵ.

If $M = 0$, then $|f(x) - f(y)| = 0 < \epsilon$ for all x and y in (a, b) (any δ will work). Otherwise choose $\delta = \epsilon/M > 0$:

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| \leq M\,|x - y| < M\delta = \epsilon \quad \forall x, y \in (a, b).$$

4.C Monotonic Functions

Definition: A function is said to be *increasing* (*decreasing*) on an interval I if

$$x, y \in I, x \leq y \Rightarrow f(x) \leq f(y) \quad (f(x) \geq f(y))$$

and *strictly increasing* (*strictly decreasing*) if

$$x, y \in I, x < y \Rightarrow f(x) < f(y) \quad (f(x) > f(y)).$$

Note that strictly increasing $\Rightarrow$ increasing.

Definition: A function is said to be *monotonic* on an interval I if it is either (i) increasing on I; or (ii) decreasing on I.

Corollary 4.4.5 (Monotonic Functions): Suppose f is differentiable on an interval I. Then

(i) f is increasing on $I \iff f'(x) \geq 0$ on I;

(ii) f is decreasing on $I \iff f'(x) \leq 0$ on I.

Proof:
"$\Rightarrow$" Without loss of generality let f be increasing on I. Then for each $x \in I$,

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x} \geq 0.$$

"$\Leftarrow$" Suppose $f' \geq 0$ on I. Let $x, y \in I$ with $x < y$. The MVT $\Rightarrow \exists c \in (x, y) \ni$

$$\frac{f(y) - f(x)}{y - x} = f'(c) \geq 0$$
$$\Rightarrow f(y) - f(x) \geq 0.$$

Hence f is increasing on I.

Remark: Corollary 4.4.5 only provides sufficient, not necessary, conditions for a function to be increasing (since it might not be differentiable).

- Consider the function $f(x) = \lfloor x \rfloor$, which returns the greatest integer less than or equal to x. Note that f is increasing (on $\mathbb{R}$) but $f'(x) \nexists$ at integer values of x.

Q. If we replace "increasing" with "strictly increasing" in Corollary 4.4.5 (i), can we then change "$\geq$" to "$>$"?

A. No, consider the strictly increasing function $f(x) = x^3$. We can only say $f'(x) = 3x^2 \geq 0$ since $f'(0) = 0$.

Problem 4.11: Prove that if f is continuous on $[a, b]$ and $f'(x) > 0$ for all $x \in (a, b)$, then f is strictly increasing on $[a, b]$.

Corollary 4.4.6 (Horse-Race Theorem): Suppose

(i) f and g are continuous on $[a, b]$,

(ii) f' and g' exist on (a, b),

(iii) $f(a) \geq g(a)$,

(iv) $f'(x) \geq g'(x) \quad \forall x \in (a, b)$.

Then $f(x) \geq g(x) \quad \forall x \in [a, b]$.

Proof: Consider

$$\begin{aligned}
&\phi(x) = f(x) - g(x),\\
&\phi(a) \geq 0,\\
&\phi'(x) \geq 0 \quad \forall x \in (a, b).
\end{aligned}$$

Suppose $\exists x \in (a, b]$ such that $\phi(x) < 0$. But then

$$\text{MVT} \Rightarrow \exists c \in (a, x) \ni \phi'(c) = \frac{\phi(x) - \phi(a)}{x - a} < 0,$$

which contradicts the fact that $\phi'(x) \geq 0$ on (a, b). Hence $\phi(x) \geq 0$ on $[a, b]$. That is,

$$f(x) \geq g(x) \quad \forall x \in [a, b].$$

- Consider $f(x) = x$ and $g(x) = \sin x$ on $[0, b]$, where $b > 0$. Then

(i) holds,

(ii) holds,

(iii) $f(0) = g(0)$,

(iv) $f'(x) = 1 \geq \cos x = g'(x)\ \forall x \in (0, b)$.

Since we can choose b arbitrarily large, we might then be convinced from Corollary 4.4.6 that $x \geq \sin x$ for all $x \geq 0$. However, since we already used this result to compute the derivative of $\sin x$, this does not really constitute an independent proof of this important inequality.

4.D First Derivative Test

We have seen that points where the derivative of a function vanishes may or may not be extrema. How do we decide which ones are extrema and, of those, which are maxima and which are minima? One answer is provided by the First Derivative Test.

Definition: A point where the derivative of f is zero or does not exist is called a *critical* point.

Theorem 4.4 $\Rightarrow$ Local interior maxima and minima occur at critical points.

Remark: Not all critical points are extrema: consider $f(x) = x^3$ at $x = 0$.

Q. How do we decide which critical points c correspond to maxima, to minima, or neither?

A. If f is differentiable near c, look at the first derivative.

Corollary 4.4.7 (First Derivative Test): Suppose f is differentiable near a critical point c (except possibly at c, provided f is continuous at c). If there exists a $\delta > 0$ such that

(i) $f'(x) \begin{cases} \leq 0 & \forall x \in (c-\delta, c) \quad (f \text{ decreasing}), \\ \geq 0 & \forall x \in (c, c+\delta) \quad (f \text{ increasing}), \end{cases}$
then f has a local minimum at c;

(ii) $f'(x) \begin{cases} \geq 0 & \forall x \in (c-\delta, c) \quad (f \text{ increasing}), \\ \leq 0 & \forall x \in (c, c+\delta) \quad (f \text{ decreasing}), \end{cases}$
then f has a local maximum at c;

(iii) $f'(x) > 0$ on $(c-\delta, c) \cup (c, c+\delta)$ or $f'(x) < 0$ on $(c-\delta, c) \cup (c, c+\delta)$,
then f does not have a local extremum at c.

Problem 4.12: Give examples of differentiable functions which have the behaviours described in each of the cases above.

Definition: The derivative of f', that is $(f')'$, when it exists, is denoted either f'' or $f^{(2)}$ and is called the *second derivative* of f. In general, we denote the n-th derivative of f, the function obtained by differentiating f with respect to its argument n times, by $f^{(n)}$. The parentheses help us avoid confusion with exponent notation. It is also convenient to define $f^{(0)} = f$ itself.

Remark: Observe that $f^{(n+1)} = (f^{(n)})'$ and that if $f^{(n+1)}$ exists at a point a, then by Theorem 4.1 $f^{(n)}$ and all lower-order derivatives must exist and be continuous at a.

Problem 4.13: Suppose that two functions f and g are differentiable n times at the point a. Prove *Leibniz's formula*:

$$(fg)^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)}(x) g^{(k)}(x).$$

4.E Second Derivative Test

In cases where the second derivative of f can be easily computed, the following test provides simple conditions for classifying critical points.

Corollary 4.4.8 (Second Derivative Test): Suppose f is twice differentiable at a critical point c (this implies $f'(c) = 0$). If

(i) $f''(c) > 0$, then f has a local minimum at c;

(ii) $f''(c) < 0$, then f has a local maximum at c.

Proof:

(i) $f''(c) > 0 \Rightarrow \lim\limits_{x\to c} \dfrac{f'(x) - f'(c)}{x - c} > 0 \Rightarrow \lim\limits_{x\to c} \dfrac{f'(x)}{x - c} > 0$

$\Rightarrow\ \exists \delta > 0 \ni f'(x) \begin{cases} < 0 & \forall x \in (c - \delta, c), \\ > 0 & \forall x \in (c, c + \delta) \end{cases}$

$\Rightarrow f$ has a local minimum at c by the First Derivative Test.

(ii) Exercise.

Remark: If $f''(c) = 0$, then anything is possible.

- $f(x) = x^3$,
  $f'(x) = 3x^2 = 0$ at $x = 0$,
  $f''(x) = 6x = 0$ at $x = 0$,
  f has neither a maximum nor minimum at $x = 0$.

- $f(x) = x^4$,
  $f'(x) = 4x^3 = 0$ at $x = 0$,
  $f''(x) = 12x^2 = 0$ at $x = 0$,
  f has a minimum at $x = 0$.

- $f(x) = -x^4$ has a maximum at $x = 0$.

Remark: The First Derivative Test can sometimes be helpful in cases where the Second Derivative Test fails, e.g. in showing that $f(x) = x^4$ has a minimum at $x = 0$.

Remark: The Second Derivative Test establishes only the local behaviour of a function, whereas the First Derivative Test can sometimes be used to establish that an extremum is global:

$$f(x) = x^2, \qquad f'(x) = 2x \begin{cases} < 0 & \forall x < 0, \\ > 0 & \forall x > 0. \end{cases}$$

Since f is decreasing for $x < 0$ and increasing for $x > 0$, we see that f has a global minimum at $x = 0$.

Corollary 4.4.9 (Cauchy Mean Value Theorem): Suppose

(i) f and g are continuous on $[a, b]$,

(ii) f' and g' exist on (a, b).

Then there exists a number $c \in (a, b)$ for which

$$f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)].$$

Proof: Consider

$$\phi(x) = [f(x) - f(a)][g(b) - g(a)] - [f(b) - f(a)][g(x) - g(a)].$$

Note that ϕ is continuous on $[a, b]$ and differentiable on (a, b). Since $\phi(a) = \phi(b) = 0$, we know from Rolle's Theorem that $\phi'(c) = 0$ for some $c \in (a, b)$; from this we immediately deduce the desired result.

4.F L'Hôpital's Rule

Corollary 4.4.10 (L'Hôpital's Rule for $\frac{0}{0}$): Suppose f and g are differentiable on (a, b), $g'(x) \neq 0$ for all $x \in (a, b)$, $\lim_{x \to b^-} f(x) = 0$, and $\lim_{x \to b^-} g(x) = 0$. Then

$$\lim_{x \to b^-} \frac{f'(x)}{g'(x)} = L \Rightarrow \lim_{x \to b^-} \frac{f(x)}{g(x)} = L.$$

This result also holds if

(i) $\lim_{x\to b^-}$ is replaced by $\lim_{x\to a^+}$;

(ii) $\lim_{x\to b^-}$ is replaced by $\lim_{x\to\infty}$ and b is replaced by ∞;

(iii) $\lim_{x\to b^-}$ is replaced by $\lim_{x\to-\infty}$ and a is replaced by $-\infty$.

Proof: Theorem 4.1 $\Rightarrow$ f and g are continuous on (a, b). Consider

$$F(x) = \begin{cases} f(x) & a < x < b, \\ 0 & x = b. \end{cases}$$

$$G(x) = \begin{cases} g(x) & a < x < b, \\ 0 & x = b. \end{cases}$$

Since $\lim_{x\to b^-} f(x) = 0$, and $\lim_{x\to b^-} g(x) = 0$, we know for any $x \in (a, b)$ that F and G are continuous on $[x, b]$ and differentiable on (x, b). We can also be sure that G is nonzero on (a, b): if $G(x) = 0 = G(b)$ for some $x \in (a, b)$, Rolle's Theorem would imply that G', and hence g', vanishes somewhere in (x, b).

Given $\epsilon > 0$, we know there exists a number δ with $0 < \delta < b - a$ such that

$$x \in (b - \delta, b) \Rightarrow \left| \frac{f'(x)}{g'(x)} - L \right| < \epsilon.$$

If $x \in (b - \delta, b)$, Corollary 4.4.9 then implies that there exists a point $c \in (x, b)$ such that

$$\frac{f(x)}{g(x)} = \frac{F(x)}{G(x)} = \frac{F(x) - F(b)}{G(x) - G(b)} = \frac{F'(c)}{G'(c)} = \frac{f'(c)}{g'(c)},$$

so that

$$\left| \frac{f(x)}{g(x)} - L \right| = \left| \frac{f'(c)}{g'(c)} - L \right| < \epsilon.$$

That is, $\lim_{x\to b^-} \frac{f(x)}{g(x)} = L$.

- Using L'Hôpital's Rule, we find

$$\lim_{x\to 0} \frac{\tan x}{x} = 1 \Leftarrow \lim_{x\to 0} \frac{\sec^2 x}{1} = 1,$$

- $$\lim_{x\to 1} \frac{x^n - 1}{x - 1} = n \Leftarrow \lim_{x\to 1} \frac{nx^{n-1}}{1} = n.$$

Remark: L'Hôpital's Rule should only be used where it applies. For example, it should not be used for when the limit does not have the $\frac{0}{0}$ form. For example, $0 = \lim_{x\to 1} \frac{x-1}{x} \neq \lim_{x\to 1} \frac{1}{1} = 1$.

Definition: If for every $M > 0$ there exists a $\delta > 0$ such that $x \in (b-\delta, b) \Rightarrow f(x) > M$, we say $\lim_{x\to b-} f(x) = \infty$.

Corollary 4.4.11 (L'Hôpital's Rule for $\frac{\infty}{\infty}$): Suppose f and g are differentiable on (a,b), $g'(x) \neq 0$ for all $x \in (a,b)$, and $\lim_{x\to b^-} f(x) = \infty$, and $\lim_{x\to b^-} g(x) = \infty$. Then

$$\lim_{x\to b^-} \frac{f'(x)}{g'(x)} = L \Rightarrow \lim_{x\to b^-} \frac{f(x)}{g(x)} = L.$$

This result also holds if

(i) $\lim_{x\to b^-}$ is replaced by $\lim_{x\to a^+}$;

(ii) $\lim_{x\to b^-}$ is replaced by $\lim_{x\to\infty}$ and b is replaced by ∞;

(iii) $\lim_{x\to b^-}$ is replaced by $\lim_{x\to-\infty}$ and a is replaced by $-\infty$.

Proof: We only need to make minor modifications to the proof used to establish Corollary 4.4.10. Choose δ such that $f(x) > 0$ and $g(x) > 0$ on $(b-\delta, b)$ and redefine

$$F(x) = \begin{cases} \frac{1}{f(x)} & b-\delta < x < b, \\ 0 & x = b, \end{cases} \qquad G(x) = \begin{cases} \frac{1}{g(x)} & b-\delta < x < b, \\ 0 & x = b. \end{cases}$$

Problem 4.14: Determine which of the following limits exist as a finite number, which are ∞, which are $-\infty$, and which do not exist at all. Where possible, compute the limit.

(a)

$$\lim_{x\to 1} \frac{\sin(x^{99}) - \sin(1)}{x-1}$$

One could use L'Hôpital's Rule here, but it is even simpler to note that this is just the definition of the derivative of the function $f(x) = \sin(x^{99})$ at $x = 1$. Since

$$f'(x) = \cos(x^{99})99x^{98},$$

the limit reduces to $f'(1) = 99\cos(1)$.

(b)

$$\lim_{x\to\pi/4} \frac{\tan x - 1}{x - \pi/4}$$

Letting $f(x) = \tan x$, we see that this is just the definition of $f'(\pi/4) = \sec^2(\pi/4) = 2$.

(c)
$$\lim_{x\to 0}\frac{\tan x - x}{x^3}$$

$$= \lim_{x\to 0}\frac{\sec^2 x - 1}{3x^2} = \lim_{x\to 0}\frac{2\sec^3 x\sin x}{6x} = \lim_{x\to 0}\frac{-6\sec^2 x\sin^2 x + 2\sec^2 x\cos x}{6} = \frac{1}{3},$$

on applying the 0/0 form of L'Hôpital's Rule three times. Alternatively, after the second application of L'Hôpital's Rule, one can use the fact that $\lim_{x\to 0}\sin x/x = 1$.

4.G Taylor's Theorem

Corollary 4.4.12 (Taylor's Theorem): Let $n \in \mathbb{N}$. Suppose

(i) $f^{(n-1)}$ exists and is continuous on $[a, b]$,

(ii) $f^{(n)}$ exists on (a, b).

Then there exists a number $c \in (a, b)$ such that

$$f(b) = \sum_{k=0}^{n-1}\frac{(b-a)^k}{k!}f^{(k)}(a) + \underbrace{\frac{(b-a)^n}{n!}f^{(n)}(c)}_{R_n}.$$

That is,

$$f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2}f''(a) + \ldots + \frac{(b-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + R_n.$$

Remark: This is known as the *Taylor expansion* of f at b about a. The term R_n is known as the *remainder* after n terms.

- For $n = 1$:
$$f(b) = \frac{(b-a)^0}{0!}f(a) + \frac{(b-a)^1}{1!}f'(c)$$
i.e. $f(b) = f(a) + (b-a)f'(c)$ (MVT).

- For $n = 2$:
$$f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2}f''(c).$$

Remark: It is interesting to notice that for functions with continuous second derivatives, Theorem 4.4.12 implies the Second Derivative Test. If $f'(a) = 0$ and $f''(a) > 0$, then $\exists \delta > 0 \ni f''(c) > 0\ \forall c \in (a-\delta, a+\delta)$. Hence

$$|b-a| < \delta \Rightarrow f(b) = f(a) + \frac{(b-a)^2}{2} f''(c) \geq f(a),$$

i.e. f has a local minimum at a. Likewise, if $f'(a) = 0$ and $f''(a) < 0$, f has a maximum at a.

Proof (of Taylor's Theorem): We will apply Rolle's Theorem to

$$\varphi(x) = f(x) + \sum_{k=1}^{n-1} \frac{(b-x)^k}{k!} f^{(k)}(x) + M(b-x)^n,$$

where M is a constant. Noting that $\varphi(b) = f(b)$, we choose M so that $\varphi(a) = f(b)$ also:

$$f(b) = \varphi(a) = f(a) + \sum_{k=1}^{n-1} \frac{(b-a)^k}{k!} f^{(k)}(a) + M(b-a)^n. \tag{4.1}$$

That is, we choose

$$M = \frac{1}{(b-a)^n}\left[f(b) - f(a) - \sum_{k=1}^{n-1} \frac{(b-a)^k}{k!} f^{(k)}(a)\right].$$

Note that $\varphi(x)$ is continuous on $[a, b]$. Using the Chain Rule, we find that

$$\begin{aligned}
\varphi'(x) &= f'(x) + \sum_{k=1}^{n-1}\left[-\frac{(b-x)^{k-1}}{(k-1)!} f^{(k)}(x) + \frac{(b-x)^k}{k!} f^{(k+1)}(x)\right] - n(b-x)^{n-1}M \\
&= f'(x) - \sum_{k=\boxed{1}}^{n-1} \frac{(b-x)^{k-1}}{(k-1)!} f^{(k)}(x) + \sum_{k=2}^{\boxed{n}} \frac{(b-x)^{k-1}}{(k-1)!} f^{(k)}(x) - n(b-x)^{n-1}M \\
&= f'(x) - f'(x) + \frac{(b-x)^{n-1}}{(n-1)!} f^{(n)}(x) - n(b-x)^{n-1}M
\end{aligned}$$

exists $\forall x \in (a, b)$. We then apply Rolle's Theorem to deduce that there exists a number $c \in (a, b)$ such that

$$\begin{aligned}
0 = \varphi'(c) &= \frac{(b-c)^{n-1}}{(n-1)!} f^{(n)}(c) - n(b-c)^{n-1}M \\
\Rightarrow M &= \frac{1}{n!} f^{(n)}(c).
\end{aligned}$$

Upon substituting this result into Eq. (4.1), we obtain Taylor's Theorem:

$$f(b) = f(a) + \sum_{k=1}^{n-1} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^n}{n!} f^{(n)}(c).$$

Remark: If $\left|f^{(n)}(c)\right| \leq M$ for all c between b and a then $|R_n| \leq M\frac{|b-a|^n}{n!}$ on this same interval.

- Compute the first three digits of $\sin 1$ after the decimal point and determine its value correctly rounded to two digits.

Step 1: Let $f(x) = \sin x$. Choose a value a reasonably close to 1 at which the value of f and its derivatives are known, such as $a = 0$.

Step 2: Write down the Taylor expansion to enough terms so that $|R_n|$ is less than or equal to the allowed error. Set $b = x$.

$$\begin{aligned}\sin x &= \sin 0 + (x-0)\cos 0 - \frac{(x-0)^2}{2!}\sin 0 - \frac{(x-0)^3}{3!}\cos 0 + \frac{(x-0)^4}{4!}\sin 0 \\ &\quad + \frac{(x-0)^5}{5!}\cos 0 - \frac{(x-0)^6}{6!}\sin 0 + R_7, \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} + R_7,\end{aligned}$$

where $R_7 = -\frac{1}{7!}(x-0)^7\cos c$ for some $c \in (0,x)$. Since $|\cos c| < 1$ we know for $x = 1$ that

$$|R_7| < \frac{1}{7!} = \frac{1}{5040} < 0.0002$$

and

$$\sin 1 \approx 1 - \frac{1}{6} + \frac{1}{120} = \frac{101}{120} = 0.841\bar{6}.$$

Hence $\sin 1 \approx 0.841\bar{6} \pm 0.0002$, so the first three digits of $\sin 1$ are 0.841. If we round this result to two digits after the decimal place, we obtain $\sin 1 \approx 0.84$.

Definition: If $\lim\limits_{n\to\infty} R_n = 0$ in the Taylor expansion of f at b about a then

$$f(b) = \sum_{k=0}^{\infty} \frac{(b-a)^k}{k!} f^{(k)}(a)$$

This is known as the *Taylor Series* of f at b about a.

Definition: The special case of a Taylor Series about $a = 0$ is sometimes known as a *Maclaurin Series*.

Problem 4.15: Let $f(x) = \sqrt[3]{1+x}$.

(a) Determine the first three terms of the Taylor expansion of $f(x)$ about the point $a = 0$,

$$f(x) = \sum_{k=0}^{2} \frac{(x-a)^k}{k!} f^{(k)}(a) + R_3.$$

Recall that the remainder term R_3 is given by

$$R_3 = \frac{(x-a)^3}{3!} f^{(3)}(c).$$

We find

$$f^{(1)}(x) = \left(\frac{1}{3}\right)(1+x)^{-2/3},$$

$$f^{(2)}(x) = \left(\frac{1}{3}\right)\left(-\frac{2}{3}\right)(1+x)^{-5/3},$$

and

$$f^{(3)}(c) = \left(\frac{1}{3}\right)\left(-\frac{2}{3}\right)\left(-\frac{5}{3}\right)(1+c)^{-8/3}.$$

Hence

$$f(x) = 1 + \left(\frac{1}{3}\right)\frac{x}{1!} - \left(\frac{2}{9}\right)\frac{x^2}{2!} + R_3 = 1 + \frac{x}{3} - \frac{x^2}{9} + R_3.$$

(b) Use part (a) to find a lower bound for $\sqrt[3]{\frac{3}{2}}$ and show that your result approximates the exact value to within 1%. (You may leave your answer as a fraction.)

$$\sqrt[3]{\frac{3}{2}} = f\left(\frac{1}{2}\right) = 1 + \left(\frac{1}{3}\right)\frac{(\frac{1}{2})}{1!} - \left(\frac{2}{9}\right)\frac{(\frac{1}{2})^2}{2!} + R_3 = 1 + \frac{1}{6} - \frac{1}{36} + R_3 = \frac{41}{36} + R_3,$$

where

$$R_3 = \frac{(\frac{1}{2})^3}{3!} f^{(3)}(c).$$

for some number $c \in (0, \frac{1}{2})$. The third derivative of f at c can be easily bounded:

$$0 \le f^{(3)}(c) \le \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)\left(\frac{5}{3}\right) = \frac{10}{27},$$

so

$$0 \le R_3 \le \frac{(\frac{1}{2})^3}{3!}\left(\frac{10}{27}\right) = \frac{5}{24 \times 27} < \frac{1}{24 \times 5} < \frac{1}{100} < \frac{1}{100}\sqrt[3]{\frac{3}{2}}.$$

Thus $\sqrt[3]{\frac{3}{2}}$ lies in the interval

$$\left[\frac{41}{36}, \frac{41}{36} + \frac{1}{100}\right].$$

Problem 4.16: Find the Taylor series for $f(x) = \sin^2 x$ about $x = 0$. Hint: after computing the first derivative, simplify the result before proceeding to take further derivatives.

4.H Convex and Concave Functions

Definition: A function f is *convex* (sometimes called *concave up*) on an interval I if the secant line segment joining $(a, f(a))$ and $(b, f(b))$ lies on or above the graph of f for all $a, b \in I$.

Definition: A function f is *concave* (sometimes called *concave down*) on an interval I if $-f$ is convex on I.

Definition: An *inflection point* is a point on the graph of a function f at which the behaviour of f changes from convex to concave. For example, since $f(x) = x^3$ is concave on $(-\infty, 0]$ and convex on $[0, \infty)$, the point $(0, 0)$ is an inflection point.

Remark: Since the equation of the line through $(a, f(a))$ and $(b, f(b))$ is

$$y = f(a) + \frac{f(b) - f(a)}{b - a}(x - a),$$

the definition of convex says

$$f(x) \le f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \qquad \forall x \in [a, b], \quad \forall a, b \in I. \tag{4.2}$$

The *linear interpolation* of f between $[a, b]$ on the right-hand side of Eq. (4.2) may be rewritten as:

$$f(x) \le \left(\frac{b - x}{b - a}\right) f(a) + \left(\frac{x - a}{b - a}\right) f(b) \qquad \forall x \in [a, b], \quad \forall a, b \in I \tag{4.3}$$

or as

$$f(x) \le f(b) + \frac{f(b) - f(a)}{b - a}(x - b) \qquad \forall x \in [a, b], \quad \forall a, b \in I. \tag{4.4}$$

It is sometimes convenient to introduce the parameter $t = \dfrac{b - x}{b - a}$, in terms of which we may express $x = b - (b - a)t$ and

$$\frac{x - a}{b - a} = \frac{(b - a) - (b - a)t}{b - a} = 1 - t.$$

This allow us to restate Eq. (4.3) in *parametric form*:

$$f(ta+(1-t)b) \le tf(a)+(1-t)f(b) \qquad \forall t \in [0,1], \quad \forall a,b \in I. \tag{4.5}$$

The convexity condition may also be expressed directly in terms of the slope of a secant:

$$\frac{f(x)-f(a)}{x-a} \le \frac{f(b)-f(a)}{b-a} \le \frac{f(b)-f(x)}{b-x} \qquad \forall x \in (a,b), \quad \forall a,b \in I. \tag{4.6}$$

The left-hand inequality follows directly from Eq. (4.2) and the right-hand inequality follows from Eq. (4.4).

Theorem 4.5 (First Convexity Criterion): *Suppose f is differentiable on an interval I. Then*

(i) *f is convex $\iff$ f' is increasing on I;*

(ii) *f is concave $\iff$ f' is decreasing on I.*

Proof: Without loss of generality we only need to consider the case where f is convex.

"$\Rightarrow$" Suppose f is convex. Let $a,b \in I$, with $a<b$, and define

$$m(x)=\frac{f(x)-f(a)}{x-a} \quad (x \ne a), \qquad M(x)=\frac{f(b)-f(x)}{b-x} \quad (x \ne b).$$

From Eq. (4.6) we know that

$$m(x) \le m(b) = M(a) \le M(x)$$

whenever $a<x<b$. Hence

$$f'(a)=\lim_{x\to a} m(x) = \lim_{x\to a^+} m(x) \le m(b) = M(a) \le \lim_{x\to b^-} M(x) = \lim_{x\to b} M(x) = f'(b).$$

Thus f' is increasing on I.

"$\Leftarrow$" Suppose f' is increasing on I. Let $a,b \in I$, with $a<b$ and $x \in (a,b)$. By the MVT,

$$\frac{f(x)-f(a)}{x-a} = f'(c_1), \qquad \frac{f(b)-f(x)}{b-x} = f'(c_2)$$

for some $c_1 \in (a,x)$ and $c_2 \in (x,b)$. Since f' is increasing and $c_1<c_2$, we know that $f'(c_1) \le f'(c_2)$. Hence

$$\frac{f(x)-f(a)}{x-a} \le \frac{f(b)-f(x)}{b-x}$$

$$\Rightarrow f(x)\left[\frac{1}{x-a}+\frac{1}{b-x}\right] \le \frac{f(b)}{b-x}+\frac{f(a)}{x-a} = \frac{f(b)(x-a)+f(a)(b-x)}{(b-x)(x-a)},$$

which reduces to Eq. (4.3), so f is convex.

Corollary 4.5.1 (Second Convexity Criterion): Suppose f'' $\exists$ on an interval I. Then

(i) f is convex on $I \iff f''(x) \geq 0 \quad \forall x \in I$;

(ii) f is concave on $I \iff f''(x) \leq 0 \quad \forall x \in I$.

Proof: Apply Corollary 4.4.5 to f'.

Corollary 4.5.2 (Tangent to a Convex Function): If f is convex and differentiable on an interval I, the graph of f lies above the tangent line to the graph of f at every point of I.

Proof: Let $a \in I$. The equation of the tangent line to the graph of f at the point $(a, f(a))$ is $y = f(a) + f'(a)(x - a)$. Given $x \in I$, the MVT implies that $f(x) - f(a) = f'(c)(x - a)$, for some c between a and x. Since f is convex on I, we also know, from Theorem 4.5, that f' is increasing on I:

$$x < a \Rightarrow c < a \Rightarrow f'(c) \leq f'(a),$$

$$x > a \Rightarrow c > a \Rightarrow f'(c) \geq f'(a).$$

In either case $f(x) - f(a) = f'(c)(x - a) \geq f'(a)(x - a)$. Hence

$$f(x) \geq f(a) + f'(a)(x - a) \quad \forall x \in I.$$

Remark: For a function f with a continuous second derivative, Corollary 4.5.1 and 4.5.2 $\Rightarrow$ Second Derivative Test: $f''(a) > 0$ and f'' continuous $\Rightarrow f'' > 0$ in $(a - \delta, a + \delta)$ for some $\delta > 0$; this implies that f is *locally convex*. Given $f'(a) = 0$, we deduce $f(x) \geq f(a)$ for all $x \in (a - \delta, a + \delta)$.

Corollary 4.5.3 (Global Second Derivative Test): Suppose f is twice differentiable on I and $f'(c) = 0$ at some $c \in I$. If

(i) $f''(x) \geq 0 \quad \forall x \in I$, then f has a global minimum at c;

(ii) $f''(x) \leq 0 \quad \forall x \in I$, then f has a global maximum at c.

Proof: These results follow from Corollaries 4.5.1 and 4.5.2, upon noting that the tangent line to the graph of f at c is the line $y = f(c)$.

- Consider $f(x) = \dfrac{1}{1 + x^2}$ on $\mathbb{R}$.

  Observe that $f(0) = 1$ and $f(x) > 0$ for all $x \in \mathbb{R}$ and $\lim_{x \to \pm\infty} f(x) = 0$. Note that f is even: $f(-x) = f(x)$. Also, $1 + x^2 \geq 1 \Rightarrow f(x) \leq 1 = f(0)$, so f has a maximum at $x = 0$. Alternatively, we can use either the First Derivative Test or the Second Derivative Test to establish this. We find

$$f'(x) = -\frac{2x}{(1 + x^2)^2}$$

and,

$$f''(x) = \frac{-2}{(1+x^2)^2} + \frac{2(2x)2x}{(1+x^2)^3} = \frac{-2-2x^2+8x^2}{(1+x^2)^3} = \frac{2(3x^2-1)}{(1+x^2)^3}.$$

First Derivative Test: $\begin{cases} f'(x) > 0 \text{ on } (-\infty,0) \Rightarrow f \text{ is increasing on } (-\infty,0), \\ f'(x) < 0 \text{ on } (0,\infty) \Rightarrow f \text{ is decreasing on } (0,\infty) \end{cases}$
$\Rightarrow f$ has a maximum at 0.

Second Derivative Test: $f'(0) = 0, f''(0) = -2 < 0 \Rightarrow f$ has a maximum at 0.

$$\text{Convexity: } \begin{cases} f''(x) \geq 0 \text{ for } |x| \geq \frac{1}{\sqrt{3}}, \text{ i.e. } f \text{ is convex on } \left(-\infty, -\frac{1}{\sqrt{3}}\right] \cup \left[\frac{1}{\sqrt{3}}, \infty\right), \\ \\ f''(x) \leq 0 \text{ for } |x| \leq \frac{1}{\sqrt{3}}, \text{ i.e. } f \text{ is concave on } \left[-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right], \\ \\ f''(x) = 0 \text{ at } \pm\frac{1}{\sqrt{3}}; \text{ these correspond to inflection points.} \end{cases}$$

Problem 4.17: Consider the function $f(x) = (x+1)x^{2/3}$ on $[-1,1]$.

(a) Find $f'(x)$.
On rewriting $f(x) = x^{5/3} + x^{2/3}$, we find

$$f'(x) = \frac{5}{3}x^{2/3} + \frac{2}{3}x^{-1/3} = \frac{x^{-1/3}}{3}(5x+2) \quad (x \neq 0).$$

(b) Determine on which intervals f is increasing and on which intervals f is decreasing.
Since

$$f'(x) \begin{cases} > 0, & -1 \leq x < -2/5, \\ = 0, & x = -2/5, \\ < 0, & -2/5 < x < 0, \\ \nexists & x = 0, \\ > 0, & 0 < x \leq 1, \end{cases}$$

we know that f is increasing on $[-1,-2/5]$ and $[0,1]$. It is decreasing on $[-2/5,0]$.

(c) Does f have any interior local extrema on $[-1,1]$? If so, where do these occur? Which are maxima and which are minima?

Note that f has two critical points: $x = -2/5$ and $x = 0$. By the First Derivative Test, f has a local maximum at $x = -2/5$ and a local minimum at $x = 0$.

(d) What are the global minimum and maximum values of f and at what points do these occur?

On comparing the endpoint function values to the function values at the critical points, we conclude that f achieves its global minimum value of 0 at $x = -1$ and at $x = 0$. It has a global maximum value of 2 at $x = 1$.

(e) Determine on which intervals f is convex and on which intervals f is concave.
Since $f''(x) = \frac{10}{9}x^{-1/3} - \frac{2}{9}x^{-4/3} = \frac{2}{9}x^{-4/3}(5x-1)$, we see that

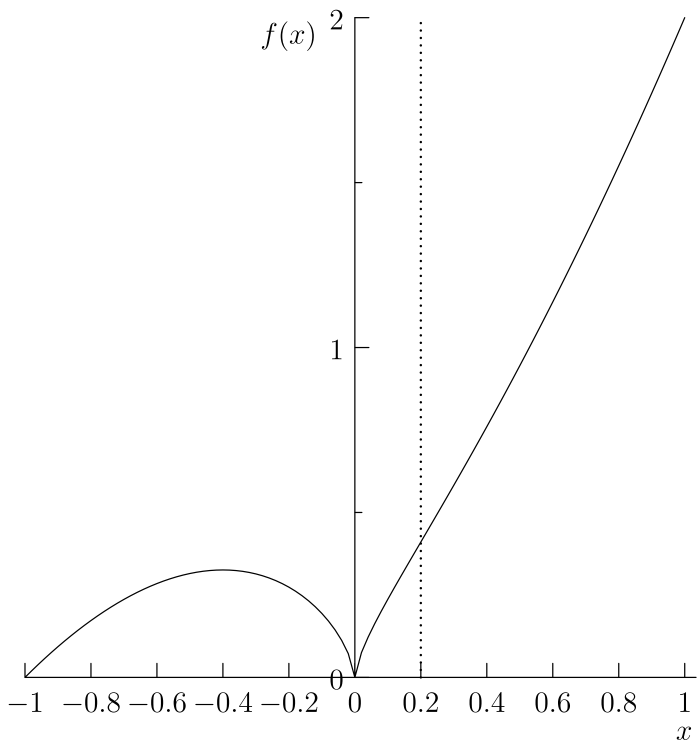
$$f''(x) \begin{cases} < 0, & -1 \le x < 0, \\ \nexists, & x = 0, \\ < 0, & 0 < x < 1/5, \\ = 0, & x = 1/5, \\ > 0, & 1/5 < x \le 1. \end{cases}$$

Thus, f is concave on $[-1, 0]$ and $[0, 1/5]$ and convex on $[1/5, 1]$. Note that f is not concave on the interval $[-1, 1/5]$.

(f) Does f have any inflection points? If so, where?

Yes: f has an inflection point at $x = 1/5$.

(g) Sketch a graph of f using the above information.



4.I Inverse Functions and Their Derivatives

[Spivak 1994, p. 227]
[Muldowney 1990, p. 103–111]

This section addresses the question: given a function f, when is it possible to find a function g that undoes the effect of f, so that

$$y = f(x) \iff x = g(y)?$$

Recall that a function is a collection of pairs of numbers (x, y) such that if (x, y_1) and (x, y_2) are in the collection, then $y_1 = y_2$.

Definition: A function $f : A \to B$ is *one-to-one* on its domain A if, whenever (x_1, y) and (x_2, y) are in the collection, then $x_1 = x_2$. That is,

$$x_1 = x_2 \iff f(x_1) = f(x_2).$$

We say that such a function is 1–1 or *invertible*.

This can be restated using the *horizontal line test*: a set of ordered pairs (x, y) is a one-to-one function if every horizontal and every vertical line intersects their graph at most once.

Remark: Equivalently, a 1–1 function f satisfies

$$x_1 \neq x_2 \iff f(x_1) \neq f(x_2).$$

- $f(x) = x$ and $f(x) = x^3$ are 1–1 functions.

- $f(x) = x^2$ and $f(x) = \sin x$ are not 1–1 functions.

Remark: Sometimes a *noninvertible* function can be made invertible by restricting its domain.

- $f = \sin x$ restricted to the domain $[-\frac{\pi}{2}, \frac{\pi}{2}]$ is 1–1.

Remark: If $f : A \to B$ is 1–1 then the collection of pairs of numbers (y, x) such that (x, y) belong to f is also a function.

Definition: The function defined by the pairs $\{(y, x) : (x, y) \in f\}$ is the *inverse* function $f^{-1} : B \to A$ of f.

Problem 4.18: Show that the inverse of a 1–1 function is itself an invertible function; that is, it satisfies both the horizontal and vertical line tests.

- The inverse of the function $\sin x$ restricted to the domain $[-\frac{\pi}{2}, \frac{\pi}{2}]$ is denoted $\arcsin x$ or $\sin^{-1} x$; it is itself a 1–1 function on $[-1, 1]$, yielding values in the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

Remark: Do not confuse the notation $\sin^{-1} x$ with $\frac{1}{\sin x}$; they are not the same function! Because of this rather unfortunate notational ambiguity, we will use the short-hand notation $f^n(x)$ to denote $(f(x))^n$ only when $n \geq 0$; in particular, we reserve the notation $f^{-1}(x)$ for the inverse of f.

Problem 4.19: Suppose that f and g are inverse functions of each other. Show that $g(f(x)) = x$ for all x in the domain of f and $f(g(y)) = y$ for all y in the range of f.

Theorem 4.6 (Continuous Invertible Functions): *Suppose f is continuous on I. Then f is one-to-one on $I \iff f$ is strictly monotonic on I.*

Proof:

"$\Rightarrow$" Let f be a continuous one-to-one function. If f were not strictly monotonic, we could find points $a, x, b \in I$ with $a < x < b$ such that

$$f(a) < f(x) \quad \text{and} \quad f(x) > f(b)$$

or

$$f(a) > f(x) \quad \text{and} \quad f(x) < f(b).$$

Since f is one-to-one on I, we know that $f(a) \neq f(b)$. Consider the first case. If $f(a) < f(b) < f(x)$, IVT $\Rightarrow$ there exists $c \in (a, x)$ such that $f(c) = f(b)$. If $f(b) < f(a) < f(x)$, IVT $\Rightarrow$ there exists $c \in (x, b)$ such that $f(c) = f(a)$. These statements contradict the fact that f is one-to-one on I. Upon reversing these inequalities, we obtain a contradiction for the second case also.

"$\Leftarrow$" Without loss of generality, suppose f is strictly increasing on I. Consider any points $a, b \in I$. Since $a < b \Rightarrow f(a) < f(b)$ and $a > b \Rightarrow f(a) > f(b)$, we see that $f(a) = f(b) \Rightarrow a = b$.

Problem 4.20: Prove that if f is continuous on $[a, b]$ and $f'(x) > 0$ for all $x \in (a, b)$, then f is 1–1 on $[a, b]$.

Corollary 4.6.1 (Continuity of Inverse Functions): Suppose f is continuous and one-to-one on an interval I. Then its inverse function f^{-1} is continuous on $f(I) = \{f(x) : x \in I\}$.

Proof: Theorem 4.6 $\Rightarrow$ f is strictly monotonic on I.

It is sufficient to consider only the case of a strictly increasing function:

$$x_1 < x_2 \iff f(x_1) < f(x_2).$$

Denoting $y_1 = f(x_1)$ and $y_2 = f(x_2)$ we can rewrite this as

$$f^{-1}(y_1) < f^{-1}(y_2) \iff y_1 < y_2.$$

Hence f^{-1} is also increasing. Note that this means that we can apply f^{-1} to both sides of an inequality.

Let a be an interior point of I. Given $\epsilon > 0$ small enough so that $[a-\epsilon, a+\epsilon] \subset I$, let $\delta = \min\{f(a) - f(a-\epsilon), f(a+\epsilon) - f(a)\} > 0$. The IVT implies that the image of the interval $[a-\epsilon, a+\epsilon]$ under the continuous function f is also an interval (cf. Corollary 3.4.1), so every point $y \in [f(a-\epsilon), f(a+\epsilon)]$ is in the domain of f^{-1}. Hence

$$\begin{aligned}
|y - f(a)| < \delta &\Rightarrow f(a-\epsilon) \leq f(a) - \delta < y < f(a) + \delta \leq f(a+\epsilon) \\
&\Rightarrow a - \epsilon < f^{-1}(y) < a + \epsilon \\
&\Rightarrow f^{-1}(f(a)) - \epsilon < f^{-1}(y) < f^{-1}(f(a)) + \epsilon \\
&\Rightarrow \left|f^{-1}(y) - f^{-1}(f(a))\right| < \epsilon.
\end{aligned}$$

Thus, f^{-1} is continuous at $f(a)$.

If a is an endpoint of I, a one-sided version of this argument shows that f^{-1} has the appropriate one-sided continuity.

Corollary 4.6.2 (Differentiability of Inverse Functions): Suppose f is continuous and one-to-one on an interval I and differentiable at $a \in I$. Let $b = f(a)$ and denote the inverse function of f on I by g. If

(i) $f'(a) = 0$, then g is **not** differentiable at b;

(ii) $f'(a) \neq 0$, then g is differentiable at b and $g'(b) = \dfrac{1}{f'(a)}$.

Proof:

(i) If g were differentiable at b then $g(f(x)) = x \Rightarrow g'(f(a))f'(a) = 1 \Rightarrow 0 = 1$, a contradiction.

(ii) For $x \neq a$, let $y = f(x)$ and consider

$$\frac{g(y) - g(b)}{y - b} = \frac{x - a}{f(x) - f(a)},$$

where $b = f(a)$ and $a = g(b)$.

Corollary 4.6.1 and Corollary 3.1.6 $\Rightarrow$ there corresponds to each sequence $\{y_n\} \to b$ a sequence $\{x_n\} \to a$, where $x_n = g(y_n)$. Theorem 3.1 then tells us that

$$\lim_{y \to b} \frac{g(y) - g(b)}{y - b} = \lim_{x \to a} \frac{x - a}{f(x) - f(a)} = \frac{1}{f'(a)};$$

that is,

$$\boxed{g'(b) = \frac{1}{f'(a)} = \frac{1}{f'(g(b))}.}$$

- The inverse of the function $f(x) = x^3$ is $f^{-1}(y) = y^{1/3}$ since $y = x^3 \Rightarrow x = y^{\frac{1}{3}}$. Notice that $f'(x) = 3x^2 \neq 0$ for $x \neq 0$ (i.e. $y \neq 0$). We can then verify that

$$\frac{d}{dy} f^{-1}(y) = \frac{1}{3} y^{-\frac{2}{3}} = \frac{1}{3y^{\frac{2}{3}}} = \frac{1}{3[f^{-1}(y)]^2} = \frac{1}{f'(f^{-1}(y))}.$$

- What is the derivative of $y = \arctan x$ (or $y = \tan^{-1} x$, the inverse function of $x = \tan y$?

Corollary 4.6.2 $\Rightarrow \dfrac{dy}{dx} = \dfrac{1}{\frac{dx}{dy}}$ where $x = \tan y$ and $\dfrac{dx}{dy} = \dfrac{1}{\cos^2 y}$. That is,

$$\frac{dy}{dx} = \frac{1}{\frac{1}{\cos^2 y}} = \cos^2 y.$$

Normally, we will want to re-express the derivative in terms of x. Recalling that $\tan^2 y + 1 = \frac{1}{\cos^2 y}$ and $x = \tan y$, we see that

$$\frac{dy}{dx} = \frac{1}{1 + \tan^2 y} = \frac{1}{1 + x^2}.$$

$$\therefore \boxed{\frac{d}{dx} \arctan x = \frac{1}{1 + x^2} \text{ on } (-\infty, \infty).}$$

Remark: Although $f(x) = \tan x$ does not satisfy the horizontal line test on $\mathbb{R}$, it does if we restrict $\tan x$ to the domain $(-\frac{\pi}{2}, \frac{\pi}{2})$. We call $\tan x$ on $(-\frac{\pi}{2}, \frac{\pi}{2})$ the *principal branch* of $\tan x$, which is sometimes denoted $\operatorname{Tan} x$. Its inverse, which is sometimes written $\operatorname{Arctan} x$ or $\operatorname{Tan}^{-1} x$, maps $\mathbb{R}$ to the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$.

Problem 4.21: Compute π to 1 digit accuracy by using the fact that $\pi = 4 \operatorname{Arctan} 1$ and considering the Taylor expansion for $\operatorname{Arctan} x$ about $a = 0$.

Remark: To go to 2 or more digits, you will probably want to write a small program to sum up the series you get. This is not a very efficient way of computing π!

- Consider $f(x) = \sqrt{1 - x^2}$, which is 1–1 on $[0, 1]$.

  Note that $f'(x) = -\frac{x}{\sqrt{1-x^2}}$ exists on $[0, 1)$.
  Now $y = \sqrt{1 - x^2} \Rightarrow x = \sqrt{1 - y^2} \Rightarrow x = f^{-1}(y) = f(y)$.
  In this case f and f^{-1} are identical functions of their respective arguments!

$$
\begin{aligned}
\frac{d}{dy} f^{-1}(y) &= \frac{1}{f'(x)} \\
&= -\frac{\sqrt{1 - x^2}}{x} \\
&= -\frac{\sqrt{1 - [f^{-1}(y)]^2}}{f^{-1}(y)} \\
&= -\frac{\sqrt{1 - [f(y)]^2}}{f(y)} \\
&= -\frac{\sqrt{1 - (1 - y^2)}}{\sqrt{1 - y^2}} \\
&= -\frac{y}{\sqrt{1 - y^2}} \qquad \text{on } [0, 1).
\end{aligned}
$$

- $y = \sin x$ is 1–1 on $\left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$.

  $\dfrac{dy}{dx} = \cos x \neq 0$ on $(-\frac{\pi}{2}, \frac{\pi}{2})$.
  The inverse function (as a function of y) is

$$
x = \arcsin y \ (\text{or } x = \sin^{-1} y),
$$

with derivative

$$
\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{\cos x}.
$$

We can express $\cos x$ as a function of y:

$$\begin{aligned}\cos x &= \sqrt{1-\sin^2 x}\\ &= \sqrt{1-y^2},\end{aligned}$$

noting that $\cos x > 0$ on $(-\frac{\pi}{2}, \frac{\pi}{2})$, to find

$$\frac{d}{dy}\arcsin y = \frac{1}{\sqrt{1-y^2}} \text{ on } (-1,1).$$

That is,

$$\boxed{\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1-x^2}} \text{ on } (-1,1).}$$

- $y = \cos x$ is 1–1 on $[0, \pi]$.

  $\dfrac{dy}{dx} = -\sin x \neq 0$ on $(0, \pi)$.

  The inverse function $x = \arccos y$ (or $x = \cos^{-1} y$) has derivative

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{-\sin x},$$

which we can express as a function of y, noting that $\sin x > 0$ on $(0, \pi)$,

$$\sin x = \sqrt{1-\cos^2 x} = \sqrt{1-y^2}.$$

$$\therefore\ \frac{d}{dy}\arccos y = -\frac{1}{\sqrt{1-y^2}},$$

i.e.

$$\boxed{\frac{d}{dx}\arccos x = -\frac{1}{\sqrt{1-x^2}} \text{ on } (-1,1).}$$

It is not surprising that $\frac{d}{dx}\arccos x = -\frac{d}{dx}\arcsin x$ since $\arccos x = \frac{\pi}{2} - \arcsin x$, as can readily be seen by taking the cosine of both sides and using $\cos y = \sin(\frac{\pi}{2} - y)$.

- Prove that $\cos^{-1} x + \sin^{-1} x = \frac{\pi}{2}$ for all $x \in [-1, 1]$. Let

$$\begin{aligned}f(x) &= \cos^{-1} x + \sin^{-1} x\\ f'(x) &= \frac{-1}{\sqrt{1-x^2}} + \frac{1}{\sqrt{1-x^2}} = 0\\ \Rightarrow f(x) &= c, \quad \text{a constant.}\end{aligned}$$

Set $x = 0$ to find c:

$$c = f(0) = \cos^{-1} 0 = \frac{\pi}{2}.$$

$$\therefore\ f(x) = \frac{\pi}{2} \text{ for all } x \in [-1,1].$$

Problem 4.22: Let $f(x) = \sin^{-1}(x^2 - 1)$. Find

(a) the domain of f;

The inverse function $y = \sin^{-1} x$ has domain $[-1, 1]$, and $x^2 - 1 \in [-1, 1]$ implies $x^2 \in [0, 2]$. Hence, the domain of f is $[-\sqrt{2}, \sqrt{2}]$.

(b) $f'(x)$;

Letting $y = \sin^{-1}(x^2 - 1)$, we first find the derivative for $x > 0$:

$$\begin{aligned}
x^2 - 1 &= \sin y \\
\Rightarrow x &= \sqrt{\sin y + 1} \\
\Rightarrow \frac{dx}{dy} &= \frac{\cos y}{2\sqrt{\sin y + 1}} \\
&= \frac{\sqrt{1 - (x^2 - 1)^2}}{2\sqrt{x^2}} = \frac{\sqrt{2x^2 - x^4}}{2x} \\
\Rightarrow \frac{dy}{dx} &= \frac{2x}{\sqrt{2x^2 - x^4}}.
\end{aligned}$$

Since the derivative of an even function is odd (and *vice-versa*) we see that the same result holds for $x < 0$ as well.

Alternatively, one could use the formula for the derivative of $\sin^{-1} x$ together with the Chain Rule.

(c) the domain of f'.

The domain of $f' = dy/dx$ is the set of x such that $2x^2 - x^4 > 0$:

$$2x^2 > x^4 \Rightarrow 2 > x^2 \quad \text{if } x \neq 0.$$

$\therefore$ domain of f' is $\left\{x : 0 < |x| < \sqrt{2}\right\} = \left(-\sqrt{2}, 0\right) \cup \left(0, \sqrt{2}\right)$.

Problem 4.23: Suppose that there exists a positive differentiable function $y = f(x)$ on $\mathbb{R}$ such that $f'(x) = f(x)$ for all $x \in \mathbb{R}$.

(a) Prove that f is one-to-one.

As we have previously seen from the Mean Value Theorem, $f'(x) > 0$ implies that f is strictly increasing. Hence, f is one-to-one.

(b) Prove that the inverse function g given by $x = g(y)$ is differentiable.

This follows from Corollary 4.6.2 since $f'(x) > 0$ for all $x \in \mathbb{R}$.

(c) Compute $g'(y)$. As usual, express your answer as a function of the argument y.

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{f(x)} = \frac{1}{y}.$$

This is an example of a function that differentiates to $1/y$, as we sought earlier. In Math 118, we will establish the existence of $g(y)$, which we will come to know as the *logarithmic* function $\log(y)$ (we have already seen that g cannot be written as a finite sum of powers of y).

4.J Implicit Differentiation

Suppose that a variable y is defined implicitly in terms of x and we wish to know dy/dx. For example, given the *implicit equation*

$$y^3 + 3y^2 + 3y + 1 = x^5 + x, \tag{4.7}$$

we could solve for y to find

$$\begin{aligned}(y+1)^3 &= x^5 + x \\ \Rightarrow y + 1 &= (x^5 + x)^{\frac{1}{3}} \\ \Rightarrow \frac{dy}{dx} &= \frac{1}{3}(x^5 + x)^{-\frac{2}{3}}(5x^4 + 1).\end{aligned} \tag{4.8}$$

But what happens if you can't (or don't want to) solve for y? You might try first to solve for x in terms of y and then find the derivative dx/dy of the inverse function. But what if this is also difficult?

It is often easier in these cases to differentiate both sides of Eq. (4.7) with respect to x, noting that $y = y(x)$:

$$\frac{d}{dx}\left[y^3(x) + 3y^2(x) + 3y(x) + 1\right] = \frac{d}{dx}(x^5 + x).$$

By the Chain Rule, we find

$$\big(3y^2 + 6y + 3\big)y'(x) = 5x^4 + 1,$$

which we can easily solve to obtain dy/dx as a function of x and y,

$$\frac{dy}{dx} = \frac{5x^4 + 1}{3y^2 + 6y + 3} = \frac{5x^4 + 1}{3(y+1)^2}. \tag{4.9}$$

Once we know an (x, y) pair that satisfies Eq. (4.7), we can immediately compute the derivative from Eq. (4.9).

It is instructive to verify that Eqs. (4.8) and (4.9) agree:

$$\frac{dy}{dx} = \frac{5x^4 + 1}{3(y+1)^2} = \frac{5x^4 + 1}{3(x^5 + x)^{\frac{2}{3}}}.$$

Problem 4.24: Suppose that f and its inverse g are twice differentiable functions on $\mathbb{R}$. Let $a \in \mathbb{R}$ and denote $b = f(a)$.

(a) Implicitly differentiate both sides of the identity $g(f(x)) = x$ with respect to x.
By the Chain Rule,

$$g'(f(x))f'(x) = 1.$$

(b) Using part(a), prove that $f'(a) \neq 0$.
If $f'(a) = 0$, we would obtain a contradiction:

$$0 = g'(f(a))f'(a) = 1.$$

(c) Using parts (a) and (b), find a formula expressing $g'(b)$ in terms of $f'(a)$.

$$g'(b) = \frac{1}{f'(a)}.$$

(d) Show that

$$g''(b) = -\frac{f''(a)}{[f'(a)]^3}.$$

On differentiating the expression in part (a), we find that

$$g''(f(x))[f'(x)]^2 + g'(f(x))f''(x) = 0.$$

On setting $x = a$ and using part(c), we find that

$$g''(f(a))[f'(a)]^2 + \frac{f''(a)}{f'(a)} = 0,$$

from which the desired result immediately follows.

Chapter 5

Integration

[Muldowney 1990, pp. 125]
[Spivak 1994, Chapter 13]

5.A The Riemann Integral

Suppose, given a bounded function $f(x) \geq 0$ on $[a, b]$, that we wish to determine the area of the region

$$\mathcal{S} = \{(x, y) : a \leq x \leq b,\ 0 \leq y \leq f(x)\}.$$

That is, we want to find the area of the region bounded by the graph of $f(x)$, the x axis, and the lines $x = a$ and $x = b$.

Consider a union $\mathcal{L}$ of non-overlapping rectangles that are each **contained within** $\mathcal{S}$, as illustrated in Figure 5.1. Also consider a union $\mathcal{U}$ of non-overlapping rectangles that together **contain** $\mathcal{S}$, as illustrated in Figure 5.2.

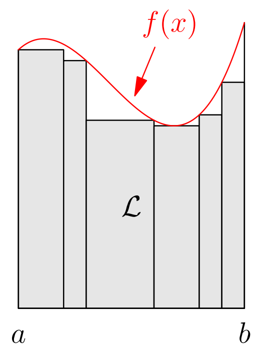


Figure 5.1: Lower rectangles

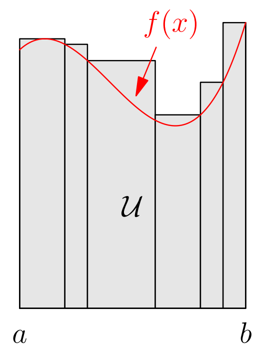


Figure 5.2: Upper rectangles

Notice that

$$\text{area}\ \mathcal{L} \leq \text{area}\ \mathcal{S} \leq \text{area}\ \mathcal{U}.$$

We now use this construction to give a precise definition for the concept of *area*.

Definition: If there is a **unique** number α satisfying

$$\text{area}\,\mathcal{L} \le \alpha \le \text{area}\,\mathcal{U}$$

for all such rectangular unions, we define

$$\text{area}\,\mathcal{S} = \alpha$$

and write $\alpha = \int_a^b f$, the Riemann integral over $[a, b]$ of f.

The following definitions will help us formalize this notion.

Definition: Let $[a, b]$ be a closed interval. A *partition* of $[a, b]$ is a finite set of distinct points from $[a, b]$ that includes the endpoints a and b. It is convenient to list the points of a partition $P = \{x_0, x_1, \ldots, x_n\}$ in increasing order: $a = x_0 < x_1 < \ldots < x_n = b$.

Definition: A partition Q of $[a, b]$ is a *refinement* of a partition P of $[a, b]$ if $Q \supset P$.

Definition: Let f be a bounded function on $[a, b]$ and $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. For each interval $i = 1, 2, \ldots, n$ we let

$$\begin{aligned} m_i &= \inf\,\{f(x) : x_{i-1} \le x \le x_i\}, \\ M_i &= \sup\,\{f(x) : x_{i-1} \le x \le x_i\}. \end{aligned}$$

Then we define

$$\begin{aligned} \mathcal{L}(P, f) &\doteq m_1(x_1 - x_0) + m_2(x_2 - x_1) + \ldots + m_n(x_n - x_{n-1}), \\ \mathcal{U}(P, f) &\doteq M_1(x_1 - x_0) + M_2(x_2 - x_1) + \ldots + M_n(x_n - x_{n-1}) \end{aligned}$$

to be the *lower* and *upper* sums, respectively, of f with respect to P.

Remark: Since $m_i \le M_i$ for each $i = 1, 2, \ldots, n$, we see for any partition P that $\mathcal{L}(P, f) \le \mathcal{U}(P, f)$.

Lemma 5.1 (Partition Refinement): *If P and Q are partitions of $[a, b]$ such that $Q \supset P$, then*

$$\mathcal{L}(P, f) \le \mathcal{L}(Q, f) \le \mathcal{U}(Q, f) \le \mathcal{U}(P, f).$$

That is, refinement increases lower sums and decreases upper sums.

Proof: It is sufficient to prove the lemma when Q contains just one more point that P. (Why?) Then for some $n \in \mathbb{N}$ and $k \in \{1, 2, \ldots, n\}$ we have

$$P = \{x_0, x_1, \ldots, x_{k-1}, x_k, \ldots, x_n\},$$

$$Q = \{x_0, x_1, \ldots, x_{k-1}, q, x_k, \ldots, x_n\}.$$

Recall

$$m_k = \inf\{f(x) : x \in [x_{k-1}, x_k]\}.$$

Let

$$\begin{aligned} m'_k &= \inf\{f(x) : x \in [x_{k-1}, q]\}, \\ m''_k &= \inf\{f(x) : x \in [q, x_k]\}. \end{aligned}$$

Note that $m_k \le m'_k$ and $m_k \le m''_k$. Then

$$\begin{aligned} \mathcal{L}(Q, f) &= \sum_{i=1}^{k-1} m_i(x_i - x_{i-1}) + m'_k(q - x_{k-1}) + m''_k(x_k - q) + \sum_{i=k+1}^{n} m_i(x_i - x_{i-1}) \\ &\ge \sum_{i=1}^{k-1} m_i(x_i - x_{i-1}) + m_k(q - x_{k-1}) + m_k(x_k - q) + \sum_{i=k+1}^{n} m_i(x_i - x_{i-1}) \\ &= \sum_{i=1}^{k-1} m_i(x_i - x_{i-1}) + m_k(x_k - x_{k-1}) + \sum_{i=k+1}^{n} m_i(x_i - x_{i-1}) = \mathcal{L}(P, f). \end{aligned}$$

The proof for upper sums follows upon replacing inf with sup and reversing the sign of the inequalities.

Lemma 5.2 (Upper Sums Bound Lower Sums): *Let f be bounded on $[a, b]$. If P and Q are* **any** *partitions of $[a, b]$, then*

$$\mathcal{L}(P, f) \le \mathcal{U}(Q, f).$$

That is, any upper sum of f is an upper bound for all lower sums and any lower sum of f is a lower bound for all upper sums.

Proof: Consider the partition $P \cup Q$. Then $P \cup Q$ is a refinement of P and also of Q:

$$P \subset P \cup Q, \quad Q \subset P \cup Q.$$

Hence Lemma 5.1 $\Rightarrow$

$$\mathcal{L}(P, f) \le \mathcal{L}(P \cup Q, f) \le \mathcal{U}(P \cup Q, f) \le \mathcal{U}(Q, f).$$

Definition: The *lower integral* and *upper integral* of f are, respectively,

$$\begin{aligned} \underline{\int_a^b} f &= \sup\{\mathcal{L}(P, f) : \forall \text{ partitions } P \text{ of } [a, b]\}, \\ \overline{\int_a^b} f &= \inf\{\mathcal{U}(P, f) : \forall \text{ partitions } P \text{ of } [a, b]\}. \end{aligned}$$

Q. Do the sup and inf here always exist? Note that the set of lower sums is nonempty and bounded above by any given upper sum (Lemma 5.2).

Lemma 5.3 (Lower Integrals vs. Upper Integrals): *Let f be bounded on $[a,b]$. Then*

$$\underline{\int_a^b} f \leq \overline{\int_a^b} f.$$

Proof: We first show that $\overline{\int_a^b} f$ is an upper bound for all lower sums. Consider any lower sum. Being a lower bound for the entire set of upper sums (by Lemma 5.2), it must be less than or equal to the greatest lower bound $\overline{\int_a^b} f$ of the set of upper sums. That is, $\overline{\int_a^b} f$ provides an upper bound for **all** lower sums.

But $\underline{\int_a^b} f$ is the **least** upper bound of all lower sums. Hence

$$\underbrace{\underline{\int_a^b} f}_{\sup\limits_P \{\mathcal{L}(P,f)\}} \quad \leq \quad \underbrace{\overline{\int_a^b} f}_{\inf\limits_P \{\mathcal{U}(P,f)\}}.$$

Definition: If

$$\underline{\int_a^b} f = \overline{\int_a^b} f = \alpha,$$

we say that f is *Riemann integrable* on $[a,b]$ and define $\int_a^b f \doteq \alpha$. The number $\int_a^b f$ is called the *Riemann integral* of f on $[a,b]$.

Since, for any partition P,

$$\mathcal{L}(P,f) \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq \mathcal{U}(P,f),$$

this definition is seen to be equivalent to our previous definition of the Riemann integral. That is, if there exists a **unique** number α such that

$$\mathcal{L}(P,f) \leq \alpha \leq \mathcal{U}(P,f)$$

for all partitions P of $[a,b]$, then f is Riemann integrable on $[a,b]$ and $\int_a^b f \doteq \alpha$.

- Consider $f(x) = c, \quad a \le x \le b$.
  Consider $P = \{a, b\}$. Then

$$\mathcal{L}(P, f) = c(b-a), \qquad \mathcal{U}(P, f) = c(b-a)$$

and

$$\underline{\int_a^b} f = \sup\{\mathcal{L}(Q, f) : Q \text{ is a partition of } [a, b]\} \ge \mathcal{L}(P, f) = c(b-a),$$

$$\overline{\int_a^b} f = \inf\{\mathcal{U}(Q, f) : Q \text{ is a partition of } [a, b]\} \le \mathcal{U}(P, f) = c(b-a),$$

so that

$$\underline{\int_a^b} f \ge \overline{\int_a^b} f.$$

But Lemma 5.3 $\Rightarrow$

$$\underline{\int_a^b} f \le \overline{\int_a^b} f.$$

Thus

$$\underline{\int_a^b} f = \overline{\int_a^b} f = c(b-a) \Rightarrow \int_a^b f = c(b-a).$$

Remark: By definition, for $\int_a^b f$ to exist it is necessary that f have an infimum and supremum on all possible subintervals of $[a, b]$. That is, a function must be bounded on $[a, b]$ in order to be Riemann integrable on $[a, b]$.

Theorem 5.1 (Integrability): *$\int_a^b f$ exists and equals $\alpha \iff$ there exists a sequence of partitions $\{P_n\}_{n=1}^{\infty}$ of $[a, b]$ such that*

$$\lim_{n\to\infty} \mathcal{L}(P_n, f) = \alpha = \lim_{n\to\infty} \mathcal{U}(P_n, f).$$

Proof:

"$\Rightarrow$" Suppose $\int_a^b f = \alpha$. That is,

$$\underline{\int_a^b} f = \overline{\int_a^b} f = \alpha.$$

This means that for each $n \in \mathbb{N}$, there exist partitions Q_n, R_n of $[a, b]$ such that

$$\alpha - \frac{1}{n} < \mathcal{L}(Q_n, f) \le \alpha$$ (otherwise α would not be the least upper bound of the set of all lower sums)

$$\alpha \le \mathcal{U}(R_n, f) < \alpha + \frac{1}{n}.$$

Let $P_n = Q_n \cup R_n$. Note $P_n \supset Q_n$ and $P_n \supset R_n$.
Lemma 5.1 $\Rightarrow$

$$\mathcal{L}(Q_n, f) \le \mathcal{L}(P_n, f) \le \mathcal{U}(P_n, f) \le \mathcal{U}(R_n, f).$$

Hence

$$\alpha - \frac{1}{n} < \mathcal{L}(P_n, f) \le \mathcal{U}(P_n, f) < \alpha + \frac{1}{n}.$$

The Squeeze Principle then guarantees that the sequences $\{\mathcal{L}(P_n, f)\}_{n=1}^{\infty}$, $\{\mathcal{U}(P_n, f)\}_{n=1}^{\infty}$ both converge to α.

"$\Leftarrow$" Suppose that there exist partitions P_n of $[a, b]$ for $n = 1, 2, \ldots \ni$

$$\text{(i)} \lim_{n\to\infty} \mathcal{L}(P_n, f) = \alpha,$$
$$\text{(ii)} \lim_{n\to\infty} \mathcal{U}(P_n, f) = \alpha.$$

Then

(i) $\Rightarrow \underline{\int_a^b} f \ge \alpha$, (what would happen if $\underline{\int_a^b} f = \alpha - \epsilon$, for $\epsilon > 0$?)

(ii) $\Rightarrow \overline{\int_a^b} f \le \alpha$.

Hence

$$\underline{\int_a^b} f \ge \alpha \ge \overline{\int_a^b} f.$$

But Lemma 5.3 $\Rightarrow$

$$\underline{\int_a^b} f \le \overline{\int_a^b} f.$$

Thus

$$\underline{\int_a^b} f = \overline{\int_a^b} f = \alpha.$$

That is, $\int_a^b f = \alpha$.

- Consider $f(x) = x, \quad 0 \le x \le 1$.
  Given $n \in \mathbb{N}$, let $P_n = \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right\}$. This is an example of a *uniform partition* into n subintervals. Denoting the points of P_n by $x_i = i/n$ for $i = 0, 1, 2, \ldots n$, we see that

$$\mathcal{L}(P_n, f) = \sum_{i=1}^{n} x_{i-1} \cdot \frac{1}{n} = \frac{1}{n^2}\sum_{i=1}^{n}(i-1) = \frac{1}{n^2}\left(\frac{n(n+1)}{2} - n\right) = \frac{n-1}{2n},$$
$$\mathcal{U}(P_n, f) = \sum_{i=1}^{n} x_i \cdot \frac{1}{n} = \frac{1}{n^2}\sum_{i=1}^{n} i = \frac{n+1}{2n}.$$

Note that

$$\lim_{n\to\infty} \mathcal{L}(P_n, f) = \frac{1}{2} = \lim_{n\to\infty} \mathcal{U}(P_n, f).$$

Theorem 5.1 $\Rightarrow$

$$\int_0^1 f = \frac{1}{2}.$$

- If

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \notin \mathbb{Q}, \end{cases}$$

then $\int_a^b f$ does not exist for any interval $[a, b]$ since $\mathcal{L}(P, f) = 0$ and $\mathcal{U}(P, f) = b - a$ for every partition P of $[a, b]$. Thus Theorem 5.1 $\Rightarrow \int_a^b f \nexists$. Note that

$$\underline{\int_a^b} f = 0 < b - a = \overline{\int_a^b} f.$$

Remark: We would arrive at the wrong conclusion if we attempted to compute the integral in the previous example by sampling the function at uniformly spaced points. For example, when $a = 0$ and $b = 1$,

$$\lim_{n\to\infty} \sum_{i=1}^{n} f\left(\frac{i}{n}\right)\frac{1}{n} = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{n} = \lim_{n\to\infty} 1 = 1.$$

Problem 5.1: Suppose f is integrable on $[a, b]$. Let $g(x) = f(x - c)$ for some $c \in \mathbb{R}$. By considering lower and upper sums, prove that g is integrable on $[a + c, b + c]$ and

$$\int_{a+c}^{b+c} g = \int_a^b f.$$

Problem 5.2: Suppose f is integrable on $[a, b]$. Let $g(x) = f(x/c)$ for some $c > 0$. By considering lower and upper sums, prove that g is integrable on $[ac, bc]$ and

$$\int_{ac}^{bc} g = c\int_a^b f.$$

5.B Cauchy Criterion

Theorem 5.2 (Cauchy Criterion for Integrability): *Suppose f is bounded on $[a, b]$. Then $\int_a^b f$ exists $\iff$ for each $\epsilon > 0$ there exists a partition P of $[a, b]$ such that*

$$\mathcal{U}(P, f) - \mathcal{L}(P, f) < \epsilon.$$

Proof:

"⇒" From Theorem 5.1, we know for any $\epsilon > 0$ we can find a partition P such that

$$\alpha - \frac{\epsilon}{2} < \mathcal{L}(P, f) \le \mathcal{U}(P, f) < \alpha + \frac{\epsilon}{2}.$$

Hence

$$\mathcal{U}(P, f) - \mathcal{L}(P, f) < \alpha + \frac{\epsilon}{2} - \left(\alpha - \frac{\epsilon}{2}\right) = \epsilon.$$

"⇐" Suppose that for each $\epsilon > 0$ there exists a partition P such that

$$\mathcal{U}(P, f) - \mathcal{L}(P, f) < \epsilon.$$

Then

$$\overline{\int_a^b} f \le \mathcal{U}(P, f) \quad \text{and} \quad -\underline{\int_a^b} f \le -\mathcal{L}(P, f)$$

$$\Rightarrow \overline{\int_a^b} f - \underline{\int_a^b} f \le \mathcal{U}(P, f) - \mathcal{L}(P, f) < \epsilon \quad \text{for each } \epsilon > 0.$$

Hence

$$\overline{\int_a^b} f - \underline{\int_a^b} f \le 0.$$

But Lemma 5.3 ⇒

$$\overline{\int_a^b} f - \underline{\int_a^b} f \ge 0.$$

Hence

$$\overline{\int_a^b} f = \underline{\int_a^b} f \Rightarrow \int_a^b f \;\exists.$$

Corollary 5.2.1 (Piecewise Integration): Suppose $a < c < b$. Then

$$\int_a^b f \;\exists \iff \int_a^c f \;\exists \text{ and } \int_c^b f \;\exists.$$

Furthermore, when either side holds,

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Proof [of existence]:

"$\Rightarrow$" Theorem 5.2: Given $\epsilon > 0$, there exists a partition Q of $[a, b]$ such that
$$\mathcal{U}(Q, f) - \mathcal{L}(Q, f) < \epsilon.$$
Now by Lemma 5.1, the refinement $P = Q \cup \{c\}$ of [a,b] satisfies
$$\mathcal{U}(P, f) - \mathcal{L}(P, f) \leq \mathcal{U}(Q, f) - \mathcal{L}(Q, f) < \epsilon.$$

Define $P' = P \cap [a, c]$ and $P'' = P \cap [c, b]$. Note that P' is a partition of $[a, c]$ and P'' is a partition of $[c, b]$. Observe that
$$\mathcal{U}(P', f) + \mathcal{U}(P'', f) = \mathcal{U}(P, f),$$
$$\mathcal{L}(P', f) + \mathcal{L}(P'', f) = \mathcal{L}(P, f)$$
$$\Rightarrow \underbrace{\mathcal{U}(P', f) - \mathcal{L}(P', f)}_{\geq 0} + \underbrace{\mathcal{U}(P'', f) - \mathcal{L}(P'', f)}_{\geq 0} = \mathcal{U}(P, f) - \mathcal{L}(P, f) < \epsilon$$
$$\Rightarrow \mathcal{U}(P', f) - \mathcal{L}(P', f) < \epsilon \quad \text{and} \quad \mathcal{U}(P'', f) - \mathcal{L}(P'', f) < \epsilon.$$
Hence Theorem 5.2 $\Rightarrow$
$$\int_a^c f \ \exists \quad \text{and} \quad \int_c^b f \ \exists.$$

"$\Leftarrow$" Suppose $\int_a^c f$ and $\int_c^b f$ both exist. Then given $\epsilon > 0$, there exists partitions P' of $[a, c]$ and P'' of $[c, b] \ni$
$$\mathcal{U}(P', f) - \mathcal{L}(P', f) < \frac{\epsilon}{2} \quad \text{and} \quad \mathcal{U}(P'', f) - \mathcal{L}(P'', f) < \frac{\epsilon}{2}.$$
Consider the partition $P = P' \cup P''$ of $[a, b]$. Then
$$\mathcal{U}(P, f) - \mathcal{L}(P, f) = \mathcal{U}(P', f) - \mathcal{L}(P', f) + \mathcal{U}(P'', f) - \mathcal{L}(P'', f) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence Theorem 5.2 $\Rightarrow \int_a^b f \ \exists$.

Problem 5.3: Complete the proof of Corollary 5.2.1 to show, given the existence of either side, that
$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Definition: If $a \leq b$, define $\int_b^a f = -\int_a^b f$.

Remark: This implies that $\int_a^a f = 0$.

Problem 5.4: Given any $c \in \mathbb{R}$, show that

$$\int_a^c f \,\exists \text{ and } \int_c^b f \,\exists \Rightarrow \int_a^b f \,\exists = \int_a^c f + \int_c^b f.$$

5.C Riemann Sums

Definition: Let f be a function on $[a, b]$ and $P = \{x_0, x_1, \ldots, x_n\}$ a partition of $[a, b]$. Then any sum of the form

$$\mathcal{S}(P, f) = \sum_{i=1}^{n} f(\bar{x}_i)(x_i - x_{i-1}),$$

where $\bar{x}_i$ are points chosen from $[x_{i-1}, x_i]$, is called a *Riemann sum* of f with respect to P.

Remark: Given any partition P and a continuous function f, the upper sum $\mathcal{U}(P, f)$ and lower sum $\mathcal{L}(P, f)$ are particular examples of Riemann sums $\mathcal{S}(P, f)$. To what choices of points $\bar{x}_i$ do these sums correspond?

Remark: Since $m_i \le f(\bar{x}_i) \le M_i$, every Riemann sum $\mathcal{S}(P, f)$ satisfies

$$\mathcal{L}(P, f) \le \mathcal{S}(P, f) \le \mathcal{U}(P, f).$$

Problem 5.5: Suppose that f is integrable on $[a, b]$. Show that there always exists a sequence of partitions P_n such that $\lim\limits_{n\to\infty} \mathcal{S}(P_n, f) = \int_a^b f$ (independent of how the points $\bar{x}_i$ are chosen).

Theorem 5.3 (Darboux Integrability Theorem): *$\int_a^b f$ exists and equals α $\iff$ for* **any** *sequence of partitions P_n having subinterval widths approaching zero as $n \to \infty$, all Riemann sums $\mathcal{S}(P_n, f)$ converge to α.*

Proof:

"$\Rightarrow$" We know since f is integrable that $|f(x)| \le M$ for some bound $M > 0$. Let $\alpha = \int_a^b f$. For any $\epsilon > 0$, Theorem 5.1 implies that there exists a partition Q of $[a, b]$ such that

$$\alpha - \frac{\epsilon}{2} < \mathcal{L}(Q, f) \le \mathcal{U}(Q, f) < \alpha + \frac{\epsilon}{2}.$$

Suppose Q partitions $[a, b]$ into q subintervals. Let P be any partition with subinterval widths all less than $\delta = \epsilon/(8qM)$. Consider the refinement $P \cup Q$ of P and Q. The contributions to the difference

$$\mathcal{U}(P, f) - \mathcal{U}(P \cup Q, f) \geq 0$$

come only from subintervals of $P \cup Q$ having an endpoint in the set Q. The $q - 1$ interior points of Q can be an endpoint of at most 2 such (adjacent) subintervals, while the points a and b can each belong to at most one subinterval. Hence there are certainly no more than $2(q - 1) + 2 = 2q$ such subintervals, each of which will contribute less than $2M\delta$ to the difference. Thus

$$\mathcal{U}(P, f) - \mathcal{U}(P \cup Q, f) < 4qM\delta = \frac{\epsilon}{2}.$$

We then deduce

$$\mathcal{U}(P, f) < \mathcal{U}(P \cup Q, f) + \frac{\epsilon}{2} \leq \mathcal{U}(Q, f) + \frac{\epsilon}{2} < \alpha + \frac{\epsilon}{2} + \frac{\epsilon}{2} = \alpha + \epsilon.$$

Similarly, we find

$$\mathcal{L}(P, f) > \alpha - \epsilon,$$

so that any Riemann sum $\mathcal{S}(P, f)$ of f on the partition P must satisfy

$$\alpha - \epsilon < \mathcal{L}(P, f) \leq \mathcal{S}(P, f) \leq \mathcal{U}(P, f) < \alpha + \epsilon.$$

"$\Leftarrow$" Given $n \in \mathbb{N}$, choose a partition P_n sufficiently fine so that for all possible choices of the points $\bar{x}_i$, we have $|\mathcal{S}(P_n, f) - \alpha| < 1/(2n)$. In particular, we may choose $\bar{x}_i \in [x_{i-1}, x_i]$ such that

$$f(\bar{x}_i) > M_i - \frac{1}{2n(b-a)}. \qquad \text{(Why?)}$$

On denoting the number of points in P_n by $|P_n|$ we find

$$\begin{aligned}
\mathcal{U}(P_n, f) &= \sum_{i=1}^{|P_n|} M_i(x_i - x_{i-1}) \\
&< \sum_{i=1}^{|P_n|} \left[f(\bar{x}_i) + \frac{1}{2n(b-a)} \right](x_i - x_{i-1}) = \mathcal{S}(P_n, f) + \frac{1}{2n(b-a)} \sum_{i=1}^{|P_n|}(x_i - x_{i-1}) \\
&< \alpha + \frac{1}{2n} + \frac{1}{2n(b-a)}(b-a) = \alpha + \frac{1}{n}.
\end{aligned}$$

Using another similar choice for the points $\bar{x}_i$, we may show

$$\mathcal{L}(P_n, f) > \alpha - \frac{1}{n}.$$

Thus, we have found a sequence of partitions P_n for which

$$\lim_{n\to\infty} \mathcal{L}(P_n, f) = \alpha = \lim_{n\to\infty} \mathcal{U}(P_n, f).$$

From Theorem 5.1 we then conclude that $\int_a^b f$ exists and equals α.

5.D Properties of Integrals

Theorem 5.4 (Linearity of Integral Operator): *Suppose* $\int_a^b f$ *and* $\int_a^b g$ *exist. Then*

(i) $\int_a^b (f+g) \exists = \int_a^b f + \int_a^b g$,

(ii) $\int_a^b (cf) \exists = c\int_a^b f$ *for any constant* $c \in \mathbb{R}$.

Proof:

(i) Let $\alpha = \int_a^b f$, $\beta = \int_a^b g$. Theorem 5.1: $\exists$ partitions P_n and Q_n of $[a,b]$ for each $n \in \mathbb{N} \ni$

$$\alpha - \frac{1}{n} < \mathcal{L}(P_n, f) \le \mathcal{U}(P_n, f) < \alpha + \frac{1}{n},$$
$$\beta - \frac{1}{n} < \mathcal{L}(Q_n, g) \le \mathcal{U}(Q_n, g) < \beta + \frac{1}{n}.$$

Let $R_n = P_n \cup Q_n$. Note that the partition R_n is a refinement of both P_n and Q_n. Now on any subinterval $I = [x_{k-1}, x_k]$ of R_n,

$$f(x) + g(x) \le \sup\{f(x) : x \in I\} + \sup\{g(x) : x \in I\}.$$

Since the right-hand side is an upper bound for $f(x) + g(x)$ for all $x \in I$,

$$\sup\{f(x) + g(x) : x \in I\} \le \sup\{f(x) : x \in I\} + \sup\{g(x) : x \in I\}.$$

Hence

$$\mathcal{U}(R_n, f+g) \le \mathcal{U}(R_n, f) + \mathcal{U}(R_n, g).$$

Similarly,

$$\mathcal{L}(R_n, f+g) \ge \mathcal{L}(R_n, f) + \mathcal{L}(R_n, g).$$

Thus

$$\begin{aligned}
\alpha + \beta - \frac{2}{n} &< \mathcal{L}(P_n, f) + \mathcal{L}(Q_n, g) \le \mathcal{L}(R_n, f) + \mathcal{L}(R_n, g) \\
&\le \mathcal{L}(R_n, f+g) \le \mathcal{U}(R_n, f+g) \le \mathcal{U}(R_n, f) + \mathcal{U}(R_n, g) \\
&\le \mathcal{U}(P_n, f) + \mathcal{U}(Q_n, g) < \alpha + \beta + \frac{2}{n}.
\end{aligned}$$

Using the Squeeze Principle, we conclude

$$\lim_{n\to\infty} \mathcal{L}(R_n, f+g) = \lim_{n\to\infty} \mathcal{U}(R_n, f+g) = \alpha + \beta.$$

Thus, by Theorem 5.1,

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g.$$

(ii) Exercise.

Theorem 5.5 (Integral Bounds): *Suppose for $a < b$ that*

(i) $\int_a^b f \,\exists$,

(ii) $m \le f(x) \le M$ *for* $x \in [a, b]$.

Then

$$m(b-a) \le \int_a^b f \le M(b-a).$$

Proof: Consider the partition $P = \{a, b\}$ of $[a, b]$. Then

$$m(b-a) \le \mathcal{L}(P, f) \le \underline{\int_a^b} f = \int_a^b f = \overline{\int_a^b} f \le \mathcal{U}(P, f) \le M(b-a).$$

Corollary 5.5.1 (Preservation of Non-Negativity): If $f(x) \ge 0$ for all $x \in [a, b]$ and $\int_a^b f$ exists then $\int_a^b f \ge 0$.

Proof: Set $m = 0$ in Theorem 5.5.

Problem 5.6: Consider the function

$$f(x) = \begin{cases} \frac{1}{n} & \text{if } x = \frac{m}{n} \text{ for relatively prime integers } m \neq 0 \text{ and } n > 0, \\ 1 & \text{if } x = 0, \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

Show that $\int_0^1 f \,\exists = 0$, even though $f \not\equiv 0$. Hence, even within the class of non-negative functions, the statement $\int_0^1 f = 0 \Rightarrow f = 0$ does not hold!

Corollary 5.5.2 (Absolute Integral Bounds): If $|f(x)| \le M$ for all $x \in [a, b]$ and $\int_a^b f$ exists then $\left|\int_a^b f\right| \le M\,|b-a|$.

Proof: Set $m = -M$ in Theorem 5.5.

Corollary 5.5.3 (Continuity of Integrals): Suppose $\int_a^b f$ exists. Then the function $F(x) = \int_a^x f$ is continuous on $[a, b]$.

Proof: If $f(x) = 0$ for all $x \in [a, b]$ then $F(x) = 0$, which is continuous. Otherwise, let $M = \sup\{|f(x)| : a \le x \le b\} > 0$ and consider $u \in [a, b]$. Given $\epsilon > 0$, for any $x \in [a, b]$ with $|x - u| < \delta$,

$$|F(x) - F(u)| = \left|\int_a^x f - \int_a^u f\right| = \left|\int_a^x f + \int_u^a f\right| = \left|\int_u^x f\right| \le M\,|x-u| < M\delta = \epsilon,$$

by Theorem 5.5, provided we choose $\delta = \epsilon/M$.

Theorem 5.6 (Integrability of Continuous Functions): *If f is continuous on $[a, b]$ then $\int_a^b f$ exists.*

Proof: Given $\epsilon > 0$, we claim that $\exists\, \delta > 0 \ni$

$$|f(x) - f(y)| < \frac{\epsilon}{b-a} \quad \forall\, x, y \in [a, b] \text{ with } |x - y| < \delta \quad \text{(we say: } f \text{ is \textit{uniformly continuous}).}$$

Suppose not. Then $\exists\, \epsilon_0 > 0$ and sequences $\{x_n\}_{n=1}^{\infty}$, $\{y_n\}_{n=1}^{\infty}$ in $[a, b] \ni$

$$|x_n - y_n| < \frac{1}{n}, \quad \text{but} \quad |f(x_n) - f(y_n)| \geq \epsilon_0.$$

Note that $\{x_n\}_{n=1}^{\infty}$ has a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ that converges to $c \in [a, b]$. Then, since

$$-\frac{1}{n_k} \leq x_{n_k} - y_{n_k} \leq \frac{1}{n_k}$$

and $\lim\limits_{k\to\infty} \dfrac{1}{n_k} = 0$, we can apply the Squeeze Principle to deduce that $\lim\limits_{k\to\infty} y_{n_k} = \lim\limits_{k\to\infty} x_{n_k} = c$. But f is continuous, so

$$\lim_{k\to\infty} [(f(x_{n_k}) - f(y_{n_k})] = f\Big(\lim_{k\to\infty} x_{n_k}\Big) - f\Big(\lim_{k\to\infty} y_{n_k}\Big) = f(c) - f(c) = 0.$$

This contradicts $\lim\limits_{k\to\infty} \Big| f(x_{n_k}) - f(y_{n_k}) \Big| \geq \epsilon_0 > 0$, so in fact our claim must hold.

Now choose a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ such that $x_i - x_{i-1} < \delta$ for $i = 1, \ldots, n$. Since f is continuous on $[x_{i-1}, x_i]$, Theorem 3.4 $\Rightarrow$ there exists numbers p_i and $q_i \in [x_{i-1}, x_i]$ such that

$$f(p_i) = \inf\{f(x) : x_{i-1} \leq x \leq x_i\},$$

$$f(q_i) = \sup\{f(x) : x_{i-1} \leq x \leq x_i\}.$$

Noting that $|p_i - q_i| < \delta$, our above claim implies that

$$\begin{aligned}\mathcal{U}(P, f) - \mathcal{L}(P, f) &= \sum_{i=1}^{n} [f(q_i) - f(p_i)](x_i - x_{i-1}) \\ &< \sum_{i=1}^{n} \frac{\epsilon}{b-a}(x_i - x_{i-1}) = \frac{\epsilon}{b-a} \sum_{i=1}^{n} (x_i - x_{i-1}) = \frac{\epsilon}{b-a}(x_n - x_0) = \epsilon.\end{aligned}$$

Therefore, by Theorem 5.2, $\int_a^b f$ exists.

Theorem 5.7 (Integrability of Monotonic Functions): *If f is monotonic on $[a,b]$ then $\int_a^b f$ exists.*

Proof: Suppose f is increasing. Then for any partition $P = \{x_0, x_1, \ldots, x_n\}$ we know that

$$\inf\{f(x) : x_{i-1} \le x \le x_i\} = f(x_{i-1}),$$

$$\sup\{f(x) : x_{i-1} \le x \le x_i\} = f(x_i).$$

If $f(b) = f(a)$ then f is constant $\Rightarrow \int_a^b f$ exists. Otherwise, given $\epsilon > 0$, choose a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a,b]$ fine enough such that for $i = 1, \ldots, n$,

$$x_i - x_{i-1} < \delta = \frac{\epsilon}{f(b) - f(a)}.$$

Then

$$\begin{aligned} \mathcal{U}(P,f) - \mathcal{L}(P,f) &= \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})](x_i - x_{i-1}) \\ &< \delta \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] = \delta[f(x_n) - f(x_0)] = \delta[f(b) - f(a)] = \epsilon. \end{aligned}$$

Hence by Theorem 5.2, $\int_a^b f \ \exists$.

- For the function
$$f(x) = \begin{cases} \frac{1}{n} & \text{if } x \in \left(\frac{1}{n+1}, \frac{1}{n}\right] \text{ for } n \in \mathbb{N}, \\ 0 & \text{if } x = 0 \end{cases}$$
on $[0,1]$, Theorem 5.7 $\Rightarrow \int_0^1 f \ \exists$.

5.E Fundamental Theorem of Calculus

Definition: A differentiable function F is called an *antiderivative* of f at an interior point x of its domain if $F'(x) = f(x)$.

Remark: If $F(x)$ is an antiderivative of f, then so is $F(x) + C$ for any constant C.

Lemma 5.4 (Families of Antiderivatives): *Let $F_0(x)$ be an antiderivative of f on an interval I. Then F is an antiderivative of f on I $\iff$ $F(x) = F_0(x) + C$ for some constant C.*

Proof:

"$\Leftarrow$" Let $F(x) = F_0(x) + C$. Then $F'(x) = F_0'(x) = f(x)$; that is, F is an antiderivative of f on I.

"$\Rightarrow$" Since

$$\frac{d}{dx}[F(x) - F_0(x)] = F'(x) - F_0'(x) = f(x) - f(x) = 0,$$

we see by Corollary 4.4.3 that $F(x) - F_0(x)$ is constant on I.

Theorem 5.8 (Antiderivatives at Points of Continuity): *Suppose*

(i) $\int_a^b f$ *exists;*

(ii) f *is continuous at* $c \in (a, b)$.

Then f *has the antiderivative* $F(x) = \int_a^x f$ *at* $x = c$.

Proof: Given $\epsilon > 0$, we know from the continuity of f at c that there exists a $\delta > 0$ such that

$$|x - c| < \delta \Rightarrow |f(x) - f(c)| < \epsilon.$$

Consider $F(x) = \int_a^x f$ for $x \in [a, b]$. Then

$$F(c + h) - F(c) - f(c)h = \int_a^{c+h} f - \int_a^c f - \int_c^{c+h} f(c) = \int_c^{c+h} [f - f(c)].$$

For $|h| < \delta$, Corollary 5.5.2 thus implies

$$0 \le |F(c + h) - F(c) - f(c)h| \le \epsilon\, |c + h - c| = \epsilon\, |h|\,.$$

That is, given $\epsilon > 0$, there exists a $\delta > 0$ such that

$$0 < |h| < \delta \Rightarrow \left| \frac{F(c + h) - F(c)}{h} - f(c) \right| \le \epsilon.$$

But this is just the statement that the limit

$$F'(c) = \lim_{h \to 0} \frac{F(c + h) - F(c)}{h}$$

exists and equals $f(c)$.

Remark: In particular, Theorem 5.8 says that, at any point $x \in (a, b)$ where an integrable function f is continuous,

$$\frac{d}{dx}\int_a^x f = f(x).$$

Thus we see that differentiation and integration are in a sense opposite processes. The actual situation is slightly complicated by the fact that antiderivatives are not unique, as we saw in Lemma 5.4. However, note that the arbitrary constant C in Lemma 5.4 disappears upon differentiation of the antiderivative.

Corollary 5.8.1 (Antiderivative of Continuous Functions): If f is continuous on $[a, b]$ then f has an antiderivative on $[a, b]$.

Proof: The antiderivative of f on $[a, b]$ is just the antiderivative $\int_a^x \bar{f}$ of the *continuous extension* $\bar{f}$ of f onto all of $\mathbb{R}$:

$$\bar{f}(x) = \begin{cases} f(a) & \text{if } x < a, \\ f(x) & \text{if } a \le x \le b, \\ f(b) & \text{if } x > b. \end{cases}$$

Theorem 5.9 (Fundamental Theorem of Calculus [FTC]): *Let f be integrable and have an antiderivative F on $[a, b]$. Then*

$$\int_a^b f = F(b) - F(a).$$

Proof: Let $P = \{x_0, x_1, \ldots, x_n\}$ be any partition of $[a, b]$. Since F is differentiable on $[a, b]$, the MVT tells us that for each $i = 1, \ldots n$ there exists a $c_i \in (x_{i-1}, x_i)$ such that

$$F(x_i) - F(x_{i-1}) = F'(c_i)(x_i - x_{i-1}) = f(c_i)(x_i - x_{i-1}).$$

Consider the Riemann sum $\mathcal{S}(P, f) \doteq \displaystyle\sum_{i=1}^{n} f(c_i)(x_i - x_{i-1})$. Then

$$\mathcal{S}(P, f) = \sum_{i=1}^{n}[F(x_i) - F(x_{i-1})] = F(x_n) - F(x_0) = F(b) - F(a),$$

so that

$$\mathcal{L}(P, f) \le F(b) - F(a) \le \mathcal{U}(P, f).$$

That is, $F(b) - F(a)$ is a lower bound for all upper sums and an upper bound for all lower sums. Hence

$$\underline{\int_a^b} f \le F(b) - F(a) \le \overline{\int_a^b} f.$$

But since $\int_a^b f = \underline{\int_a^b} f = \overline{\int_a^b} f$, it follows that $\int_a^b f = F(b) - F(a)$.

Remark: It is possible for a function to be integrable, but have no antiderivative. But by Theorem 5.8, we know that such a function cannot be continuous. Consider the function

$$f(x) = \begin{cases} -1 & \text{if } -1 \leq x < 0, \\ \\ 1 & \text{if } 0 \leq x \leq 1. \end{cases}$$

Being monotonic, f is certainly integrable (Theorem 5.7). However, the FTC implies that f cannot be the derivative of another function F. For if $f = F'$, then $\int_0^x f = F(x) - F(0)$, so that

$$\begin{aligned} F(x) = F(0) + \int_0^x f = F(0) + &\begin{cases} \int_0^x (-1) & \text{if } -1 \leq x < 0, \\ \\ \int_0^x 1 & \text{if } 0 \leq x \leq 1, \end{cases} \\ = F(0) + &\begin{cases} -1(x-0) & \text{if } -1 \leq x < 0, \\ \\ 1(x-0) & \text{if } 0 \leq x \leq 1, \end{cases} \\ = F(0) + &\left| x \right|, \end{aligned}$$

which we know is not differentiable at $x = 0$, regardless of what $F(0)$ is.

Problem 5.7: It is also possible for an integrable function f to be discontinuous at a point but still have an antiderivative F. Consider $f = F'$, where

$$F(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Prove that f is discontinuous at 0 but that f is still integrable on any finite interval.

Corollary 5.9.1 (FTC for Continuous Functions): Let f be continuous on $[a, b]$ and let F be any antiderivative of f on $[a, b]$. Then

$$\int_a^b f = F(b) - F(a).$$

Proof: This follows directly from Theorem 5.6 and the FTC.

Remark: The FTC says that a *definite integral* $\int_a^b f$ is equal to the value of any antiderivative F of f at b minus the value of the **same** function F at a. That is, $\int_a^b f = [F(x)]_a^b$, where the notation $[F(x)]_a^b$ or $F(x)|_a^b$ is shorthand for the difference $F(b) - F(a)$.

- Let $f(x) = x$. Then

$$\int_0^1 f = \left[\frac{x^2}{2} + c \right]_0^1 = \frac{1}{2} + c - (0 + c) = \frac{1}{2}.$$

Remark: We need a convenient notation for an antiderivative.

Definition: If an integrable function f has antiderivative F, we write $F = \int f$ and say F is the *indefinite integral* of f.

$$\int f = F \text{ means } f = F'.$$

Remark: When we write $\int f$ we understand that f is a function of some variable. Let's call it x. We want to partition a portion of the x axis into $\{x_0, x_1, \ldots, x_n\}$ and compute upper and lower sums based on function values and the interval widths $x_i - x_{i-1}$. Similarly, when we write f', it is clear that we mean the derivative of f with respect to its argument, whatever that may be. However, if we want to differentiate the function $y = f(u)$, where $u = x^2$, it is important to know whether we are differentiating with respect to u or with respect to x. Likewise, suppose we wish to calculate the integral of f. It is equally important to know whether we are calculating the integral with respect to u or with respect to x, because the area under the graph of $y = f(u)$ with respect to u will in general differ from the area under the graph of $y = f(x^2)$ with respect to x. Since we can differentiate with respect to different variables, it is only reasonable, in view of the FTC, that we should be able to integrate with respect to different variables as well. It will often be helpful to indicate explicitly with respect to which variable we are integrating, that is, which variable do we use to construct the differences $x_i - x_{i-1}$ in the lower and upper sums.

Definition: We can specify the integration variable by writing $\int_0^1 f(x)\,dx$ instead of just $\int_0^1 f$. The notation $f(x)\,dx$ reminds us that the lower sums and upper sums consists of function values (strictly speaking, infima and suprema) multiplied by interval widths, $x_i - x_{i-1}$.

Remark: The same notation is also used for indefinite integrals. For example,

$$\int x\,dx = \frac{x^2}{2} + C \quad \text{ means } \quad x = \frac{d}{dx}\left(\frac{x^2}{2} + C\right).$$

Remark: Remember that the definite integral $\int_a^b f(x)\,dx$ is a number, whereas the indefinite integral $\int f(x)\,dx$ represents a *family* of functions that differ from each other by a constant.

- Since
$$\frac{d}{dx}\frac{x^{p+1}}{p+1} = x^p,$$
we know that
$$\int_a^b x^p\,dx = \left.\frac{x^{p+1}}{p+1}\right|_a^b = \frac{b^{p+1}-a^{p+1}}{p+1} \qquad \text{if } p \neq -1.$$

- Also,
$$\int_0^\pi \sin x\,dx = [-\cos x]_0^\pi = -[\cos x]_0^\pi = -[-1-1] = 2.$$

- But
$$\int_0^{2\pi} \sin x\,dx = [-\cos x]_0^{2\pi} = [-1-(-1)] = 0.$$

- The function
$$F(x) = \int_0^x \frac{1}{\cos t}\,dt$$
is differentiable for $x \in [0, \frac{\pi}{2})$.

  We don't yet know F, but we do know its derivative. Thm 5.8 $\Rightarrow$
$$F'(x) = \frac{1}{\cos x}.$$

Furthermore, suppose
$$G(x) = \int_0^{x^2} \frac{1}{\cos t}\,dt = F(x^2)$$
for $x \in \left[0, \sqrt{\frac{\pi}{2}}\right)$. Then
$$G'(x) = F'(x^2)\,2x = \frac{2x}{\cos(x^2)} \quad \text{by the Chain Rule.}$$

- Consider the inverse trigonometric function $y = \sin^{-1} x$ for $x \in [-1, 1]$. Recall that
$$\frac{d}{dx}\sin^{-1} x = \frac{1}{\sqrt{1-x^2}} \text{ for } (-1,1)$$
and
$$\frac{d}{dx}\tan^{-1} x = \frac{1}{1+x^2} \text{ for } x \in (-\infty, \infty).$$
These results yield two important antiderivatives:
$$\int \frac{1}{\sqrt{1-x^2}}\,dx = \sin^{-1} x + C$$
and
$$\int \frac{1}{1+x^2}\,dx = \tan^{-1} x + C.$$

Problem 5.8: Suppose f is a continuous function and g and b are differentiable functions on $[a, b]$. Prove that

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(t)\,dt = f(b(x))b'(x) - f(a(x))a'(x).$$

Let F be an antiderivative for f. Theorem FTC states that

$$\int_{a(x)}^{b(x)} f(t)\,dt = F(b(x)) - F(a(x)).$$

Hence, using the Chain Rule,

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(t)\,dt = F'(b(x))b'(x) - F'(a(x))a'(x) = f(b(x))b'(x) - f(a(x))a'(x).$$

5.F Average Value of a Function

Definition: We define the *average value* of an integrable function f on $[a, b]$ to be

$$\frac{1}{b-a}\int_a^b f.$$

Remark: If f is continuous on $[a, b]$, the following theorem states that f takes on its average value somewhere in the interval $[a, b]$.

Theorem 5.10 (Mean Value Theorem for Integrals [MVTI]): *Suppose f is continuous on $[a, b]$. Then*

$$\int_a^b f = f(c)(b-a)$$

for some number $c \in [a, b]$.

Proof: We know from Theorem 3.4 that there exists points $p, q \in [a, b]$ such that

$$f(p) \le f(x) \le f(q) \quad \forall x \in [a, b].$$

Theorem 5.5 then implies that

$$f(p)(b-a) \le \int_a^b f \le f(q)(b-a),$$

so that

$$f(p) \le \frac{1}{b-a}\int_a^b f \le f(q).$$

Finally, the IVT implies that $\exists c \in [p, q] \subset [a, b]$, such that

$$f(c) = \frac{1}{b-a}\int_a^b f.$$

Problem 5.9: Find a proof for a slightly stronger version of Theorem 5.10: if f is continuous on $[a, b]$, then

$$\int_a^b f = f(c)(b-a)$$

for some number $c \in (a, b)$.

- Since $x^3/3$ is an antiderivative of $f(x) = x^2$, it follows that $\int_0^1 f = 1/3$. The MVTI implies that there exists a number $c \in [0, 1]$ such that

$$\frac{1}{3} = \int_0^1 f = f(c)(1-0) = f(c).$$

In this case we can even determine the value of $c \in [0, 1]$:

$$\frac{1}{3} = c^2 \Rightarrow c = \frac{1}{\sqrt{3}}.$$

Chapter 6

Logarithmic and Exponential Functions

6.A Exponentials and Logarithms

Suppose we would like to fill in Table 6.1 with a continuous function that satisfies

$$10^x 10^y = 10^{x+y}.$$

For example, we might like to find a general method to take the n^{th} root of 10. Also, most computers express numbers in terms of powers of 2, so we might need a method to convert numbers from their decimal to binary representations and back again.

x	10^x
-2	$\frac{1}{100}$
-1	$\frac{1}{10}$
0	1
$1/3$	$\sqrt[3]{10}$
$1/2$	$\sqrt{10}$
1	10
$4/3$	$10\sqrt[3]{10}$
$3/2$	$10\sqrt{10}$
2	100
3	1000

Table 6.1: Values of 10^x for rational x

Q. What should it mean to take irrational powers of 10, such as $10^{\sqrt{2}}$?

General Problem: Find a continuous function $f(x)$ that satisfies $f(0) = 1$, $f(1) > 0$ and $f(x)f(y) = f(x+y)$ for all real numbers x and y. In particular, this implies that

$$f(n) = b^n = \underbrace{b \times b \times \ldots \times b}_{n \text{ times}} \qquad \text{for } n \in \mathbb{N},$$

for some positive number b called the *base*.

Calculus provides us with an answer: if $f(x)$ is differentiable then $f(x)$ will be continuous, as desired, and

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{f(x)f(h) - f(x)}{h} \\ &= f(x) \lim_{h \to 0} \frac{f(h) - 1}{h}. \end{aligned}$$

Since $f(0) = 1$ we see that

$$f'(0) = \lim_{h \to 0} \frac{f(h) - 1}{h},$$

so we seek a function that satisfies

$$f'(x) = f(x)f'(0), \qquad f(0) = 1.$$

Q. Does such a function exist?

A. Yes, there even exists one with $f'(0) = 1$, so that this special function is its own derivative!

Q. How do we find f such that $f'(x) = f(x)$? From the FTC, we see that

$$f(x) - 1 = f(x) - f(0) = \int_0^x f'(t)\,dt = \underbrace{\int_0^x f(t)\,dt.}_{\substack{\text{we can't integrate this until} \\ \text{we know } f \text{ (vicious circle)}}}$$

A. Inverse functions come to our rescue. Letting $y = f(x)$, we see that

$$\frac{dy}{dx} = y \Rightarrow \frac{dx}{dy} = \frac{1}{y}.$$

For any positive real numbers a and y, we know from Theorem 5.8 that

$$x = \int_a^y \frac{1}{t}\,dt \Rightarrow \frac{dx}{dy} = \frac{1}{y}.$$

When $x = 0$, we want $y = 1$. Enforce this by choosing $a = 1$. This yields x as a function of $y > 0$:

$$x = f^{-1}(y) = \int_1^y \frac{1}{t}\,dt.$$

We define this function, which measures the area under the graph of the function $\frac{1}{t}$ vs. t between 1 and y, as the *natural logarithm*,

$$\log(y) = \int_1^y \frac{1}{t}\,dt.$$

Note that $\log(1) = \int_1^1 \frac{1}{t}\,dt = 0$.

Remark: For an integrable function g and any constant $c > 0$, we showed in an earlier exercise that

$$\int_{ac}^{bc} g(t)\,dt = c\int_a^b g(ct)\,dt.$$

By letting $g(t) = 1/t$, $a = 1$, and $b = c = y$, we can use this result to highlight an important property of the logarithm:

$$\log(y^2) - \log(y) = \int_y^{y^2} \frac{1}{t}\,dt = y\int_1^y \frac{1}{yt}\,dt = \int_1^y \frac{1}{t}\,dt = \log(y).$$

Thus $\log(y^2) = 2\log(y)$ for any number $y > 0$. By induction, we find for $n \in \mathbb{N}$ that

$$\log(y^n) = n\log(y).$$

Remark: Since $y > z$ implies $\log y - \log z = \int_z^y \frac{1}{t}\,dt \geq \frac{1}{y}(y - z) > 0$, we see that $\log y$ is a strictly increasing function of y.

Problem 6.1: For $n \in \mathbb{N}$ use the facts that $\log 2^n = n\log 2 > 0$ and $\log\left(\frac{1}{2}\right)^n = n\log\frac{1}{2} < 0$, in view of the strictly increasing nature of $\log x$, to prove that $\lim\limits_{y\to\infty}\log y = \infty$ and $\lim\limits_{y\to 0+}\log y = -\infty$.

Remark: Since Corollary 5.5.3 guarantees that $\log y$ is continuous, the IVT then implies that the range of $\log y$ is all of $\mathbb{R}$, Moreover, $\log y > 0$ for $y > 1$ and $\log y < 0$ for $y < 1$, as shown in Figure 6.1.

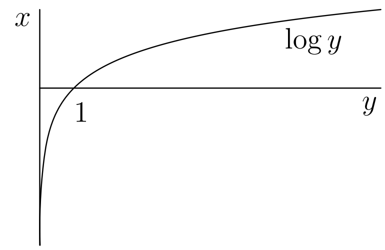


Figure 6.1: The natural logarithm

Q: The function $f^{-1}(y) = \log y$ is thus invertible on $(0, \infty)$. Its inverse function $f(x)$, defined on all of $\mathbb{R}$, is a solution to our problem for a particular value of $b = f(1)$. But what is this special value of b? In other words, for what value of y does $\log y = 1$?

A: By construction,

$$\frac{1}{y} = \frac{d}{dy}\log(y) = \lim_{h\to 0}\frac{\log(y+h)-\log(y)}{h}.$$

At $y = 1$, we find

$$\begin{aligned}
1 &= \lim_{h\to 0}\frac{\log(1+h)-\log 1}{h} = \lim_{h\to 0}\frac{1}{h}\log(1+h) \\
&= \lim_{n\to\infty} n\log\left(1+\frac{1}{n}\right) = \lim_{n\to\infty}\log\left(1+\frac{1}{n}\right)^n \\
&= \log\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \qquad \text{(continuity)} \\
&= \log e,
\end{aligned}$$

where we recall $e \approx 2.718281828459\ldots$ is the limit of the convergent sequence $\{\left(1+\frac{1}{n}\right)^n\}_{n=1}^{\infty}$.

Remark: We have thus found a unique differentiable function f such that

$$\begin{aligned}
f(0) &= 1, \\
f(1) &= e, \\
f(n) &= e^n = \underbrace{e\times e\times\ldots\times e}_{n\text{ times}}
\end{aligned}$$

and

$$f(x)f(y) = f(x+y).$$

Remark: From these properties we see that

$$f(x)f(-x) = f(0) = 1 \Rightarrow f(-x) = 1/f(x) > 0.$$

Note that $f(x) > 0$ for all real x.

Definition: For any real x, we define $\exp(x) \doteq f(x)$ to be the inverse of the function $x = \log y$ (which measures the area under the graph of the function $1/t$ from $t = 1$ to $t = y$). Note for $n \in \mathbb{Z}$ and $m \in \mathbb{N}$ that $\exp(n/m) = \sqrt[m]{e^n}$. We thus see that the function $\exp(x)$ extends the notion of *exponentiation* from $\mathbb{Q}$ to $\mathbb{R}$. To emphasize this connection, it is often convenient to denote the *exponential* function $\exp(x)$ as e^x. The graph of e^x is shown in Figure 6.2.

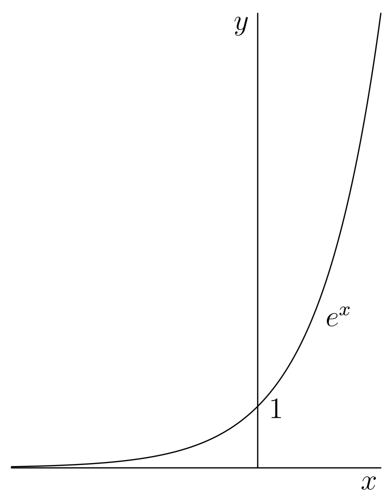


Figure 6.2: The natural exponential function

Remark: There are other such exponential function (e.g. 10^x or 2^x) corresponding to different choices of the base b, but we have found a special one, e^x, such that its derivative is the same as its value everywhere.

Problem 6.2: Given that an invertible function f obeys $f(x+y) = f(x)f(y)$ for all x and y in its domain, prove that f^{-1} satisfies $f^{-1}(XY) = f^{-1}(X) + f^{-1}(Y)$ for all X and Y in the range of f. Deduce that $\log(xy) = \log x + \log y$ for all positive numbers x and y.

Remark: We now use Taylor's Theorem about the point $a = 0$ to show that the number e is irrational. Let $n \in \mathbb{N}$ and set $f(x) = e^x$. On using the fact that $f^{(k)}(x) = e^x$ for all $k \in \mathbb{N}$, we find that

$$e = f(1) = \sum_{k=0}^{n-1} \frac{1}{k!} e^0 + \frac{1}{n!} e^{c_n}$$

for some $c_n \in (0, 1)$. That is,

$$e - \sum_{k=0}^{n-1} \frac{1}{k!} = \frac{1}{n!} e^{c_n}.$$

We know that e^x is strictly increasing on $\mathbb{R}$ since its derivative e^x is positive on $\mathbb{R}$. Hence

$$0 < c_n < 1 \Rightarrow 1 < e^{c_n} < e^1 < 3.$$

Thus

$$\frac{1}{n!} < e - \sum_{k=0}^{n-1} \frac{1}{k!} < \frac{3}{n!}.$$

On multiplying this result by $(n-1)!$ we find

$$0 < \frac{1}{n} < (n-1)!\, e - \sum_{k=0}^{n-1} \frac{(n-1)!}{k!} < \frac{3}{n} \leq 1 \qquad \text{(for } n \geq 3\text{)}.$$

Note that each term of the summation here is an integer. If e were rational, this would imply for n sufficiently large that there exists an integer in the interval $(0, 1)$, which is absurd! Hence e must be irrational. It can also be shown that e is *transcendental*; that is, e does not satisfy any algebraic equation with integer coefficients [Spivak 1994, Ch. 21].

Base Conversion: Sometimes, we need to work in another base $b > 0$, e.g. $b = 2$ or $b = 10$.

Define $b^x \doteq e^{x \log b}$ (this agrees with our notation for integral powers and roots of b since $e^{n \log b} = e^{\log b^n} = b^n$). Note that if $f(x) = e^{x \log b}$ then $f'(x) = f(x) \log b$. The factor $\log b$ is just the expression $\lim\limits_{h \to 0} \frac{f(h)-1}{h} = \lim\limits_{h \to 0} \frac{e^{h \log b}-1}{h}$ that we encountered in our construction of exponential functions.

Remark: For integral (and some rational) values of x, the definition of b^x can be extended to negative values of b.

Properties: If x, $y \in \mathbb{R}$ and $b > 0$, then

1. $b^{x+y} = b^x b^y$.
   Proof: $b^{x+y} = e^{(x+y)\log b} = e^{x\log b + y\log b} = e^{x\log b}e^{y\log b} = b^x b^y$.
2. $b^{x-y} = \frac{b^x}{b^y}$.
   Proof: $b^{x-y} = e^{(x-y)\log b} = e^{x\log b - y\log b} = \frac{e^{x\log b}}{e^{y\log b}} = \frac{b^x}{b^y}$.
3. $(b^x)^y = b^{xy}$.
   Proof: $\left(e^{x\log b}\right)^y = e^{y\log\left(e^{x\log b}\right)} = e^{yx\log b} = b^{xy}$.
4. $(ab)^x = a^x b^x$.
   Proof: $e^{x\log(ab)} = e^{x\log a + x\log b} = e^{x\log a}e^{x\log b} = a^x b^x$.

Q. What is the inverse function of $y = b^x = e^{(x\log b)}$?

A. We solve for x:

$$\log y = x\log b$$
$$\Rightarrow x = \frac{\log y}{\log b} \doteq \log_b y.$$

- Note that $\log_2 e = \frac{\log e}{\log 2} = \frac{1}{\log 2}$.

- However, $\log_e 2 = \log 2$.

Properties: If x, y, and b are positive numbers,

1. $\log_b(xy) = \log_b x + \log_b y$.
   Proof: Exercise.
2. $\log_b\left(\frac{x}{y}\right) = \log_b x - \log_b y$.
   Proof: Exercise.
3. $\log_b(x^r) = r\log_b x$.
   Proof: $\frac{\log x^r}{\log b} = \frac{\log e^{r\log x}}{\log b} = \frac{r\log x}{\log b}$.

Problem 6.3: Use Taylor's Theorem about $a = 0$ and the Sequence Limit Ratio Test to establish that

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Remark: Recall that

$$\frac{1}{y} = \frac{d}{dy}\log(y) = \lim_{h\to 0}\frac{\log(y+h)-\log(y)}{h}.$$

In particular, at $y = \frac{1}{x}$ we find

$$x = \lim_{h\to 0}\frac{\log\left(\frac{1}{x}+h\right)+\log x}{h} = \lim_{h\to 0}\log(1+xh)^{\frac{1}{h}} = \log\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n.$$

We thus obtain another expression for e^x:

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n.$$

- Derivative of b^x:

$$\begin{aligned} & f(x) = b^x = e^{x\log b} \\ \Rightarrow\ & f'(x) = e^{x\log b}\log b = b^x\log b \qquad \left(\textbf{not } xb^{x-1}\right). \end{aligned}$$

- Derivative of $\log_b x$:

$$\begin{aligned} & f(x) = \log_b x = \frac{\log x}{\log b} \\ \Rightarrow\ & f'(x) = \left(\frac{1}{\log b}\right)\frac{d}{dx}\log x = \left(\frac{1}{\log b}\right)\frac{1}{x} = \frac{1}{x\log b}. \end{aligned}$$

Remark: We know that $\int_1^y \frac{1}{t}\,dt = \log y$ for any $y > 0$. What is $\int_{-1}^y \frac{1}{t}\,dt$ for $y < 0$?

Consider

$$F(x) = \log|x| = \begin{cases} \log x & x > 0, \\ \log(-x) & x < 0. \end{cases}$$

Then

$$\begin{aligned} F'(x) &= \begin{cases} \dfrac{1}{x} & x > 0, \\ \dfrac{1}{-x}(-1) & x < 0 \end{cases} \\ &= \frac{1}{x} \quad \text{for all } x \neq 0. \end{aligned}$$

Therefore, we see that

$$\int \frac{1}{x}\,dx = \log|x| + C \qquad \left(\textbf{not } \frac{x^0}{0}\right),$$

where C is an arbitrary constant. Thus

$$\int x^n\,dx = \begin{cases} \frac{x^{n+1}}{n+1} + C & \text{if } n \neq -1, \\ \\ \log|x| + C & \text{if } n = -1. \end{cases}$$

- Here are some further examples:

$$\frac{d}{dx}\left(\frac{e^{2x}}{e^{2x}+1}\right) = \frac{d}{dx}\left(\frac{1}{1+e^{-2x}}\right) = \frac{-1}{(1+e^{-2x})^2}e^{-2x}(-2) = \frac{2}{(e^x+e^{-x})^2}.$$

- $$3^{\log_3 x} = e^{\log_3 x(\log 3)} = e^{\frac{\log x}{\log 3}\log 3} = x.$$

- $$\log\int e^x\,dx = \log(e^x + C).$$

- $$e^{\int e^x\,dx} = e^{e^x+C} = e^{e^x}e^C = Ae^{e^x},$$

  where $A = e^C$ is a contant.

- $$\lim_{x\to 0^+} x^x = \lim_{x\to 0+}\exp(x\log x) = \exp\left(\lim_{x\to 0^+}\frac{\log x}{\frac{1}{x}}\right)$$
  $$= \exp\left(\lim_{x\to 0}\frac{\frac{1}{x}}{-\frac{1}{x^2}}\right) = \exp\left(\lim_{x\to 0}(-x)\right) = \exp(0) = 1,$$

  by L'Hôpital's Rule.

- $$\frac{d}{dx}\log_2(\sin x) = \left(\frac{1}{\log 2}\right)\frac{d}{dx}\log(\sin x) = \left(\frac{1}{\log 2}\right)\frac{\cos x}{\sin x}.$$

Problem 6.4: Consider a camera flash capacitor with maximum capacity Q_0. If the charge $q(t)$ at time t is given by

$$q(t) = Q_0\left(1 - e^{-\frac{t}{a}}\right),$$

at what time t will the capacitor have attained the charge Q needed to operate the flash unit?

We find

$$\frac{Q}{Q_0} = 1 - e^{-\frac{t}{a}} \Rightarrow \log\left(1 - \frac{Q}{Q_0}\right) = -\frac{t}{a}.$$

Hence $t = -a\log(1 - Q/Q_0)$. For example, if $a = 2$ seconds, $Q = 0.9\,Q_0$, we find that $t = -2\log(0.1) \approx 4.6$ seconds.

Problem 6.5: Find

(a)
$$\lim_{x\to 0}\frac{e^x-1}{x}$$

$$=\lim_{x\to 0}\frac{e^x}{1}=1,$$

on using L'Hôpital's Rule.

(b)
$$\lim_{x\to 0}\frac{e^x-1-x}{x^2}$$

$$=\lim_{x\to 0}\frac{e^x-1}{2x}=\frac{1}{2}\lim_{x\to 0}\frac{e^x-1}{x}=\frac{1}{2}.$$

(c)
$$\frac{d}{dx}\left(\sqrt{x}\right)^x$$

$$=e^{\frac{1}{2}x\log x}\frac{1}{2}(\log x+1)=\left(\sqrt{x}\right)^x\frac{1}{2}(\log x+1).$$

(d)
$$\frac{d}{dx}\int_{\arctan x}^{\log x}\frac{1}{t^7-1}\,dt$$

$$=\left(\frac{1}{\log^7 x-1}\right)\frac{1}{x}-\left(\frac{1}{\arctan^7 x-1}\right)\frac{1}{1+x^2}.$$

6.B Logarithmic Differentiation

Because they can be used to transform multiplication problems into addition problems, logarithms are frequently exploited in calculus to facilitate the calculation of derivatives of complicated products or quotients. For example, if we need to calculate the derivative of a *positive* function $f(x)$, the following procedure may simplify the task:

1. Take the logarithm of both sides of $y=f(x)$.

2. Differentiate each side implicitly with respect to x.

3. Solve for dy/dx.

• Differentiate $y = x^{\sqrt{x}}$ for $x > 0$.

We have

$$\log y = \log x^{\sqrt{x}} = \sqrt{x}\log x.$$

Thus

$$\begin{aligned} \frac{1}{y}\frac{dy}{dx} &= \frac{1}{2\sqrt{x}}\log x + \sqrt{x}\left(\frac{1}{x}\right). \\ \Rightarrow \frac{dy}{dx} &= y\left(\frac{\log x}{2\sqrt{x}} + \frac{1}{\sqrt{x}}\right) \\ &= x^{\sqrt{x}}\left(\frac{\log x + 2}{2\sqrt{x}}\right). \end{aligned}$$

Problem 6.6: Show that the same result follows on differentiating $y = e^{\sqrt{x}\log x}$ directly.

• For $x > 0$ differentiate

$$y = -\frac{x^{\frac{3}{4}}\sqrt{x^2+1}}{(3x+2)^5}.$$

Since

$$\log(-y) = \frac{3}{4}\log x + \frac{1}{2}\log(x^2+1) - 5\log(3x+2),$$

we find

$$\begin{aligned} \frac{1}{y}\frac{dy}{dx} &= \frac{3}{4}\left(\frac{1}{x}\right) + \frac{1}{2}\left(\frac{1}{x^2+1}\right)(2x) - \frac{5}{3x+2}(3) \\ \Rightarrow \frac{dy}{dx} &= -\frac{x^{\frac{3}{4}}\sqrt{x^2+1}}{(3x+2)^5}\left(\frac{3}{4x} + \frac{x}{x^2+1} - \frac{15}{3x+2}\right). \end{aligned}$$

6.C Hyperbolic Functions

Hyperbolic functions are combinations of e^x and e^{-x}:

$$\sinh x = \frac{e^x - e^{-x}}{2}, \cosh x = \frac{e^x + e^{-x}}{2}, \tanh x = \frac{\sinh x}{\cosh x},$$

$$\text{csch} x = \frac{1}{\sinh x}, \quad \text{sech}\, x = \frac{1}{\cosh x}, \quad \coth x = \frac{1}{\tanh x}.$$

Recall that the points $(x, y) = (\cos t, \sin t)$ generate a circle, as t is varied from 0 to 2π, since $x^2 + y^2 \cos^2 t + \sin^2 t = 1$. In contrast, the points $(x, y) = (\cosh t, \sinh t)$

generate a hyperbola, as t is varied over all real values, since $x^2 - y^2 = \cosh^2 t - \sinh^2 t = 1$ (hence the name hyperbolic functions). That is,

$$\left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2 = \frac{e^{2x} + 2 + e^{-2x}}{4} - \frac{e^{2x} - 2 + e^{-2x}}{4} = 1.$$

Note that

$$\frac{d}{dx}\sinh x = \frac{e^x + e^{-x}}{2} = \cosh x,$$

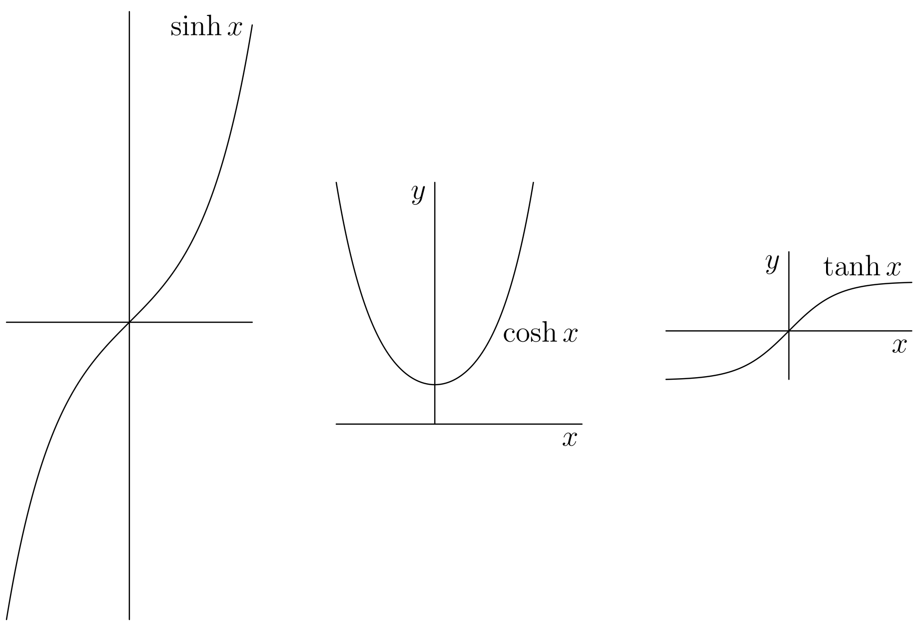
but

$$\frac{d}{dx}\cosh x = \frac{e^x - e^{-x}}{2} = \sinh x,$$

(without any minus sign). Also,

$$\frac{d}{dx}\tanh x = \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} = \frac{1}{\cosh^2 x}.$$

Note that $\sinh x$ and $\tanh x$ are strictly monotonic, whereas $\cosh x$ is strictly decreasing on $(-\infty, 0]$ and strictly increasing on $[0, \infty)$.



Just as the inverse of e^x is $\log x$, the inverse of $\sinh x$ also involves $\log x$. Letting $y = \sinh^{-1} x$, we see that

$$x = \sinh y = \frac{e^y - e^{-y}}{2},$$

so that $e^y - e^{-y} - 2x = 0$. To solve for y, it is convenient to make the substitution $z = e^y$:

$$\begin{aligned} z - \frac{1}{z} - 2x &= 0 \\ \Rightarrow z^2 - 2xz - 1 &= 0. \end{aligned}$$

Thus

$$z = \frac{2x \pm \sqrt{(2x)^2 + 4}}{2},$$

so that $e^y = x \pm \sqrt{x^2+1}$. But since $e^y > 0$ for all $y \in \mathbb{R}$, only the positive square root is relevant. That is, for all real x,

$$\sinh^{-1} x = \log(x + \sqrt{x^2+1}).$$

Problem 6.7: Prove that the two solutions for $\cosh^{-1} x$ are given by $\log(x \pm \sqrt{x^2-1})$. Show directly that $\log(x + \sqrt{x^2-1}) = -\log(x - \sqrt{x^2-1})$.

Problem 6.8: Show that

$$\tanh^{-1} x = \frac{1}{2} \log\left(\frac{1+x}{1-x}\right).$$

Problem 6.9: Show that

$$\frac{d}{dx} \sinh^{-1} x = \frac{d}{dx} \log\left(x + \sqrt{x^2+1}\right) = \frac{1}{\sqrt{x^2+1}}.$$

Also verify this result directly from the fact that

$$\frac{d}{dy} \sinh y = \cosh y.$$

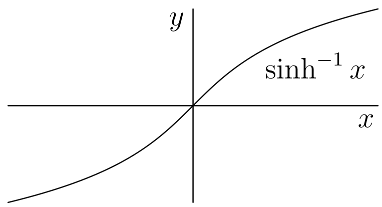
- Thus

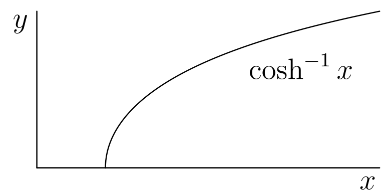
$$\int_0^1 \frac{dx}{\sqrt{1+x^2}} = \left[\sinh^{-1} x\right]_0^1 = \left[\log(x + \sqrt{x^2+1})\right]_0^1 = \log\left(1 + \sqrt{2}\right).$$

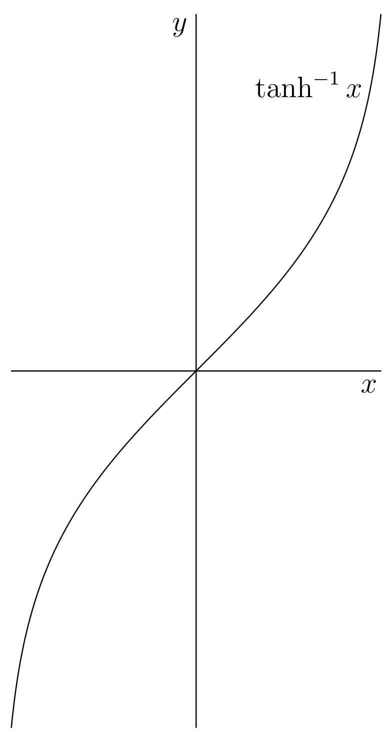
- To find $\frac{d}{dx} \cosh^{-1} x$, we can use the relation $\cosh^2 y - \sinh^2 y = 1$:

$$\begin{aligned} y &= \cosh^{-1} x \\ \Rightarrow x &= \cosh y \\ \Rightarrow \frac{dx}{dy} &= \sinh y = \sqrt{\cosh^2 y - 1} = \sqrt{x^2-1} \\ \Rightarrow \frac{dy}{dx} &= \frac{1}{\sqrt{x^2-1}}, \end{aligned}$$

which gives us an antiderivative for $\displaystyle\int \frac{1}{\sqrt{x^2-1}}\,dx$.


y
$\sinh^{-1} x$
x


y
$\cosh^{-1} x$
x


y
$\tanh^{-1} x$
x

Problem 6.10: Prove that

(a)

$$\cosh^2 t = \frac{\cosh 2t + 1}{2}$$

$$\cosh^2 t = \frac{(e^t + e^{-t})^2}{4} = \frac{e^{2t} + 2 + e^{-2t}}{4} = \frac{\cosh 2t + 1}{2}.$$

(b)

$$\sinh^2 t = \frac{\cosh 2t - 1}{2}$$

$$\sinh^2 t = \frac{(e^t - e^{-t})^2}{4} = \frac{e^{2t} - 2 + e^{-2t}}{4} = \frac{\cosh 2t - 1}{2}.$$

(c)

$$2 \sinh t \cosh t = \sinh 2t$$

$$2 \sinh t \cosh t = 2\frac{(e^t - e^{-t})(e^t + e^{-t})}{4} = \frac{e^{2t} - e^{-2t}}{2} = \sinh 2t.$$

Chapter 7

Techniques of Integration

In this chapter we develop several fundamental techniques of integration.

7.A Change of Variables

Q. What is $\int \tan x \, dx = \displaystyle\int \frac{\sin x}{\cos x} \, dx$?

A. On differentiating $F(x) = -\log|\cos x| + C$, we see that $\int \tan x \, dx = F(x)$.

Q. Are there systematic ways of finding such antiderivatives?

A. Yes, the following theorem (sometimes known as the *Substitution Rule*) is often helpful.

Theorem 7.1 (Change of Variables): *Suppose g' is continuous on $[a,b]$ and f is continuous on $g([a,b])$. Then*

$$\int_{x=a}^{x=b} f(\underbrace{g(x)}_{u}) \, \underbrace{g'(x) \, dx}_{du} = \int_{u=g(a)}^{u=g(b)} f(u) \, du.$$

Proof: Theorem 5.8 $\Rightarrow$ f has an antiderivative F:

$$F'(u) = f(u) \quad \forall u \in g([a,b]).$$

Consider $H(x) = F(g(x))$. Then

$$\begin{aligned} H'(x) &= F'(g(x))g'(x) \\ &= f(g(x))g'(x), \end{aligned}$$

that is, H is an antiderivative of $(f \circ g)g'$. Letting $u = g(x)$, we may then write

$$\int f(g(x))g'(x) \, dx = H(x) = F(g(x)) = F(u) = \int f(u) \, du$$

and, using the FTC,

$$\int_a^b f(g(x))g'(x) = [H(x)]_a^b = [F(g(x))]_a^b = F(g(b)) - F(g(a)) = \int_{u=g(a)}^{u=g(b)} f(u)\,du.$$

- Suppose we wish to calculate $\int_0^1 (x^2+2)^{99}\,2x\,dx$. One could expand out this polynomial and integrate term by term, but a much easier way to evaluate this integral is to make the substitution $u = g(x) = x^2+2$. To help us remember the factor $\frac{du}{dx} = g'(x) = 2x$ we formally write $du = g'(x)\,dx = 2x\,dx$,

$$\int_{x=0}^{x=1} (x^2+2)^{99} 2x\,dx = \int_{u=2}^{u=3} u^{99}\,du = \left.\frac{u^{100}}{100}\right|_2^3 = \frac{3^{100}-2^{100}}{100}.$$

- To compute $\int x\sqrt{x^2+1}\,dx$, it is helpful to substitute $u = x^2+1 \Rightarrow du = 2x\,dx$.

$$\begin{aligned}\int x\sqrt{x^2+1}\,dx = \int u^{1/2}\frac{du}{2} &= \frac{1}{\cancel{2}}\frac{\cancel{2}}{3}u^{3/2} + C \qquad \longleftarrow \text{(don't leave in this form)}\\ &= \frac{1}{3}(x^2+1)^{3/2} + C.\end{aligned}$$

Check:

$$\frac{d}{dx}\left[\frac{1}{3}(x^2+1)^{3/2} + C\right] = \frac{1}{\cancel{3}}\frac{\cancel{3}}{\cancel{2}}(x^2+1)^{1/2}\,\cancel{2}\,x = x\sqrt{x^2+1}.$$

- The change of variables $u = e^t \Rightarrow du = e^t\,dt$ allows us to evaluate

$$\int \frac{e^t}{e^t+1}\,dt = \int \frac{du}{u+1} = \log|u+1| + C = \log(e^t+1) + C.$$

- The substitution $u = \frac{x}{a} \Rightarrow x = au \Rightarrow dx = a\,du$, where a is a constant, allows us to evaluate any integral of the form

$$\begin{aligned}\int \frac{1}{x^2+a^2}\,dx &= \int \frac{1}{a^2\left(\frac{x^2}{a^2}+1\right)}\,dx\\ &= \frac{1}{a^2}\int \frac{1}{u^2+1}\,a\,du\\ &= \frac{1}{a}\int \frac{1}{u^2+1}\,du\\ &= \frac{1}{a}\tan^{-1} u + C\\ &= \frac{1}{a}\tan^{-1}\left(\frac{x}{a}\right) + C.\end{aligned}$$

Problem 7.1: Find

$$\int \cos t \sqrt{\sin t}\, dt.$$

$$= \frac{2}{3} \sin^{3/2} t + C.$$

Problem 7.2: Let α be a real number. Find

$$\int x^{-\alpha} e^{-\alpha x} \left(\frac{1}{x} + 1 \right) dx.$$

We use the substitution $u = \log x + x$ to rewrite

$$\int e^{-\alpha(\log x + x)} \left(\frac{1}{x} + 1 \right) dx = \int e^{-\alpha u}\, du = \begin{cases} \frac{-e^{-\alpha u}}{\alpha} + C & \text{if } \alpha \neq 0, \\ u + C & \text{if } \alpha = 0. \end{cases}$$
$$= \begin{cases} -\frac{x^{-\alpha} e^{-\alpha x}}{\alpha} + C & \text{if } \alpha \neq 0, \\ \log x + x + C & \text{if } \alpha = 0. \end{cases}$$

Problem 7.3: Let $f : [0, 1] \to \mathbb{R}$ be a continuous function. Prove that

$$\int_0^{\pi/2} f(\sin x)\, dx = \int_0^{\pi/2} f(\cos x)\, dx.$$

This follows on using the substitution $u = \pi/2 - x$:

$$\int_0^{\pi/2} f(\sin x)\, dx = \int_0^{\pi/2} f\left(\cos\left(\frac{\pi}{2} - x\right)\right)\, dx = -\int_{\pi/2}^{0} f(\cos u)\, du = \int_0^{\pi/2} f(\cos u)\, du.$$

7.B Integration by Parts

The Change of Variables theorem in the last section is seen to be just an integral version of the Chain Rule. Another important and frequently used rule in differential calculus is the product rule.

$\dfrac{d}{dx}$	$\displaystyle\int dx$
Chain Rule	Change of Variables
Product Rule	Integration by Parts

Table 7.1: Techniques of Integration

Q. Does the product rule also have an integral version?

A. Yes, it is called Integration by Parts, as illustrated in Table 7.1.

Theorem 7.2 (*Integration by Parts*): *Suppose f' and g' are continuous functions. Then*

$$\int fg' = fg - \int f'g.$$

Proof: Let $F = fg$. Then $F' = f'g + fg'$, so fg is an antiderivative of $f'g + fg'$:

$$\int (f'g + fg') = fg \qquad \text{to with a constant.}$$

That is,

$$\boxed{\int fg'\,dx = fg - \int f'g\,dx.}$$

Remark: Letting $u = f(x)$, so that $du = f'(x)\,dx$, and $v = g(x)$, so that $dv = g'(x)\,dx$, we may rewrite the integration by parts formula as

$$\int u\,dv = uv - \int v\,du.$$

Remark: For definite integrals we have, by the FTC,

$$\int_a^b fg'\,dx = \Big[fg\Big]_a^b - \int_a^b f'g\,dx.$$

- We can integrate $\int x \sin x \, dx$ using Integration by Parts:

$$\int \underbrace{x}_{f} \underbrace{\sin x}_{g'} \, dx = \underbrace{x}_{f} \underbrace{(-\cos x)}_{g} - \int \underbrace{1}_{f'} \underbrace{(-\cos x)}_{g} \, dx$$

$$\therefore \int x \sin x \, dx = -x \cos x + \sin x + C$$

Try to pick f so that f' is simple and g' has a known antiderivative. If instead we pick

$$f = \sin x \quad (\Rightarrow f' = \cos x)$$

and

$$g' = x \quad \left(\Rightarrow g = \frac{x^2}{2}\right)$$

then the Integration by Parts formula leads to an even more complicated integral:

$$\int \sin x \; x \, dx = \sin x \left(\frac{x^2}{2}\right) - \int \cos x \left(\frac{x^2}{2}\right) = \dots .$$

So this choice of f and g' was not fruitful.

- Noting that

$$\int \log x \, dx = \int 1 \cdot \log x \, dx,$$

we might be tempted to try Integration by Parts, setting $f = 1$ and $g' = \log x$:

$$\begin{aligned} \int \log x \, dx &= 1 \int \log x \, dx - \int 0 \left[\int \log x \, dx \right] dx \\ &= \int \log x \, dx. \end{aligned}$$

This doesn't help! Instead, we could try $f = \log x$ and $g' = 1$:

$$\begin{aligned} \int \underbrace{\log x}_{f} \cdot \underbrace{1}_{g'} \, dx &= \underbrace{\log x}_{f} \cdot \underbrace{x}_{g} - \int \underbrace{\frac{1}{x}}_{f'} \underbrace{x}_{g} \, dx \\ &= x \log x - x + C. \end{aligned}$$

- Similarly we can integrate $\tan^{-1}$ by parts and use the substitution $u = x^2$ to find

$$\int_0^1 \underbrace{\tan^{-1} x}_{f} \cdot \underbrace{1}_{g'} \, dx = \Big[x \tan^{-1} x \Big]_0^1 - \int_0^1 \frac{x}{1+x^2} \, dx$$
$$= 1 \tan^{-1} 1 - 0 - \int_{0^2}^{1^2} \frac{1}{1+u} \frac{du}{2}$$
$$= \frac{\pi}{4} - \frac{1}{2} [\log |1+u|]_0^1$$
$$= \frac{\pi}{4} - \frac{1}{2} \log 2.$$

- In order to find

$$\int \underbrace{x^2}_{f} \underbrace{e^x}_{g'} \, dx = x^2 e^x - \int 2x e^x \, dx,$$

we need to know $\int 2xe^x \, dx$. But that integral is just twice $\int xe^x \, dx$, which we can find by applying integration by parts again:

$$\int xe^x \, dx = xe^x - \int e^x \, dx = xe^x - e^x + C.$$

Thus

$$\int x^2 e^x \, dx = x^2 e^x - 2(xe^x - e^x + C)$$
$$= x^2 e^x - 2xe^x + 2e^x + C_2, \qquad \text{where } C_2 = -2C.$$

- We can even find integrals of the form

$$I = \int \underbrace{\sin x}_{f} \underbrace{e^x}_{g'} \, dx = \sin x \, e^x - \int \cos x \, e^x \, dx.$$

What is $\int \cos x \, e^x \, dx$?

$$\int \underbrace{\cos x}_{f} \underbrace{e^x}_{g'} \, dx = \cos x \, e^x - \int (-\sin x) \, e^x \, dx$$
$$= \cos x \, e^x + I.$$

Thus $I = \sin x \, e^x - (\cos x \, e^x + I)$, from which we find $I = \frac{1}{2} \sin x \, e^x - \frac{1}{2} \cos x \, e^x + C$.

- For nonzero real numbers a and b find

$$\begin{aligned} I &= \int e^{ax} \sin bx \, dx, \\ J &= \int e^{ax} \cos bx \, dx. \end{aligned}$$

On integrating by parts, we obtain

$$I = \int e^{ax} \sin bx \, dx = \frac{1}{a} e^{ax} \sin bx - \frac{b}{a} \underbrace{\int e^{ax} \cos bx \, dx}_{J}.$$

$$J = \int \underbrace{e^{ax}}_{g'} \underbrace{\cos bx}_{f} \, dx = \frac{1}{a} e^{ax} \cos bx + \frac{b}{a} \underbrace{\int e^{ax} \sin bx \, dx}_{I},$$

We thus need to solve the system of equations

$$\begin{aligned} I &= \frac{1}{a} e^{ax} \sin bx - \frac{b}{a} J, \\ J &= \frac{1}{a} e^{ax} \cos bx + \frac{b}{a} I. \end{aligned}$$

$$\begin{aligned} \Rightarrow J &= \frac{1}{a} e^{ax} \cos bx + \frac{b}{a^2} e^{ax} \sin bx - \frac{b^2}{a^2} J \\ \left(1 + \frac{b^2}{a^2}\right) J &= \left(\frac{1}{a} \cos bx + \frac{b}{a^2} \sin bx\right) e^{ax}. \end{aligned}$$

$$\begin{aligned} \Rightarrow J &= \frac{a \cos bx + b \sin bx}{a^2 + b^2} e^{ax} + C_1, \\ I &= \frac{a \sin bx - b \cos bx}{a^2 + b^2} e^{ax} + C_2. \end{aligned}$$

- We now compute, for $n \geq 2$,

$$\begin{aligned} I_n &= \int \sin^n x \, dx \\ &= \int \underbrace{\sin^{n-1} x}_{f} \underbrace{\sin x}_{g'} \, dx \\ &= \sin^{n-1} x(-\cos x) - \int \underbrace{(n-1) \sin^{n-2} x(\cos x)}_{f'}(-\cos x) \, dx \end{aligned}$$

Now

$$\int \sin^{n-2} x \underbrace{\cos^2 x}_{1-\sin^2 x} \, dx = \int (\sin^{n-2} x - \sin^n x)\, dx = I_{n-2} - I_n.$$

Thus

$$\begin{aligned} I_n &= -\sin^{n-1} x \cos x + (n-1)(I_{n-2} - I_n) \\ \Rightarrow I_n &= -\sin^{n-1} x \cos x + (n-1) I_{n-2} - nI_n + I_n \\ \Rightarrow nI_n &= -\sin^{n-1} x \cos x + (n-1) I_{n-2}. \end{aligned}$$

That is,

$$\boxed{\int \sin^n x \, dx = -\frac{1}{n} \sin^{n-1} x \, \cos x + \frac{n-1}{n} \int \sin^{n-2} x \, dx.}$$

This is known as a *reduction formula*.

- For $n = 2$, the reduction formula states that

$$\begin{aligned} \int \sin^2 x \, dx &= -\frac{1}{2} \sin x \cos x + \frac{1}{2} \int 1 \, dx \\ &= -\frac{1}{2} \sin x \cos x + \frac{1}{2} x + C. \end{aligned}$$

Alternatively, one can evaluate this integral using trigonometric identities:

$$\begin{aligned} \int \sin^2 x \, dx = \int \frac{1 - \cos 2x}{2} \, dx &= \frac{1}{2} x - \frac{1}{4} \sin 2x + C \\ &= \frac{1}{2} x - \frac{1}{2} \sin x \cos x + C. \end{aligned}$$

- For $n = 3$,

$$\begin{aligned} \int \sin^3 x \, dx &= -\frac{1}{3} \sin^2 x \cos x + \frac{2}{3} \int \sin x \, dx \\ &= -\frac{1}{3} \sin^2 x \cos x - \frac{2}{3} \cos x + C. \end{aligned}$$

- Find $I = \int x^3 \log x \, dx$.

$$I = \frac{x^4}{4} \log x - \int \frac{x^4}{4} \frac{1}{x} \, dx = \frac{x^4}{4} \log x - \frac{x^4}{16} + C.$$

- We can also find a reduction formula for integrals of the form

$$
\begin{aligned}
I_n &= \int \underbrace{x^\mu}_{g'} \underbrace{\log^n x}_{f}\, dx, \qquad \text{where } \mu \neq -1,\ \mu \in \mathbb{R},\ n = 0,1,2,3,\ldots \\
&= \underbrace{\frac{x^{\mu+1}}{\mu+1}}_{g} \underbrace{\log^n x}_{f} - \int \underbrace{\frac{x^{\mu+1}}{\mu+1}}_{g} \underbrace{n \log^{n-1} x \left(\frac{1}{x}\right)}_{f'} dx \\
&= \frac{x^{\mu+1}}{\mu+1} \log^n x - \frac{n}{\mu+1} \int x^\mu \log^{n-1} x\, dx.
\end{aligned}
$$

Thus,

$$\boxed{I_n = \frac{x^{\mu+1}}{\mu+1} \log^n x - \frac{n}{\mu+1} I_{n-1}.}$$

When $\mu = -1$, we can evaluate I_n directly, by making the substitution $u = \log x$:

$$I_n = \int \frac{\log^n x}{x}\, dx = \frac{\log^{n+1} x}{n+1} + C.$$

- For $n = 0$ and $\mu \neq -1$,

$$I_0 = \int x^\mu\, dx = \frac{x^{\mu+1}}{\mu+1} + C.$$

- For $n = 1$ and $\mu \neq -1$,

$$
\begin{aligned}
I_1 = \int x^\mu \log x\, dx &= \frac{x^{\mu+1}}{\mu+1} \log x - \frac{1}{\mu+1} I_0 \\
&= \frac{x^{\mu+1}}{\mu+1} \log x - \frac{x^{\mu+1}}{(\mu+1)^2} + C.
\end{aligned}
$$

- An important integral that will soon need is, for $n \geq 1$ and $a \neq 0$,

$$
\begin{aligned}
J_n &= \int \frac{1}{(x^2+a^2)^n} \cdot 1\, dx \\
&= \frac{x}{(x^2+a^2)^n} + 2n \int \frac{x^2}{(x^2+a^2)^{n+1}}\, dx \\
&= \frac{x}{(x^2+a^2)^n} + 2n \int \frac{(x^2+a^2)-a^2}{(x^2+a^2)^{n+1}}\, dx \\
&= \frac{x}{(x^2+a^2)^n} + 2n \left(\int \frac{1}{(x^2+a^2)^n}\, dx - \int \frac{a^2}{(x^2+a^2)^{n+1}}\, dx \right) \\
&= \frac{x}{(x^2+a^2)^n} + 2n(J_n - a^2 J_{n+1}) \\
\Rightarrow (1-2n) J_n &= \frac{x}{(x^2+a^2)^n} - 2na^2 J_{n+1}.
\end{aligned}
$$

The resulting reduction formula,

$$\boxed{J_{n+1} = \frac{1}{2na^2}\frac{x}{(x^2+a^2)^n} + \frac{2n-1}{2na^2}J_n} \quad (n \geq 1, a \neq 0),$$

together with the result (for $a \neq 0$)

$$J_1 = \int \frac{1}{x^2+a^2}\,dx = \frac{1}{a}\arctan\frac{x}{a} + C,$$

allows us to compute J_n for any $n \geq 1$.

Problem 7.4: Find

$$\int_0^1 \arcsin x\,dx.$$

$$\int_0^1 1\cdot\arcsin x\,dx = [x\arcsin x]_0^1 - \int_0^1 \frac{x}{\sqrt{1-x^2}}\,dx = \arcsin 1 + \left[(1-x^2)^{1/2}\right]_0^1 = \frac{\pi}{2} - 1.$$

Problem 7.5: Let $P(x)$ be a polynomial of degree n. Prove that

$$\int P(x)e^x\,dx = e^x\sum_{k=0}^{n}(-1)^k P^{(k)}(x) + C,$$

where $P^{(k)}$ denotes the kth derivative of P. Give an explicit reason why the sum terminates at $k = n$.

This follows immediately on integrating by parts n times, using $f(x) = P(x)$ and $g(x) = e^x$. Alternatively, we can verify the result by noting that the derivative of the right-hand side is

$$\begin{aligned} e^x\sum_{k=0}^{n}(-1)^k P^{(k)}(x) + e^x\sum_{k=0}^{n}(-1)^k P^{(k+1)}(x) &= e^x\sum_{k=0}^{n}(-1)^k P^{(k)}(x) - e^x\sum_{k=1}^{n+1}(-1)^k P^{(k)}(x) \\ &= e^x P^{(0)}(x) - e^x(-1)^{n+1}P^{(n+1)}(x) = e^x P(x) \end{aligned}$$

since $P^{(n+1)}(x) = 0$.

7.C Integrals of Trigonometric Functions

Often we encounter integrals of the form

$$\int \sin^m x\cos^n x\,dx,$$

where m and n are integers. Here is an integration strategy:

Case I. If either of the integers m or n is odd, separate out one factor of $\sin x$ or $\cos x$ so that the rest of the integrand may be written entirely as a polynomial in $\cos x$ or a polynomial in $\sin x$, as the case may be. Then make the appropriate substitution. (Note: if both m and n are odd there will be two possible ways of doing this.)

Case II. If m and n are both even use the addition formulae

$$\sin^2 x = \frac{1-\cos 2x}{2}, \quad \cos^2 x = \frac{1+\cos 2x}{2}, \quad 2\sin x\cos x = \sin 2x,$$

possibly repeatedly, to reduce the problem to the form of Case I.

- Find $\int \sin^3 x \cos^2 x\, dx$, using the substitution $u = \cos x$ $(du = -\sin x\, dx)$,

$$\begin{aligned}\int \sin^3 x\cos^2 x\, dx &= \int \sin x(1-\cos^2 x)\cos^2 x\, dx \\ &= -\int (1-u^2)u^2\, du = -\int (u^2-u^4)\, du \\ &= -\left(\frac{u^3}{3}-\frac{u^5}{5}\right)+C = \frac{1}{5}\cos^5 x - \frac{1}{3}\cos^3 x + C.\end{aligned}$$

- Find $\int \sin^2 x \cos^3 x\, dx$, using the substitution $u = \sin x$ $(du = \cos x\, dx)$,

$$\begin{aligned}\int \sin^2 x\cos^3 x\, dx &= \int \sin^2 x(1-\sin^2 x)\cos x\, dx \\ &= \int u^2(1-u^2)\, du \\ &= \frac{u^3}{3}-\frac{u^5}{5}+C = \frac{1}{3}\sin^3 x - \frac{1}{5}\sin^5 x + C.\end{aligned}$$

- We can use the fact that $\sin^2 2x = (1-\cos 4x)/2$ to compute

$$\begin{aligned}\int \sin^2 x\cos^2 x\, dx &= \int \left(\frac{1}{2}\sin 2x\right)^2 dx \\ &= \frac{1}{4}\int \frac{1-\cos 4x}{2}\, dx \\ &= \frac{1}{8}\left[x - \frac{\sin 4x}{4}\right] + C.\end{aligned}$$

Problem 7.6: Find

$$\int \cos^7 x \sin^2 x \, dx$$

Let $u = \sin x$. The integral then evaluates to

$$\int (1-u^2)^3 u^2 \, du = \int (1-3u^2+3u^4-u^6)u^2 \, du = \int (u^2-3u^4+3u^6-u^8) \, du$$
$$= \frac{u^3}{3} - \frac{3u^5}{5} + \frac{3u^7}{7} - \frac{u^9}{9} + C = \frac{\sin^3 x}{3} - \frac{3\sin^5 x}{5} + \frac{3\sin^7 x}{7} - \frac{\sin^9 x}{9} + C.$$

Remark: In view of Problem 6.10, the same technique can be used to compute $\int \cosh^m t \sinh^n t \, dt$ for integer values of m and n.

Remark: One can use a similar technique to compute certain integrals of the form

$$\int \tan^m x \sec^n x \, dx,$$

by exploiting the Pythagorean relation $\tan^2 x + 1 = \sec^2 x$, along with the derivatives

$$\frac{d}{dx} \tan x = \sec^2 x$$

and

$$\frac{d}{dx} \sec x = \frac{d}{dx}\left(\frac{1}{\cos x}\right) = \frac{+\sin x}{\cos^2 x} = \sec x \tan x.$$

For example, if n is an even natural number, the substitution $u = \tan x$ will reduce the integrand to a polynomial. If m is an odd natural number, the substitution $u = \sec x$ will work.

- Letting $u = \tan x$, we find

$$\begin{aligned} &\int \tan^4 x \sec^4 x \, dx \\ &= \int \tan^4 x (1+\tan^2 x) \underbrace{\sec^2 x \, dx}_{du} \\ &= \int u^4(1+u^2) \, du = \frac{u^5}{5} + \frac{u^7}{7} + C = \frac{\tan^5 x}{5} + \frac{\tan^7 x}{7} + C. \end{aligned}$$

- Letting $u = \sec x$, we find

$$\begin{aligned}
&\int \tan^3 x \sec^3 x\,dx \\
&= \int \tan^2 x \sec^2 x \underbrace{\sec x \tan x\,dx}_{du} \\
&= \int (u^2-1)u^2\,du = \frac{u^5}{5} - \frac{u^3}{3} + C \\
&= \frac{\sec^5 x}{5} - \frac{\sec^3 x}{3} + C.
\end{aligned}$$

Remark: The technique of extracting out an odd factor of $\cos x$ or $\sin x$ can even be applied if one or both of the integers m and n are negative. For example, we can compute the indefinite integral of $\sec x$ by rewriting the integrand and using the substitution $u = \sin x$,

$$\begin{aligned}
&\int \sec x\,dx = \int \frac{\cos x}{\cos^2 x}\,dx = \int \frac{\cos x}{1-\sin^2 x}\,dx \\
&= \int \frac{1}{1-u^2}\,dx = \int \left(\frac{\frac{1}{2}}{1+u} + \frac{\frac{1}{2}}{1-u}\right)dx \\
&= \frac{1}{2}\log|1+u| - \frac{1}{2}\log|1-u| + C \\
&= \frac{1}{2}\log\left|\frac{1+\sin x}{1-\sin x}\right| + C \\
&= \frac{1}{2}\log\left|\frac{(1+\sin x)^2}{1-\sin^2 x}\right| + C \\
&= \log\left|\frac{1+\sin x}{\cos x}\right| + C \\
&= \log|\sec x + \tan x| + C. \qquad (7.1)
\end{aligned}$$

Alternatively, we could use the following trick to obtain this integral:

$$\begin{aligned}
\int \sec x\,dx &= \int \sec x \left(\frac{\sec x + \tan x}{\sec x + \tan x}\right) dx \\
&= \int \frac{(\sec^2 x + \sec x \tan x)\,dx}{\sec x + \tan x} \\
&= \log|\sec x + \tan x| + C.
\end{aligned}$$

- Compute

$$\int \frac{1}{\cos^3 x}\,dx = \int \sec^3 x\,dx = \int \underbrace{\sec x}_{f}\underbrace{\sec^2 x}_{g'}\,dx.$$

On integrating by parts, we find that

$$\begin{aligned}\int \frac{1}{\cos^3 x}\,dx &= \sec x \tan x - \int (\sec x \tan x)\tan x\,dx\\ &= \sec x \tan x - \int \sec x(\sec^2 x - 1)\,dx\\ &= \sec x \tan x - \int \sec^3 x\,dx + \int \sec x\,dx.\end{aligned}$$

Thus

$$\begin{aligned}2\int \sec^3 x\,dx &= \sec x \tan x + \log|\sec x + \tan x| + C\\ \Rightarrow \int \sec^3 x\,dx &= \frac{1}{2}(\sec x \tan x + \log|\sec x + \tan x|) + C.\end{aligned}$$

Definition: A *birational function* $R(x,y)$ is a rational function of the form

$$R(x,y) = \frac{\displaystyle\sum_{ij} a_{ij}x^i y^j}{\displaystyle\sum_{kl} b_{k\ell}x^k y^\ell},$$

where a_{ij} and $b_{k\ell}$ are real numbers (at least one $b_{k\ell}$ must be nonzero). Note that a birational function is a rational function of each of its arguments (holding the other argument fixed).

Remark: Any birational function $R(\sin x, \cos x)$ of $\sin x$ and $\cos x$ can be converted to a rational function of t with the *universal substitution* $t = \tan\frac{x}{2}$ since

$$\cos x = \frac{1-t^2}{1+t^2}, \quad \sin x = \frac{2t}{1+t^2}, \quad dx = \frac{2}{1+t^2}\,dt. \tag{7.2}$$

Problem 7.7: Use the trigonometric addition formulae to prove that $\cos x$, $\sin x$, and dx transform according to Eq. (7.2).

- We can use the universal substitution to find an alternative expression for $\int \sec x\,dx$:

$$\begin{aligned}\int \sec x\,dx &= \int \frac{1+t^2}{1-t^2}\left(\frac{2}{1+t^2}\right)dt = \int \left(\frac{2}{1-t^2}\right)dt = \int \left(\frac{1}{1+t} + \frac{1}{1-t}\right)dt\\ &= \log\left|\frac{1+\tan(x/2)}{1-\tan(x/2)}\right| + C = \log\left|\tan\left(\frac{x}{2}+\frac{\pi}{4}\right)\right| + C.\end{aligned} \tag{7.3}$$

Problem 7.8: Show directly that

$$\tan\left(\frac{x}{2}+\frac{\pi}{4}\right) = \sec x + \tan x.$$

Hence, Eq. (7.3) and Eq. (7.1) are equivalent.

Remark: Although the universal substitution will always work (i.e. it can always be used to reduce the integrand to a rational function), it should be viewed only as a last resort; often, easier methods are available.

7.D Partial Fraction Decomposition

[Muldowney 1990, p. 211]
[Spivak 1994, p. 374]

Consider the following techniques for integrating rational functions.

- Find
$$\int \frac{x+1}{x}\,dx = \int \left(1+\frac{1}{x}\right)dx = x + \log|x| + C.$$

- Similarly,
$$\int \frac{x}{x+1}\,dx = \int \left(\frac{x+1}{x+1} - \frac{1}{x+1}\right)dx = x - \log|x+1| + C.$$

- In general,
$$\begin{aligned}\int \frac{ax+b}{cx+d}\,dx &= \int \frac{ax+\frac{ad}{c}-\frac{ad}{c}+b}{cx+d}\,dx\\ &= \int \frac{\frac{a}{c}(cx+d)-\frac{ad}{c}+b}{cx+d}\,dx = \int \left[\frac{a}{c}+\left(b-\frac{ad}{c}\right)\frac{1}{cx+d}\right]dx\\ &= \left(\frac{a}{c}\right)x + \left(\frac{bc-ad}{c^2}\right)\log|cx+d| + C.\end{aligned}$$

- We can also evaluate integrals like
$$\int \frac{1}{x(x+1)}\,dx = \int \frac{(x+1)-x}{x(x+1)}\,dx = \int \left(\frac{1}{x}-\frac{1}{x+1}\right)dx = \log|x| - \log|x+1| + C.$$

• Here is a more complicated example:

$$\begin{aligned}\int \frac{1}{x^2(1+x^2)^2}\,dx &= \int \frac{(1+x^2)-x^2}{x^2(1+x^2)^2}\,dx = \int \left[\frac{1}{x^2(1+x^2)} - \frac{1}{(1+x^2)^2}\right] dx \\ &= \int \left[\frac{(1+x^2)-x^2}{x^2(1+x^2)} - \frac{1}{(1+x^2)^2}\right] dx \\ &= \int \left[\frac{1}{x^2} - \frac{1}{1+x^2} - \frac{1}{(1+x^2)^2}\right] dx = -\frac{1}{x} - \arctan x - J_2(1),\end{aligned}$$

where

$$J_2(a) = \int \frac{dx}{(x^2+a^2)^2}.$$

Recall that the indefinite integral

$$J_n(a) = \int \frac{dx}{(x^2+a^2)^n}$$

can be evaluated using the reduction formula

$$J_{n+1} = \frac{1}{2na^2}\frac{x}{(x^2+a^2)^n} + \frac{2n-1}{2na^2}J_n.$$

Setting $n = 1$ and $a = 1$ yields

$$J_2(1) = \frac{1}{2}\left(\frac{x}{x^2+1}\right) + \frac{1}{2}J_1,$$

where $J_1 = \arctan x + C$. Hence,

$$\begin{aligned}\int \frac{1}{x^2(1+x^2)^2}\,dx &= -\frac{1}{x} - \arctan x - \frac{1}{2}\left(\frac{x}{x^2+1}\right) - \frac{1}{2}\arctan x + C \\ &= -\frac{1}{x} - \frac{3}{2}\arctan x - \frac{1}{2}\left(\frac{x}{x^2+1}\right) + C.\end{aligned}$$

Q. Can these techniques be generalized for integrating any rational function?

A. Yes, using the general method of *partial fraction decomposition*:

Suppose we wish to evaluate the integral

$$\int \frac{P(x)}{Q(x)}\,dx,$$

where P and Q are polynomials functions of x.

Step 1: If the degree of $P \geq$ degree of Q, we rewrite the integrand in *proper form*:

$$\frac{P(x)}{Q(x)} = S(x) + \frac{R(x)}{Q(x)},$$

such that the degree of R is less than the degree of Q.

- Suppose that $P(x) = x^3 + x$ and $Q(x) = x - 1$. We see that $\deg P = 3 \geq \deg Q = 1$, so we put $P(x)/Q(x)$ in proper form, using long division:

$$\frac{P(x)}{Q(x)} = \frac{x^3+x}{x-1} = \underbrace{x^2+x+2}_{S(x)} + \underbrace{\frac{2}{x-1}}_{\frac{R(x)}{Q(x)}}.$$

We can now go ahead and integrate $S(x)$ and, in this case, also $R(x)/Q(x)$ without doing any further work:

$$\begin{aligned}\int \frac{x^3+x}{x-1}\,dx &= \int \left(x^2+x+2+\frac{2}{x-1}\right) dx \\ &= \frac{x^3}{3} + \frac{x^2}{2} + 2x + 2\log|x-1| + C.\end{aligned}$$

Remark: At this stage, we will always be able to find an antiderivative for $S(x)$ since it is just a polynomial. The following steps may be needed to integrate the remaining term $R(x)/Q(x)$.

Step 2: Factor $Q(x)$ as far as possible, into products of linear factors and irreducible quadratic factors.

- $Q(x) = x^4 - 16 = (x^2-4)(x^2+4) = (x-2)(x+2)(x^2+4)$.
- $Q(x) = (x+1)(x+2)^2(x^2+x+3)(x^2+x+4)^2$.

Step 3: Suppose $Q(x)$ has the form

$$Q(x) = A(x-a)^n \dots (x^2+\gamma x+\lambda)^m \dots,$$

where the *discriminant* $\gamma^2 - 4\lambda < 0$, so that $x^2+\gamma x+\lambda$ cannot be factorized into linear factors with real coefficients. It is then possible to express $R(x)/Q(x)$, where $\deg R < \deg Q$ in the form

$$\begin{aligned}\frac{R(x)}{Q(x)} = &\left[\frac{A_1}{(x-a)} + \frac{A_2}{(x-a)^2} + \dots + \frac{A_n}{(x-a)^n}\right] + \dots \\ &+ \left[\frac{\Gamma_1 x + \Lambda_1}{x^2+\gamma x+\lambda} + \dots + \frac{\Gamma_m x + \Lambda_m}{(x^2+\gamma x+\lambda)^m}\right] + \dots .\end{aligned}$$

Step 4: Solve for the coefficients in the numerator by equating like powers of x.

- We can solve
$$\frac{1}{x(x+1)} = \frac{A}{x} + \frac{B}{x+1} = \frac{A(x+1)+Bx}{x(x+1)}$$
for the coefficients A and B by equating like polynomial coefficients in the numerator. On setting
$$1 = A(x+1) + Bx = (A+B)x + A,$$
we see that the coefficients of x^0 and x^1 are
$$\begin{aligned} x^0 &: 1 = A, \\ x^1 &: 0 = A + B. \end{aligned}$$
The unique solution to these equations is $A = 1$ and $B = -1$.

Step 5: Integrate each term separately. Each term of the form $A_1/(x-a)$ has antiderivative $A_1 \log|x-a| + C$. Also,
$$\int \frac{A_n}{(x-a)^n} = -\left(\frac{A_n}{n-1}\right)\frac{1}{(x-a)^{n-1}} + C, \qquad \text{for } n \geq 2.$$
Finally, each integral of the form
$$\int \frac{\Gamma_m x + \Lambda_m}{(x^2+\gamma x+\lambda)^m}\,dx$$
can be evaluated by "completing the square" in the denominator,
$$\int \frac{\Gamma_m x + \Lambda_m}{\left[\left(x+\frac{\gamma}{2}\right)^2 - \frac{\gamma^2}{4} + \lambda\right]^m}\,dx.$$
On making the substitution $u = x + \frac{\gamma}{2}$, we can then express the result as the sum of two integrals,
$$\Gamma_m \int \frac{u}{(u^2+a^2)^m}\,du + \left(\Lambda_m - \frac{\gamma}{2}\Gamma_m\right)\int \frac{1}{(u^2+a^2)^m}\,du,$$
where $a^2 = \lambda - \frac{\gamma^2}{4} = -(\gamma^2 - 4\lambda)/4 > 0$. The first integral can be easily computed using the substitution $w = u^2 + a^2$ and the second integral is just $J_m(a)$.

Remark: If $\gamma = 0$ there is no need to complete the square.

- To find

$$\int \frac{1-x+2x^2-x^3}{x(x^2+1)^2}\,dx$$

we write

$$\frac{1-x+2x^2-x^3}{x(x^2+1)^2} = \frac{A}{x} + \frac{Bx+C}{x^2+1} + \frac{Dx+E}{(x^2+1)^2}.$$

This requires that

$$\begin{aligned} -x^3+2x^2-x+1 &= A(x^2+1)^2 + (Bx+C)x(x^2+1) + (Dx+E)x \\ &= A(x^4+2x^2+1) + B(x^4+x^2) + C(x^3+x) + Dx^2 + Ex \end{aligned}$$

Thus

$$\begin{aligned} x^4 &: \quad 0 = A+B, \\ x^3 &: -1 = C, \\ x^2 &: \quad 2 = 2A+B+D, \\ x^1 &: -1 = C+E \Rightarrow E = 0, \\ x^0 &: \quad 1 = A \Rightarrow B = -1 \text{ and } D = 1, \end{aligned}$$

so that

$$\begin{aligned} \int \frac{1-x+2x^2-x^3}{x(x^2+1)^2}\,dx &= \int \left(\frac{1}{x} - \frac{x+1}{x^2+1} + \frac{x}{(x^2+1)^2}\right) dx \\ &= \int \left(\frac{1}{x} - \frac{x}{x^2+1} - \frac{1}{x^2+1} + \frac{x}{(x^2+1)^2}\right) dx \\ &= \log|x| - \frac{1}{2}\log(x^2+1) - \tan^{-1}x - \frac{1}{2(x^2+1)} + K. \end{aligned}$$

- Find

$$\int \frac{x}{(x+a)(x+b)}\,dx.$$

Try to write

$$\frac{x}{(x+a)(x+b)} = \frac{A}{x+a} + \frac{B}{x+b} = \frac{A(x+b)+B(x+a)}{(x+a)(x+b)}.$$

Thus

$$\begin{aligned} x^1 &: \quad 1 = A+B \Rightarrow B = 1-A, \\ x^0 &: \quad 0 = Ab+Ba. \end{aligned}$$

Solving for A and B, we find for $a \neq b$ that

$$\begin{aligned} 0 &= Ab + (1-A)a, \\ \Rightarrow A &= \frac{a}{a-b}, \\ B &= 1 - \frac{a}{a-b} = \frac{-b}{a-b}. \end{aligned}$$

$\therefore$ If $a \neq b$ then

$$\int \frac{x}{(x+a)(x+b)}\,dx = \int \frac{\frac{a}{a-b}}{x+a}\,dx - \int \frac{\frac{b}{a-b}}{x+b}\,dx$$
$$= \frac{a}{a-b}\log|x+a| - \frac{b}{a-b}\log|x+b| + C.$$

Problem: But what if $b = a$? Then

$$\begin{aligned} 1 &= A + B \\ 0 &= Aa + Ba = (A+B)a, \end{aligned}$$

which is consistent only if $a = b = 0$.

Remedy: Write

$$\frac{x}{(x+a)^2} = \frac{A}{x+a} + \frac{B}{(x+a)^2} = \frac{A(x+a)+B}{(x+a)^2}.$$

Then

$$\begin{aligned} x^1 &: \quad 1 = A \\ x^0 &: \quad 0 = Aa + B \Rightarrow B = -a. \end{aligned}$$

$$\therefore \int \frac{x}{(x+a)^2}\,dx = \int \left[\frac{1}{x+a} - \frac{a}{(x+a)^2}\right] dx$$
$$= \log|x+a| + \frac{a}{x+a} + C.$$

- Evaluate

$$\int \frac{x^2}{(x+1)^2}\,dx.$$

Since $\deg x^2 = 2 \geq \deg(x+1)^2 = 2$, we need to rewrite

$$\frac{x^2}{(x+1)^2} = 1 - \frac{(2x+1)}{(x+1)^2}.$$

Express

$$\frac{2x+1}{(x+1)^2} = \frac{A}{x+1} + \frac{B}{(x+1)^2};$$

this requires that $2x + 1 = A(x+1) + B$. On equating like powers of x, we find

$$\begin{aligned} x^1 &: \quad 2 = A, \\ x^0 &: \quad 1 = A + B \Rightarrow B = -1, \end{aligned}$$

so that

$$\begin{aligned}\int \frac{x^2}{(x+1)^2}\,dx &= \int \left(1 - \left[\frac{A}{x+1} + \frac{B}{(x+1)^2}\right]\right) dx\\ &= \int \left(1 - \left[\frac{2}{x+1} + \frac{-1}{(x+1)^2}\right]\right) dx\\ &= x - 2\log|x+1| - \frac{1}{x+1} + C.\end{aligned}$$

Remark: Show that the substitution $u = x+1$ makes the previous problem much easier!

• Find

$$\int \frac{1}{x^3-1}\,dx.$$

Step 1: We already have $\deg P < \deg Q$.

Step 2: Noting that $Q(x) = x^3 - 1$ has a root at $x = 1$, we factor

$$Q(x) = x^3 - 1 = (x-1)(x^2+x+1).$$

The quadratic factor x^2+x+1 cannot be factored into linear factors with real coefficients since it has no real roots (the discriminant $1^2 - 4 = -3$ is negative).

Step 3: Write

$$\frac{1}{x^3-1} = \frac{A}{x-1} + \frac{Bx+C}{x^2+x+1}.$$

Step 4: Then

$$\begin{aligned}1 &= A(x^2+x+1) + (Bx+C)(x-1)\\ &= Ax^2 + Ax + A + Bx^2 - Bx + Cx - C.\end{aligned}$$

We find

$$\begin{aligned}&x^2 : 0 = A + B \Rightarrow B = -A,\\ &x^1 : 0 = A - B + C,\\ &x^0 : 1 = A - C \Rightarrow C = A - 1.\end{aligned}$$

The x^1 equation then yields $0 = A + A + (A-1)$, which implies $A = \frac{1}{3}$, $B = -\frac{1}{3}$, and $C = -\frac{2}{3}$, so that

$$\int \frac{1}{x^3-1}\,dx = \int \frac{\frac{1}{3}}{x-1}\,dx + \int \frac{-\frac{1}{3}x - \frac{2}{3}}{x^2+x+1}\,dx.$$

Step 5: On completing the square and letting $u = x + \frac{1}{2}$, we evaluate

$$\begin{aligned}\int \frac{x+2}{x^2+x+1}\,dx &= \int \frac{x+2}{(x+\frac{1}{2})^2 - \frac{1}{4} + 1}\,dx \\ &= \int \frac{u+\frac{3}{2}}{u^2+\frac{3}{4}}\,du \\ &= \int \frac{u}{u^2+\frac{3}{4}}\,du + \frac{3}{2}\int \frac{1}{u^2+\frac{3}{4}}\,du \\ &= \frac{1}{2}\log\left|u^2+\frac{3}{4}\right| + \frac{3}{2}\frac{1}{\sqrt{\frac{3}{4}}}\arctan\left(\frac{u}{\sqrt{\frac{3}{4}}}\right) \\ &= \frac{1}{2}\log\left|x^2+x+1\right| + \sqrt{3}\arctan\left(\frac{2x+1}{\sqrt{3}}\right) + C.\end{aligned}$$

Thus

$$\int \frac{1}{x^3-1}\,dx = \frac{1}{3}\log|x-1| - \frac{1}{6}\log\left|x^2+x+1\right| - \frac{1}{\sqrt{3}}\arctan\left(\frac{2x+1}{\sqrt{3}}\right) + C.$$

Problem 7.9: Evaluate

$$\int \frac{1}{1-u^2}\,du.$$

$$\int \frac{1}{(1+u)(1-u)}\,du = \int \left(\frac{\frac{1}{2}}{1+u} + \frac{\frac{1}{2}}{1-u}\right)du = \frac{1}{2}\log\left|\frac{1+u}{1-u}\right| + C$$

Note: in the complex plane, the antiderivative may also be written as $\tanh^{-1} u + C$. However, in $\mathbb{R}$, the latter solution does not exist outside of $(-1, 1)$.

Problem 7.10: Compute

$$\int \frac{1}{(u^2+1)(u+1)}\,du.$$

Express

$$\frac{1}{(u^2+1)(u+1)} = \frac{A}{u+1} + \frac{Bu+C}{u^2+1}.$$

By equating coefficients of like powers in $1 = A(u^2+1) + B(u^2+u) + C(u+1)$, we obtain the system of equations

$$\begin{aligned} u^0 &: 1 = A + C, \\ u^1 &: 0 = B + C, \\ u^2 &: 0 = A + B, \end{aligned}$$

which has the unique solution $A = C = 1/2$, $B = -1/2$. Hence the integral becomes

$$\frac{1}{2}\int\left(\frac{1}{u+1}-\frac{u}{u^2+1}+\frac{1}{u^2+1}\right)du = \frac{1}{2}\log|u+1| - \frac{1}{4}\log(u^2+1) + \frac{1}{2}\arctan u + K,$$

where K is a constant.

Problem 7.11: Find

$$\int \frac{x^3+4x^2+7x+5}{(x+1)^2(x+2)}\,dx.$$

First, note that $(x+1)^2(x+2) = x^3+4x^2+5x+2$. The integral thus becomes

$$\int 1 + \frac{2x+3}{(x+1)^2(x+2)}\,du$$

On expressing

$$\frac{2x+3}{(x+1)^2(x+2)} = \frac{A}{x+1} + \frac{B}{(x+1)^2} + \frac{C}{x+2}$$

and equating coefficients of like powers in $2x+3 = A(x+1)(x+2)+B(x+2)+C(x+1)^2$, we obtain the system of equations

$$\begin{aligned} x^0 &: 3 = 2A + 2B + C,\\ x^1 &: 2 = 3A + B + 2C,\\ x^2 &: 0 = A + C, \end{aligned}$$

which has the unique solution $A = B = 1$, $C = -1$.

The integral thus evaluates to

$$x + \log\left|\frac{x+1}{x+2}\right| - \frac{1}{x+1} + K,$$

where K is an arbitrary constant.

Q. Can the method of partial fraction decomposition be used to reduce every rational function into terms of the types discussed in Step 5?

A. Yes, Theorem A.1 (the Fundamental Theorem of Algebra) and Corollary A.1.2 in Appendix A show that the factorization of the denominator in the form described in Step 2 is always possible. Furthermore, recursive application of the following lemmas can be used to show that the decomposition described in Steps 3 and 4 is always possible. Together, these results show that we can in principle integrate any rational function in terms of other rational functions, logarithms, and inverse tangents. Note that for polynomials $Q(x)$ of degree five and higher, we may have to resort to a numerical method for approximately locating the roots of the denominator.

Lemma 7.1 (Factor Theorem): *If z_0 is a root of a polynomial $P(z)$ then $P(z)$ is divisible by $(z - z_0)$.*

Proof: Suppose z_0 is a root of a polynomial $P(z) = \sum_{k=0}^{n} a_k z^k$ of degree n. Consider

$$
\begin{aligned}
P(z) = P(z) - P(z_0) &= \sum_{k=0}^{n} a_k z^k - \sum_{k=0}^{n} a_k z_0^k = \sum_{k=0}^{n} a_k (z^k - z_0^k) \\
&= \sum_{k=0}^{n} a_k (z - z_0)(z^{k-1} + z^{k-2} z_0 + \ldots + z_0^{k-1}) = (z - z_0) Q(z),
\end{aligned}
$$

where $Q(z) = \sum_{k=0}^{n} a_k (z^{k-1} + z^{k-2} z_0 + \ldots + z_0^{k-1})$ is a polynomial of degree $n-1$.

Lemma 7.2 (Linear Partial Fractions): *Suppose that $P(x)/Q(x)$ is a proper rational function such that $Q(x) = (x-a)^n Q_0(x)$, where $Q_0(a) \neq 0$ and $n \in \mathbb{N}$. Then there exists a constant A and a polynomial P_0 with $\deg P_0 < \deg Q - 1$ such that*

$$
\frac{P(x)}{Q(x)} = \frac{A}{(x-a)^n} + \frac{P_0(x)}{(x-a)^{n-1} Q_0(x)}.
$$

Proof: Let $A = P(a)/Q_0(a)$, so that the polynomial $P(x) - AQ_0(x)$ has the root a. Lemma 7.1 implies $P(x) - AQ_0(x) = (x-a)P_0(x)$ for some polynomial P_0, where $\deg P_0 \le \max(\deg P, \deg Q_0) - 1 < \deg Q - 1$. On dividing this result by $Q(x)$, we obtain

$$
\frac{P(x)}{Q(x)} - \frac{A}{(x-a)^n} = \frac{P_0(x)}{(x-a)^{n-1} Q_0(x)}.
$$

Lemma 7.3 (Quadratic Partial Fractions): *Let $x^2 + \gamma x + \lambda$ be an irreducible quadratic polynomial (i.e. $\gamma^2 - 4\lambda < 0$). Suppose that $P(x)/Q(x)$ is a proper rational function such that $Q(x) = (x^2 + \gamma x + \lambda)^m Q_0(x)$, where $Q_0(x)$ is not divisible by $(x^2 + \gamma x + \lambda)$ and $m \in \mathbb{N}$. Then there exists constants Γ and Λ and a polynomial P_0 with $\deg P_0 < \deg Q - 2$ such that*

$$
\frac{P(x)}{Q(x)} = \frac{\Gamma x + \Lambda}{(x^2 + \gamma x + \lambda)^m} + \frac{P_0(x)}{(x^2 + \gamma x + \lambda)^{m-1} Q_0(x)}.
$$

Proof: Use long division to express

$$
\begin{aligned}
P(x) &= (x^2 + \gamma x + \lambda) S(x) + (ax + b), \\
Q_0(x) &= (x^2 + \gamma x + \lambda) T(x) + (cx + d),
\end{aligned}
$$

where S and T are polynomials. We want to find Γ and Λ such that

$$
P(x) - (\Gamma x + \Lambda) Q_0(x) = P_0(x)(x^2 + \gamma x + \lambda),
$$

i.e., we want $ax+b-(\Gamma x+\Lambda)(cx+d)$ to be divisible by $(x^2+\gamma x+\lambda)$. Using long division, we then see that the remainder $[a+(c\gamma-d)\Gamma-\Lambda c]x+(b+\Gamma c\lambda-\Lambda d)$ must be zero. We thus need to satisfy the simultaneous equations

$$\begin{aligned}(c\gamma-d)\Gamma-\Lambda c&=-a,\\ \Gamma c\lambda-\Lambda d&=-b.\end{aligned}$$

We can write this condition as the matrix equation

$$\begin{bmatrix} c\gamma-d & -c \\ c\lambda & -d \end{bmatrix}\begin{bmatrix}\Gamma\\ \Lambda\end{bmatrix}=\begin{bmatrix}-a\\ -b\end{bmatrix}. \tag{7.4}$$

Let

$$\Delta=\det\begin{bmatrix} c\gamma-d & -c \\ c\lambda & -d \end{bmatrix}=d^2-c\gamma d+c^2\lambda.$$

If $c\neq 0$, then $\Delta=c^2\left[\left(\frac{-d}{c}\right)^2+\gamma\left(\frac{-d}{c}\right)+\lambda\right]\neq 0$ since $x^2+\gamma x+\lambda\neq 0$ for all x. If $c=0$, then $\Delta=d^2\neq 0$ since the fact that $Q_0(x)$ is not divisible by $(x^2+\gamma x+\lambda)$ implies that c and d cannot both be 0. In either case, we see that $\Delta\neq 0$, so we can always solve Eq. (7.4) for values of Γ and Λ such that Eq. (7.3) will hold. Note that $\deg P_0\leq\max(\deg P,\deg Q_0+1)-2<\deg Q-2$.

7.E Trigonometric & Hyperbolic Substitution

Trigonometric and hyperbolic trigonometric substitutions are often useful for evaluating integrals containing square roots of quadratic expressions. Several common trigonometric substitutions for frequently appearing quadratic expressions are listed in Table 7.2. Note that it may be necessary to complete the square of the quadratic and shift the variable of integration to put the expression into one of these forms.

- We can use the trigonometric substitution $x=a\sin\theta$ to find the area between the half-circle $y=\sqrt{a^2-x^2}$ and $y=0$:

  $$\int_{-a}^{a}\sqrt{a^2-x^2}\,dx=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}a\cos\theta\,a\cos\theta\,d\theta=a^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{1+\cos 2\theta}{2}\,d\theta=\frac{\pi}{2}a^2.$$

  Note that this result agrees with our original definition of π as the area of the unit circle.

- We can find the indefinite integral

  $$\int\sqrt{u^2+a^2}\,du$$

Expression	x Domain	Substitution	θ or t Domain	Identity
$\sqrt{a^2 - x^2}$	$[-a, a]$	$x = a\sin\theta$	$\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$	$1 - \sin^2\theta = \cos^2\theta$
$\sqrt{a^2 + x^2}$	$(-\infty, \infty)$	$x = a\tan\theta$ $x = a\sinh t$	$\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ $(-\infty, \infty)$	$1 + \tan^2\theta = \sec^2\theta$ $1 + \sinh^2 t = \cosh^2 t$
$\sqrt{x^2 - a^2}$	$(-\infty, -a] \cup [a, \infty]$	$x = a\sec\theta$ $x = a\cosh t$	$\left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right]$ $[0, \infty)$	$\sec^2\theta - 1 = \tan^2\theta$ $\cosh^2 t - 1 = \sinh^2 t$

Table 7.2: Useful trigonometric substitutions.

with the substitution $u = a\tan\theta$ ($du = a\sec^2\theta\, d\theta$). Without loss of generality, we assume that $a > 0$. Since

$$u^2 + a^2 = a^2(\tan^2\theta + 1) = a^2\sec^2\theta,$$

we may write

$$\begin{aligned}\int \sqrt{u^2 + a^2}\, du &= \int a\sec\theta\, a\sec^2\theta\, d\theta = a^2 \int \frac{1}{\cos^3\theta}\, d\theta \\ &= a^2 \int \frac{\cos\theta}{\cos^4\theta}\, d\theta = a^2 \int \frac{\cos\theta}{(1 - \sin^2\theta)^2}\, d\theta \\ &= a^2 \int \frac{dw}{(1 - w^2)^2} = a^2 \int \frac{dw}{(1 - w)^2(1 + w)^2},\end{aligned}$$

upon making the substitution $w = \sin\theta$. We have thus reduced the integrand to a rational function that can be integrated by the method of partial fractions. Alternatively, the integral of $1/\cos^3 x$ we obtained above could have be evaluated by an integration by parts, as we did on page 170.

Problem 7.12: Evaluate

$$\int \sqrt{u^2 + a^2}\, du$$

instead with the substitution $u = a\sinh t$ and Problem 6.10.

- For $0 < x \leq 3$, the substitution $x = 3\sin\theta$ $(dx = 3\cos\theta\, d\theta)$ and the fact that $9 - x^2 = 9\cos^2\theta$ can be used to evaluate the integral

$$\int \frac{\sqrt{9-x^2}}{x^2}\,dx = \int \frac{3\cos\theta}{9\sin^2\theta} 3\cos\theta\, d\theta = \int \frac{1-\sin^2\theta}{\sin^2\theta}\,d\theta = \int \left(\frac{1}{\sin^2\theta} - 1\right) d\theta$$
$$= -\cot\theta - \theta + C = -\sqrt{\frac{9}{x^2} - 1} - \sin^{-1}\frac{x}{3} + C.$$

Problem 7.13: Find the antiderivative

$$\int \frac{\sqrt{9-x^2}}{x^2}\,dx$$

for $-3 \leq x < 0$.

- The hyperbolic substitution $x = a\cosh t$ $(dx = a\sinh t\, dt)$ and the fact that $x^2 - a^2 = a^2\sinh^2 t$ allow us to evaluate

$$\int \frac{dx}{\sqrt{x^2-a^2}} = \int \frac{1}{a\sinh t} a\sinh t\, dt = \int dt = t + C = \cosh^{-1}\frac{x}{a} + C.$$

Problem 7.14: Show that the substitution $x = a\sec\theta$ can also be used to find

$$\int \frac{dx}{\sqrt{x^2-a^2}}$$

and that the answer equals $\cosh^{-1}\frac{x}{a} + C$.

- An integral of the form

$$I = \int \frac{x^3}{(4x^2+9)^{\frac{3}{2}}}\,dx$$

can first be put in the form of the expressions listed in Table 7.2 with the substitution $u = 2x$, so that $4x^2+9 = u^2+9$. One could then apply the substitution $u = 3\sinh t$. In fact, both substitutions can be done in a single step by defining $x = \frac{3}{2}\sinh t$:

$$I = \int \left(\frac{\left(\frac{\cancel{3}}{2}\right)^3 \sinh^3 t}{\cancel{3}^{\cancel{3}} \cosh^3 t}\right)\frac{3}{2}\cosh t\, dt = \frac{3}{16}\int \frac{\sinh^3 t}{\cosh^2 t}\,dt = \frac{3}{16}\int \frac{\sinh t(\cosh^2 t - 1)}{\cosh^2 t}\,dt$$
$$= \frac{3}{16}\int \left(\sinh t - \frac{\sinh t}{\cosh^2 t}\right) dt = \frac{3}{16}\left(\cosh t + \frac{1}{\cosh t}\right) + C$$
$$= \frac{1}{16}\left(\sqrt{4x^2+9} + \frac{9}{\sqrt{4x^2+9}}\right) + C.$$

This integral could have also been evaluated with the substitution $x = \frac{3}{2}\tan\theta$.

7.F Integration of Certain Irrational Expressions

Q. How do we find integrals like
$$\int \frac{\sqrt{x+4}}{x}\,dx?$$

A. Substitute $t = \sqrt{x+4}$. Then $t^2 = x+4 \Rightarrow 2t\,dt = dx$, and
$$\begin{aligned}
\int \frac{\sqrt{x+4}}{x}\,dx &= \int \left(\frac{t}{t^2-4}\right) 2t\,dt \\
&= 2\int \frac{t^2}{t^2-4}\,dt = 2\int \left[\frac{t^2-4}{t^2-4} + \frac{4}{t^2-4}\right] dt \\
&= 2t + 8\int \frac{1}{t^2-4}\,dt \\
&= 2t + 8\int \left[\frac{A}{t-2} + \frac{B}{t+2}\right] dt \qquad \begin{cases} 1 = A(t+2) + B(t-2), \\ 0 = A + B \Rightarrow B = -A, \\ 1 = 2A - 2B = 4A \Rightarrow A = \frac{1}{4}. \end{cases} \\
&= 2t + 8\int \left[\frac{\frac{1}{4}}{t-2} - \frac{\frac{1}{4}}{t+2}\right] dt \\
&= 2t + 2\log\left|\frac{t-2}{t+2}\right| + C.
\end{aligned}$$

Thus
$$\int \frac{\sqrt{x+4}}{x}\,dx = 2\sqrt{x+4} + 2\log\left|\frac{\sqrt{x+4}-2}{\sqrt{x+4}+2}\right| + C.$$

In general, one can reduce any integral of the form
$$\int R\left(x,\ \sqrt[m]{\frac{ax+b}{cx+d}}\right) dx,$$
where R is a birational function of its arguments, to the integral of a rational function by using the substitution
$$t = \sqrt[m]{\frac{ax+b}{cx+d}}.$$

- We can evaluate
$$\int \frac{1}{x - \sqrt{x+2}}\,dx$$
with the substitution $t = \sqrt{x+2} \Rightarrow t^2 = x+2 \Rightarrow 2t\,dt = dx$,
$$\begin{aligned}
\int \frac{1}{x-\sqrt{x+2}}\,dx &= \int \left(\frac{1}{t^2-2-t}\right) 2t\,dt = 2\int \frac{t}{t^2-t-2}\,dt \\
&= 2\int \frac{t}{(t-2)(t+1)}\,dt,
\end{aligned}$$

which can then be decomposed into partial fractions.

Remark: When more than one radical appears, it is often helpful to take m to be the least common multiple of the radical indices.

Problem 7.15: Find

$$\int \frac{1}{\sqrt[2]{x} - \sqrt[3]{x}}\, dx$$

using the substitution $t = x^{\frac{1}{2\cdot 3}} = x^{\frac{1}{6}}$.

Problem 7.16: Find

$$\int \frac{1}{\sqrt[6]{x} + \sqrt[4]{x}}\, dx.$$

using the substitution $t = x^{\frac{1}{12}}$.

Q. How about integrals of the form $\int \sqrt{x^2 + x + 1}\, dx$?

A. We can first simplify the integrand somewhat by completing the square and making the substitution $u = x + 1/2$:

$$\int \sqrt{\left(x + \frac{1}{2}\right)^2 - \frac{1}{4} + 1}\; dx = \int \sqrt{u^2 + \frac{3}{4}}\, du.$$

The resulting integral is of the form $\int \sqrt{u^2 + a^2}\, du$, which we computed in section 7.E.

Remark: Integrals of the form

$$\int R(x, \sqrt{ax^2 + bx + c})\, dx,$$

where R is a birational function of its arguments, can be calculated with the aid of one of the three *Euler substitutions*

(i) $\sqrt{ax^2 + bx + c} = t \pm x\sqrt{a}$ if $a > 0$;

(ii) $\sqrt{ax^2 + bx + c} = tx \pm \sqrt{c}$ if $c > 0$;

(iii) $\sqrt{ax^2 + bx + c} = (x - \alpha)t$ if α is a real root of $ax^2 + bx + c = 0$.

(Either sign may be taken in the first two cases.)

- To find

$$I = \int \frac{dx}{1 + \sqrt{x^2 + 2x + 2}},$$

we can use the first Euler substitution since $a = 1 > 0$. Let

$$\sqrt{x^2 + 2x + 2} = t - x.$$

Then $x^2 + 2x + 2 = t^2 - 2tx + x^2$, so that

$$x = \frac{t^2 - 2}{2(t+1)}, \qquad dx = \frac{2t(t+1) - (t^2 - 2)}{2(t+1)^2}\, dt = \frac{t^2 + 2t + 2}{2(t+1)^2}\, dt.$$

The integral becomes

$$\int \frac{t^2 + 2t + 2}{\left(1 + t - \frac{t^2-2}{2(t+1)}\right) 2(t+1)^2}\, dt = \int \frac{t^2 + 2t + 2}{(t^2 + 4t + 4)(t+1)}\, dt = \int \frac{t^2 + 2t + 2}{(t+2)^2(t+1)}\, dt,$$

which we can then solve with the method of partial fractions. Decomposing

$$\frac{t^2 + 2t + 2}{(t+1)(t+2)^2} = \frac{A}{t+1} + \frac{B}{t+2} + \frac{C}{(t+2)^2},$$

we see that

$$t^2+2t+2 = A(t+2)^2+B(t+1)(t+2)+C(t+1) = A(t^2+4t+4)+B(t^2+3t+2)+C(t+1).$$

The solution to the resulting system of equations,

$$\begin{aligned} &t^0 : 2 = 4A + 2B + C, \\ &t^1 : 2 = 4A + 3B + C \Rightarrow B = 0, \\ &t^2 : 1 = A + B \Rightarrow A = 1, \end{aligned}$$

is $A = 1$, $B = 0$, and $C = -2$. Thus

$$\begin{aligned} I &= \int \frac{dt}{t+1} - 2\int \frac{dt}{(t+2)^2} = \log|t+1| + \frac{2}{t+2} + C \\ &= \log\left|x + 1 + \sqrt{x^2 + 2x + 2}\right| + \frac{2}{x + 2 + \sqrt{x^2 + 2x + 2}} + C. \end{aligned}$$

7.G Strategy for Integration

1. Simplify the integrand.

2. Look for an obvious substitution: see if you can write the integral in the form

$$\int f(g(x))g'(x)\,dx.$$

If so, try the substitution $u = g(x)$.

3. Classify the integrand.

(a) Trigonometric functions: exploit trigonometric identities to find integrals of the form

$$\left\{ \begin{array}{l} \int \sin^n x \cos^m x\,dx \\ \int \tan^n x \sec^m x\,dx \\ \int \cot^n x \csc^m x\,dx \end{array} \right\}.$$

As a last resort, use the universal substitution $t = \tan\frac{x}{2}$.

(b) Rational functions: use the Method of Partial Fractions.

(c) Polynomials (including 1) $\times$ *transcendental functions* (e.g. Trigonometric, exponential, logarithmic, and inverse functions): use Integration by Parts.

(d) Radicals:

(i) $\sqrt{\pm x^2 \pm a^2}$: use a trigonometric substitution

(ii) $\sqrt[n]{\dfrac{ax+b}{cx+d}}$: $t = \sqrt[n]{\dfrac{ax+b}{cx+d}}$

For $\sqrt[n]{g(x)}$: $t = \sqrt[n]{g(x)}$ sometimes helps.

4. Try again (maybe use several methods combined).

Problem 7.17: (a) Find

$$\int \frac{x}{\sqrt{1+x^{2/3}}}\,dx.$$

Substituting first $y = x^{2/3}$ and then $t = y + 1$ we find

$$\begin{aligned}
\int \frac{x}{\sqrt{1+x^{2/3}}}\,dx &= \int \frac{y^{3/2}}{\sqrt{1+y}}\frac{3}{2}y^{1/2}\,dy = \frac{3}{2}\int \frac{y^2}{\sqrt{1+y}}\,dy = \frac{3}{2}\int \frac{(t-1)^2}{\sqrt{t}}\,dt \\
&= \frac{3}{2}\int t^{3/2} - 2t^{1/2} + t^{-1/2}\,dt = \frac{3}{2}\left(\frac{2}{5}t^{5/2} - \frac{4}{3}t^{3/2} + 2t^{1/2}\right) + C \\
&= \frac{3}{5}\left(1+x^{2/3}\right)^{5/2} - 2\left(1+x^{2/3}\right)^{3/2} + 3\left(1+x^{2/3}\right)^{1/2} + C \\
&= \frac{1}{5}\left(1+x^{2/3}\right)^{1/2}\left[3\left(1+x^{2/3}\right)^2 - 10\left(1+x^{2/3}\right) + 15\right] + C. \\
&= \frac{1}{5}\left(1+x^{2/3}\right)^{1/2}\left(8 - 4x^{2/3} + 3x^{4/3}\right) + C.
\end{aligned}$$

Alternatively, substituting $y = x^{1/3}$, then $\sinh t = y$, and finally $u = \cosh t = \sqrt{1+y^2} = \sqrt{1+x^{2/3}}$ (which could be used as a more direct substitution), we find

$$\begin{aligned}\int \frac{x}{\sqrt{1+x^{2/3}}}\,dx &= \int \frac{y^3}{\sqrt{1+y^2}}\,3y^2\,dy = 3\int \frac{y^5}{\sqrt{1+y^2}}\,dy = 3\int \frac{\sinh^5 t}{\cosh t}\,\cosh t\,dt \\ &= 3\int (u^2-1)^2\,du = 3\int (u^4 - 2u^2 + 1)\,du \\ &= 3\Big(\frac{u^5}{5} - 2\frac{u^3}{3} + u\Big) + C \\ &= \frac{u}{5}\big(3u^4 - 10u^2 + 15\big) + C \\ &= \frac{1}{5}\Big(1 + x^{2/3}\Big)^{1/2}\Big(8 - 4x^{2/3} + 3x^{4/3}\Big) + C.\end{aligned}$$

(b) Find

$$\int_1^4 \sqrt{x}\cos\sqrt{x}\,dx.$$

Let $u = \sqrt{x}$, so that $du = dx/(2\sqrt{x}) = dx/(2u)$. The integral becomes

$$2\int_1^2 u^2\cos u\,du.$$

Integrating by parts, we first compute the indefinite integral

$$\begin{aligned}\int u^2\cos u\,du = u^2\sin u - 2\int u\sin u\,du &= u^2\sin u - 2\Big(-u\cos u + \int \cos u\,du\Big) \\ &= u^2\sin u + 2u\cos u - 2\sin u + C.\end{aligned}$$

Thus the original integral evaluates to

$$2\big[(u^2-2)\sin u + 2u\cos u\big]_1^2 = 4\sin 2 + 8\cos 2 + 2\sin 1 - 4\cos 1.$$

7.H Numerical Approximation of Integrals

There are many continuous functions such as

$$\frac{e^x}{x},\ \frac{\sin x}{x},\ \text{ and } e^{x^2},$$

for which the antiderivative cannot be expressed in terms of the elementary functions introduced so far. For applications where one needs only the value of a definite integral, one possibility is to approximate the integral numerically.

To illustrate the numerical evaluation of definite integrals, it is helpful to consider an integral for which we know the exact answer, such as $\int_0^1 f$, where $f(x) = x^2$. For the partition $P = \{0, \frac{1}{2}, 1\}$ we find

$$\mathcal{L}(P, f) = 0\left(\frac{1}{2}\right) + \frac{1}{4}\left(\frac{1}{2}\right) = \frac{1}{8} = 0.125,$$

$$\mathcal{U}(P, f) = \frac{1}{4}\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right) = \frac{5}{8} = 0.625.$$

We know that

$$\mathcal{L}(P, f) \le \int_0^1 x^2\, dx \le \mathcal{U}(P, f),$$

but neither $\mathcal{L}$ nor $\mathcal{U}$ provides us with a very good approximation to the integral. Notice that the average of $\mathcal{L}$ and $\mathcal{U}$, namely $(\mathcal{L} + \mathcal{U})/2 = 3/8 = 0.375$, is much closer to the exact value (1/3) of the definite integral and that since f is increasing, $\mathcal{L}(P, f)$ is identical to the *left Riemann sum* $\mathcal{S}_L(P, f) = \sum_{i=1}^2 f(x_{i-1})(x_i - x_{i-1})$ and $\mathcal{U}(P, f)$ is the *right Riemann sum* $\mathcal{S}_R(P, f) = \sum_{i=1}^2 f(x_i)(x_i - x_{i-1})$. This suggests that it may be better to approximate the integral by using the *Trapezoidal Rule*

$$\mathcal{T}(P_n, f) \doteq \frac{\mathcal{S}_L(P_n, f) + \mathcal{S}_R(P_n, f)}{2} = \sum_{i=1}^n \frac{f(x_{i-1}) + f(x_i)}{2}(x_i - x_{i-1}),$$

Remark: For a uniform partition with fixed a, b, and f, $\mathcal{T}(P_n, f)$ depends only on the number n of points in the partition P_n.

Q. How accurately does $\mathcal{T}(P_n, f)$, where P_n is a uniform partition of $[a, b]$ into n subintervals, approximate $\int_a^b f$? How does the error depend on n?

A. First, we look at a special case of this question where there is only one subinterval.

Theorem 7.3 (Linear Interpolation Error): *Let f be a twice-differentiable function on $[0, h]$ satisfying $|f''(x)| \le M$ for all $x \in [0, h]$. Let*

$$L(x) = f(0) + \frac{f(h) - f(0)}{h} x.$$

Then

$$\int_0^h |L(x) - f(x)|\, dx \le \frac{Mh^3}{12}.$$

Proof: Let $x \in (0,h)$ and

$$\varphi(t) = L(t) - f(t) - Ct(t-h),$$

where C is chosen so that $\varphi(x) = 0$. Then

$$\varphi(0) = L(0) - f(0) = 0,$$

$$\varphi(h) = L(h) - f(h) = 0.$$

From Rolle's Theorem, we then know that there exists $x_1 \in (0,x)$ and $x_2 \in (x,h)$ such that

$$\varphi'(x_1) = \varphi'(x_2) = 0.$$

Again by Rolle's Theorem, we know that there exists $c \in (x_1,x_2)$ such that

$$0 = \varphi''(c) = -f''(c) - 2C,$$

noting that L is linear. Therefore $C = -f''(c)/2$ and since $\varphi(x) = 0$,

$$L(x) - f(x) = \frac{-1}{2} f''(c)x(x-h),$$

where $c \in (0,h)$ depends on x. That is, for every $x \in [0,h]$ we have

$$|L(x) - f(x)| \le \frac{1}{2} Mx(h-x),$$

so

$$\int_0^h |L(x) - f(x)|\, dx \le \frac{M}{2}\int_0^h x(h-x)\, dx = \frac{M}{2}\left[\frac{x^2 h}{2} - \frac{x^3}{3}\right]_0^h = \frac{Mh^3}{12}.$$

Corollary 7.3.1 (Trapezoidal Rule Error): Let P be a uniform partition of $[a,b]$ into n subintervals of width $h = (b-a)/n$, and f be a twice-differentiable function on $[a,b]$ satisfying $|f''(x)| \le M$ for all $x \in [a,b]$. Then the error $E_n^{\mathcal{T}} \doteq \mathcal{T}(P_n, f) - \int_a^b f$ of the uniform Trapezoidal Rule

$$\mathcal{T}(P_n, f) = h\sum_{i=1}^{n} \frac{f(x_{i-1}) + f(x_i)}{2}$$

satisfies

$$\left|E_n^{\mathcal{T}}\right| \le \frac{nMh^3}{12} = \frac{M(b-a)^3}{12n^2}.$$

Proof: We need to add up the contribution to the error from each subinterval. If we temporarily relabel the endpoints of each subinterval 0 and h, we may apply Theorem 7.3 to obtain a contribution, $\left|\int_0^h L - \int_0^h f\right| \le \int_0^h |L - f| \le Mh^3/12$, from each of the n subintervals.

Remark: We can rewrite the Trapezoidal Rule as

$$\mathcal{T}(P_n, f) = \frac{h}{2}[f(x_0) + 2f(x_1) + \ldots + 2f(x_{n-1}) + f(x_n)].$$

- We can use the Trapezoidal Rule to approximate $\int_1^2 \frac{1}{x}\,dx$ with $n = 5$ subintervals of width $h = 1/5$:

$$\begin{aligned}\int_1^2 \frac{1}{x}\,dx \approx \mathcal{T}(P_n, f) &= \frac{1}{10}\left[\frac{1}{1} + \frac{2}{1.2} + \frac{2}{1.4} + \frac{2}{1.6} + \frac{2}{1.8} + \frac{1}{2}\right] \\ &\approx 0.6956.\end{aligned}$$

The exact value of the integral is $\log 2 = 0.6931\ldots$.

Remark: Typically, a more accurate method than the Trapezoidal Rule is the *Midpoint Rule*

$$\mathcal{M}(P_n, f) = \sum_{i=1}^{n} f\left(\frac{x_{i-1} + x_i}{2}\right)(x_i - x_{i-1}),$$

which has the additional advantage of requiring one less function evaluation.

Problem 7.18: Show that the Midpoint Rule has an error $E_n^{\mathcal{M}} \doteq \mathcal{M}(P_n, f) - \int_a^b f$ satisfying

$$\left|E_n^{\mathcal{M}}\right| \leq \frac{M(b-a)^3}{24n^2}.$$

Notice that this bound is a factor of 2 smaller than the error bound for the Trapezoidal Rule.

- Let us use the Midpoint Rule to approximate $\int_1^2 \frac{1}{x}\,dx$ with $n = 5$ subintervals of width $h = 1/5$:

$$\begin{aligned}\int_1^2 \frac{1}{x}\,dx \approx \mathcal{M}(P_n, f) &= \frac{1}{5}\left[\frac{1}{1.1} + \frac{1}{1.3} + \frac{1}{1.5} + \frac{1}{1.7} + \frac{1}{1.9}\right] \\ &\approx 0.6919,\end{aligned}$$

which is indeed closer than $\mathcal{T}(P_n, f)$ to the exact value of $\log 2$ (by roughly a factor of 2).

Remark: Even better are the higher-order methods, such as *Simpson's Rule*, which fits parabolas rather than line segments to the data values $f(x_0), f(x_1), \ldots, f(x_n)$, where n is even. This approximation is given by

$$\mathcal{S}(P_n, f) = \frac{h}{3}[f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+\ldots+2f(x_{n-2})+4f(x_{n-1})+f(x_n)],$$

with an error $E_n^{\mathcal{S}} \doteq \mathcal{S}(P_n, f) - \int_a^b f$ satisfying (cf. [Muldowney 1990, p. 277] with $n \to 2n$)

$$\left|E_n^{\mathcal{S}}\right| \le \frac{K(b-a)^5}{180n^4} \quad \text{if } \left|f^{(4)}(x)\right| \le K \ \forall x \in [a,b].$$

The formula for $\mathcal{S}(P_n, f)$ is obtained by dividing $[a,b]$ into n uniform subintervals of width $h = (b-a)/n$ with endoints $x_k = a + kh$, where $k = 0, \ldots, n$. In each fixed interval $[x_{2k}, x_{2k+2}]$, where $k \in \{0, \ldots, \frac{n}{2} - 1\}$, we fit a parabola of the form $y = A(x - x_{2k+1})^2 + B(x - x_{2k+1}) + C$ to the data values y_{2k}, y_{2k+1}, and y_{2k+2}:

$$\begin{aligned} y_{2k} &= Ah^2 - Bh + C, \\ y_{2k+1} &= C, \\ y_{2k+2} &= Ah^2 + Bh + C. \end{aligned}$$

On letting $u = x - x_{2k+1}$, the area under the curve in this subinterval can then be expressed as

$$\begin{aligned} \int_{-h}^{h} (Au^2 + Bu + C)\, du &= 2\int_0^h (Au^2 + C)\, du = 2\left[\frac{Au^3}{3} + Cu\right]_0^h \\ &= \frac{2Ah^3}{3} + 2Ch = \frac{h}{3}(2Ah^2 + 6C) = \frac{h}{3}(y_{2k} + 4y_{2k+1} + y_{2k+2}). \end{aligned}$$

Thus

$$\begin{aligned} \int_a^b f(x)\, dx &\approx \sum_{k=0}^{\frac{n}{2}-1} \frac{h}{3}(y_{2k} + 4y_{2k+1} + y_{2k+2}) \\ &= \frac{h}{3}(y_0 + 4y_1 + y_2) + \frac{h}{3}(y_2 + 4y_3 + y_4) + \ldots + \frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n) \\ &= \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + \ldots + 2y_{n-2} + 4y_{n-1} + y_n), \end{aligned}$$

so that

$$\mathcal{S}(P_n, f) = \frac{h}{3}[f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+\ldots+2f(x_{n-2})+4f(x_{n-1})+f(x_n)].$$

- Let $f(x) = 1/x$ and partition $[1, 2]$ into $n = 4$ subintervals. Then

$$\log 2 = \int_1^2 \frac{1}{x}\,dx \approx \mathcal{S}(P_4, f) \equiv \frac{1}{3}\left(\frac{1}{4}\right)\left[f(1) + 4f\left(\frac{5}{4}\right) + 2f\left(\frac{3}{2}\right) + 4f\left(\frac{7}{4}\right) + f(2)\right]$$
$$= \frac{1}{12}\left[1 + \frac{16}{5} + \frac{4}{3} + \frac{16}{7} + \frac{1}{2}\right] \approx 0.6932,$$

is very close to the exact value of $\log 2$. In fact, since

$$\left|\frac{d^4}{dx^4}\left(\frac{1}{x}\right)\right| = \left|\frac{24}{x^5}\right| \le 24 \quad \forall x \in [1, 2],$$

the error $E_n^{\mathcal{S}}$ in the above result can be no more than $24/(180 \cdot 4^4) = 0.0005$.

Problem 7.19: For uniform partitions, show that

$$\mathcal{S}(P_{2n}, f) = \frac{2}{3}(\mathcal{T}(P_n, f) + 2\mathcal{M}(P_n, f)).$$

Problem 7.20: Consider the function $f(x) = 1/(1+x^2)$ on $[0, 1]$. Let P be a uniform partition on $[0, 1]$ with 2 subintervals of equal width.

(a) Compute the lower sum $\mathcal{L}(P, f)$.
Since the partition is uniform,

$$\mathcal{L}(P, f) = \frac{1}{2}\left(\frac{4}{5} + \frac{1}{2}\right) = \frac{13}{20}.$$

(b) Compute the upper sum $\mathcal{U}(P, f)$.

$$\mathcal{U}(P, f) = \frac{1}{2}\left(1 + \frac{4}{5}\right) = \frac{9}{10}.$$

(c) Use your results in part (a) and (b) to find lower and upper bounds for π.
We see that

$$\frac{13}{20} = \mathcal{L}(P, f) \le \frac{\pi}{4} = \int_0^1 \frac{1}{1+x^2}\,dx \le \mathcal{U}(P, f) = \frac{9}{10}.$$

Thus $\frac{13}{5} \le \pi \le \frac{18}{5}$.

(d) Use the Trapezoidal Rule to find a numerical estimate for π.
We find that π is approximately

$$4\left(\frac{1}{2}\right)\left(\frac{1+\frac{4}{5}}{2} + \frac{\frac{4}{5}+\frac{1}{2}}{2}\right) = 2\left(\frac{9}{10} + \frac{13}{20}\right) = \frac{31}{10}.$$

(e) Obtain a better rational estimate for π by using the Midpoint Rule.
We find that π is approximately

$$4\left(\frac{1}{2}\right)\left(f\left(\frac{1}{4}\right)+f\left(\frac{3}{4}\right)\right)=2\left(\frac{16}{17}+\frac{16}{25}\right)=32\left(\frac{1}{17}+\frac{1}{25}\right)=32\left(\frac{42}{425}\right)=\frac{1344}{425}.$$

Chapter 8

Applications of Integration

8.A Areas between Curves

The area A between two continuous functions $y = f(x)$ and $y = g(x)$ on $[a, b]$, where $f(x) \geq g(x)$, is given by the difference of the respective areas between these functions and the x axis:

$$A = \int_a^b f(x) - \int_a^b g(x)\, dx = \int_a^b [f(x) - g(x)]\, dx.$$

- Find the area bounded by $f(x) = x^2 + 1$ and $g(x) = x$ between $x = 0$ and $x = 1$.

$$\begin{aligned} A &= \int_0^1 [f(x) - g(x)]\, dx = \int_0^1 [x^2 + 1 - x]\, dx \\ &= \left[\frac{x^3}{3} + x - \frac{x^2}{2}\right]_0^1 = \frac{1}{3} + 1 - \frac{1}{2} = \frac{5}{6}. \end{aligned}$$

  Sometimes we are not given a and b, but we can determine them from the points of intersections of the two curves.

- Find the area enclosed by the curves $f(x) = 2x - x^2$ and $g(x) = x^2$. Here a and b are determined by the points of intersection of $f(x)$ and $g(x)$,

$$\begin{aligned} f(x) &= g(x) \\ x^2 &= 2x - x^2 \\ \Rightarrow 2x^2 &= 2x \Rightarrow 2x(x-1) = 0 \\ &\Rightarrow x = 0 \text{ or } x = 1. \end{aligned}$$

Thus

$$A = \int_0^1 [f(x) - g(x)]\,dx = \int_0^1 (2x - x^2 - x^2)\,dx = \int_0^1 (2x - 2x^2)\,dx$$
$$= 2\int_0^1 (x - x^2)\,dx = 2\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = 2\left(\frac{1}{2} - \frac{1}{3}\right) = \frac{1}{3}.$$

Q. What happens when $f(x) \geq g(x)$ for some values of x but $g(x) \geq f(x)$ for other values?

A. We simply take the absolute value of the integrand before integrating. That is, the general formula for the area A of the region bounded by two continuous functions f and g and the vertical lines $x = a$ and $x = b$ is

$$A = \int_a^b |f(x) - g(x)|\,dx,$$

where

$$|f(x) - g(x)| = \begin{cases} f(x) - g(x) & \text{when } f(x) \geq g(x), \\ g(x) - f(x) & \text{when } f(x) < g(x). \end{cases}$$

For continuous functions f and g, the regions where $f(x) > g(x)$ and $f(x) < g(x)$ are separated by the points where $f(x) = g(x)$.

- Find the area bounded by the curves $y = \sin x$, $y = \cos x$, $x = 0$ and $x = \pi/2$. The intersection points occur when

$$\sin x = \cos x \Rightarrow \tan x = 1 \Rightarrow x = \frac{\pi}{4}.$$

We need to split the integration interval $[0, \pi/2]$ into two parts:

$$A = \int_0^{\frac{\pi}{2}} |\cos x - \sin x|\,dx = \int_0^{\frac{\pi}{4}} (\cos x - \sin x)\,dx + \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} (\sin x - \cos x)\,dx$$
$$= [\sin x + \cos x]_0^{\frac{\pi}{4}} + [-\cos x - \sin x]_{\frac{\pi}{4}}^{\frac{\pi}{2}}$$
$$= \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - 0 - 1 - 0 - 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = 2\sqrt{2} - 2.$$

- To find the area of the region bounded by $f(x) = x$ and $g(x) = x^3$, we first solve for the intersection points:

$$\begin{aligned}
&f(x) = g(x)\\
&\Rightarrow x = x^3\\
&\Rightarrow 0 = x^3 - x = x(x^2 - 1) = x(x-1)(x+1)\\
&\Rightarrow x = -1, 0, 1.
\end{aligned}$$

On $[-1, 0]$ we see that $f(x) \leq g(x)$ and on $[0, 1]$ we see that $f(x) \geq g(x)$. Thus

$$\begin{aligned}
A &= \int_{-1}^{1} |f(x) - g(x)|\, dx = \int_{-1}^{0} [g(x) - f(x)]\, dx + \int_{0}^{1} [f(x) - g(x)]\, dx\\
&= \int_{-1}^{0} (x^3 - x)\, dx + \int_{0}^{1} (x - x^3)\, dx = \left[\frac{x^4}{4} - \frac{x^2}{2}\right]_{-1}^{0} + \left[\frac{x^2}{2} - \frac{x^4}{4}\right]_{0}^{1}\\
&= 0 - \left(\frac{1}{4} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{4}\right) - 0 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.
\end{aligned}$$

Q. In the above example, what would happen if we tried to compute

$$\int_{-1}^{1} [f(x) - g(x)]\, dx$$

without first taking the absolute value of the integrand?

A. We would find

$$\int_{-1}^{1} [x - x^3]\, dx = \left[\frac{x^2}{2} - \frac{x^4}{4}\right]_{-1}^{1} = \left(\frac{1}{2} - \frac{1}{4}\right) - \left(\frac{1}{2} - \frac{1}{4}\right) = 0.$$

In general, whenever $f(x) - g(x)$ is an odd function we will find

$$\int_{-1}^{1} [f(x) - g(x)]\, dx = \int_{-1}^{0} [f(x) - g(x)]\, dx + \int_{0}^{1} [f(x) - g(x)]\, dx = 0,$$

because the two contributions are of opposite sign, even though the geometric area of the region bounded by the two functions will (normally) be positive.

- If the function is defined piecewise, we can integrate it in two pieces. For example, to find the area bounded by $y = f(x)$ and $y = 0$ between $x = 0$ and $x = 2$, where

$$f(x) = \begin{cases} x & \text{if } 0 \leq x \leq 1,\\ 2 - x & \text{if } 1 < x \leq 2, \end{cases}$$

we would perform the integral over $[0,1]$ and $[1,2]$ separately:

$$\int_0^2 |f(x)|\,dx = \int_0^1 x\,dx + \int_1^2 (2-x)\,dx = \left[\frac{x^2}{2}\right]_0^1 + \left[-\frac{(2-x)^2}{2}\right]_1^2 = \frac{1}{2} + \frac{1}{2} = 1.$$

However, it is even easier to determine the area of this region by finding the area between the inverse functions $x = f(y) = 2 - y$ and $x = g(y) = y$, where y varies from 0 to 1:

$$\int_0^1 |f(y) - g(y)|\,dy = \int_0^1 |(2-y) - y|\,dy = 2\int_0^1 (1-y)\,dy = 2\left[-\frac{(1-y)^2}{2}\right]_0^1 = 1.$$

8.B Arc Length

Suppose $x(t)$ and $y(t)$ are functions on $[a,b]$ with continuous derivatives. The equations

$$x = x(t), \qquad y = y(t)$$

provide a *parametric representation* of a *smooth curve* $(x(t), y(t))$ in $\mathbb{R}^2$ in terms of the *parameter* t.

As a special case, we could take $x(t) = t$ and $y(t) = f(t)$. The points $(t, f(t))$ describe the familiar *graph* of the function $f(t)$. However, the parametric representation allows us to describe relations, such as circles, that are not the graph of a single function.

Q. What is the length of such a curve?

A. To answer this question, we must first define the notion of what we mean by the "length" of a smooth curve. What we seek is an extension of Pythagoras' Theorem, which allows us to calculate the length of line segments in terms of their endpoints, to general curves.

Definition: The *arc length* or *path length* $s(t)$ of a smooth curve $(x(t), y(t))$ on $[a,b]$ is the unique differentiable function $s(t)$ (if it exists) that satisfies $s(a) = 0$ and the property that

$$\lim_{h\to 0^+} \frac{s(t+h) - s(t)}{\sqrt{[x(t+h) - x(t)]^2 + [y(t+h) - y(t)]^2}} = 1 \tag{8.1}$$

for all $t \in [a,b)$. That is, the difference between the path lengths $s(t+h)$ and $s(t)$ to any points $P = (x(t), y(t))$ and $Q = (x(t+h), y(t+h))$ on the curve, respectively, should reduce to the length of the straight line segment joining P and Q in the limit $h \to 0$ (in which case $Q \to P$).

Upon dividing the numerator and denominator on the left-hand side of Eq. (8.1) by h we see that

$$\frac{\displaystyle\lim_{h\to 0^+} \frac{s(t+h)-s(t)}{h}}{\sqrt{\displaystyle\lim_{h\to 0^+}\left[\frac{x(t+h)-x(t)}{h}\right]^2 + \lim_{h\to 0^+}\left[\frac{y(t+h)-y(t)}{h}\right]^2}} = 1.$$

This gives us a formula for the derivative of $s(t)$ for every $t \in [a,b]$,

$$s'(t) = \sqrt{[x'(t)]^2 + [y'(t)]^2}. \tag{8.2}$$

Upon integrating this result from a to b, we find an expression for the arc length $s(b)$ of a curve $(x(t), y(t))$ on $[a,b]$. Since $s(a) = 0$, we have

$$s(b) = s(b) - s(a) = \int_a^b s'(t)\,dt = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2}\,dt.$$

Remark: One can think of each point on the curve $(x(t), y(t))$ as the position of a point in $\mathbb{R}^2$ at each time t. The integrand $\sqrt{[x'(t)]^2 + [y'(t)]^2}$ is just the magnitude $|\boldsymbol{v}|$ of the *velocity* vector $\boldsymbol{v} = (x'(t), y'(t))$. The arc length, being the integral of the *speed* $|\boldsymbol{v}|$ with respect to time, is then seen to be the distance travelled by the point $(x(t), y(t))$ over the time interval $[a,b]$.

Remark: An easy way to remember the arc-length formula is to multiply Eq. (8.2) formally by dt and square the result:

$$ds^2 = dx^2 + dy^2.$$

This can be thought of as a statement of Pythagoras' Theorem for differentials. The arc length can then be computed by integrating ds between $t = a$ and $t = b$, remembering that $dx = x'(t)\,dt$ and $dy = y'(t)\,dt$.

- Although we defined π to be the area of the unit circle, it is also possible to express the circumference of the unit circle in terms of π. A circle of radius $r \geq 0$ centered on the origin can be described either by the equation $x^2 + y^2 = r^2$ or in parametric form as $(r\cos t, r\sin t)$ for $t \in [0, 2\pi]$. The *circumference* of the circle is then given by

$$\int_0^{2\pi} \sqrt{[x'(t)]^2 + [y'(t)]^2}\,dt = \int_0^{2\pi} \sqrt{r^2\sin^2(t) + r^2\cos^2(t)}\,dt = \int_0^{2\pi} r\,dt = 2\pi r.$$

That is, π can equivalently be defined as the ratio of the circumference of a circle of radius r to its *diameter* $2r$.

Remark: If the curve $(x(t), y(t))$ for $t \in [a, b]$ can be described by a differentiable function $y = f(x)$, then

$$s = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt = \int_{x(a)}^{x(b)} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx = \int_{x(a)}^{x(b)} \sqrt{1 + [f'(x)]^2}\, dx.$$

- We could also compute the circumference of a circle as twice the arc length of the function $f(x) = \sqrt{r^2 - x^2}$ on $[-r, r]$:

$$\begin{aligned} 2\int_{-r}^{r} \sqrt{1 + [f'(x)]^2}\, dx = 2\int_{-r}^{r} \sqrt{1 + \left(\frac{-2x}{2\sqrt{r^2 - x^2}}\right)^2}\, dx = 4\int_0^r \sqrt{1 + \frac{x^2}{r^2 - x^2}}\, dx \\ = 4\int_0^r \sqrt{\frac{r^2}{r^2 - x^2}}\, dx = 4r\int_0^r \frac{1}{\sqrt{r^2 - x^2}}\, dx \\ = 4r\int_0^1 \frac{1}{\sqrt{1 - u^2}}\, du = 4r[\arcsin u]_0^1 = 2\pi r, \end{aligned}$$

where we have used the substitution $u = x/r$.

Remark: Of course, we can also express arc length as an integral in y:

$$s = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt = \int_{y(a)}^{y(b)} \sqrt{\left(\frac{dx}{dy}\right)^2 + 1}\ dy.$$

- Find the arc length L of the parabola $y^2 = x$ between $(0, 0)$ and $(1, 1)$.

Since $dx/dy = 2y$, we know that $L = \int_0^1 \sqrt{(2y)^2 + 1}\, dy$. Let $2y = \tan\theta$, so that $2\, dy = \sec^2\theta\, d\theta$. Then

$$\int \sqrt{4y^2 + 1}\, dy = \frac{1}{2}\int \sec^3\theta\, d\theta = \frac{1}{4}(\sec\theta\tan\theta + \log|\sec\theta + \tan\theta|) + C,$$

using a result from page 170. Thus

$$L = \frac{1}{4}\left[2y\sqrt{4y^2 + 1} + \log\left|\sqrt{4y^2 + 1} + 2y\right|\right]_0^1 = \frac{1}{4}\left[2\sqrt{5} + \log\left(\sqrt{5} + 2\right)\right].$$

8.C Volumes by Cross Sections

Single-variable calculus can sometimes be used to calculate more than just lengths and areas. If an expression for the cross-sectional area of an object is known, it is possible to compute its volume by the *method of cross sections* (also known as the method of slabs or slices).

For example, we can of course easily compute the volume of a loaf of bread, where each slice has same shape and size, using the definition

$$\text{volume} = \text{area} \times \text{length}.$$

But what if the slices of bread don't all have the same size (or even the same shape)? Maybe we have a conical loaf!

Q. What is the volume of such a strange loaf of bread?

A. Slice up the loaf and sum up the area (height $\times$ width) $\times$ thickness $(x_i - x_{i-1})$ of each slice to form the Riemann sum

$$\sum_{i=1}^{n} A(\bar{x}_i)(x_i - x_{i-1}),$$

where $A(x)$ is the area of a cross section at x obtaining by slicing perpendicular to the x-axis and $\bar{x}_i$ is a point in $[x_{i-1}, x_i]$ Assuming that $A(x)$ is integrable on $[a, b]$, we can use restrict our attention to uniform partitions and then take the limit as $n \to \infty$ to find the volume:

$$V = \lim_{n\to\infty} \sum_{i=1}^{n} A(\bar{x}_i)(x_i - x_{i-1}) = \int_a^b A(x)\,dx.$$

- For a conical loaf of bread of length L, the middle slice, at $x = L/2$ has $1/2$ the height and $1/2$ the width of the largest slice, so its area is $1/4$ the area of the largest slice. If we put the apex of the cone at $x = 0$ and the largest slice, with area A at $x = L$, we see by similar triangles that the slice located at x has area $A(x) = (x/L)^2 A$. Thus

$$\begin{aligned} V &= \int_0^L A(x)\,dx = \int_0^L \frac{x^2}{L^2} A\,dx \\ &= \frac{A}{L^2}\int_0^L x^2\,dx = \frac{A}{L^2}\left[\frac{x^3}{3}\right]_0^L = \frac{A}{L^2}\frac{L^3}{3} = \frac{1}{3}AL. \end{aligned}$$

  We have thus established the formula:

$$V_{\text{cone}} = \frac{1}{3}\,\text{base area} \times \text{altitude}.$$

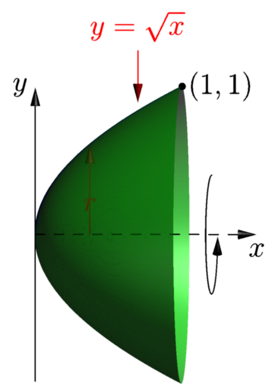
- We can compute the volume enclosed by a sphere of radius R, described by the equation $x^2+y^2+z^2=R^2$, by partitioning the x axis. This produces circular cross sections of radius $r=r(x)>0$. The value of r is the maximum possible value of y, which occurs when $z=0$:

$$x^2+y^2=R^2 \Rightarrow y=\pm\sqrt{R^2-x^2}.$$

That is, $r(x)=\sqrt{R^2-x^2}$, so that $A(x)=\pi r^2=\pi(R^2-x^2)$. Thus

$$V=\int_{-R}^{R}\pi(R^2-x^2)\,dx=2\pi\int_0^R(R^2-x^2)\,dx=2\pi\left[R^2x-\frac{x^3}{3}\right]_0^R=\frac{4}{3}\pi R^3.$$

- Find the volume of the solid obtained by rotating the area bounded by the curves $y=\sqrt{x}$ and $y=0$ from 0 to 1 about the x axis.



Since the radius of revolution is given by $r(x)=y=\sqrt{x}$, the cross-sectional area is given by

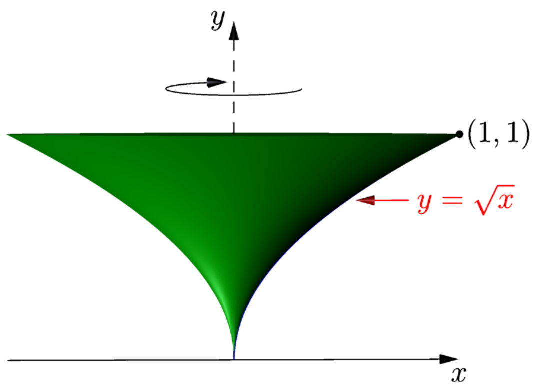
$$A(x)=\pi r^2=\pi\left(\sqrt{x}\right)^2=\pi x.$$

Thus

$$V=\int_0^1 A(x)\,dx=\int_0^1\pi x\,dx=\pi\left[\frac{x^2}{2}\right]_0^1=\frac{\pi}{2}.$$

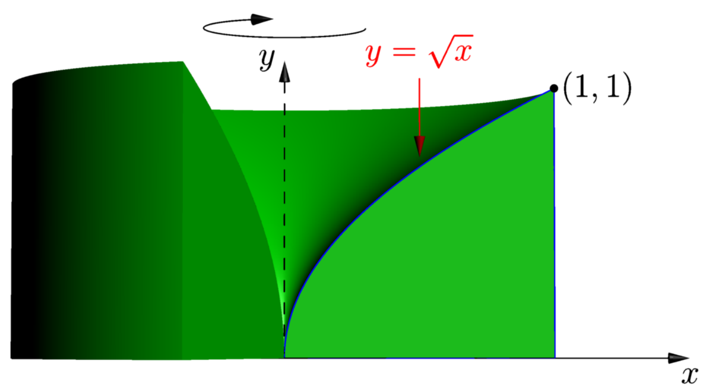
- We could instead compute the volume of the funnel-shaped object generated by rotating the region bounded by $y=\sqrt{x}$, $x=0$, and $y=1$ about the y axis. For the method of cross sections, we always slice the rotation axis (in this case the y axis) and express everything else in terms of the corresponding variable (y). We see that the radius of each circular cross section is $r(y)=x=y^2$, so that the cross-sectional area is $A(y)=\pi r^2=\pi y^4$. The resulting volume of revolution is thus

$$V=\int_0^1 A(y)\,dy=\pi\int_0^1 y^4\,dy=\pi\left[\frac{y^5}{5}\right]_0^1=\frac{\pi}{5}.$$



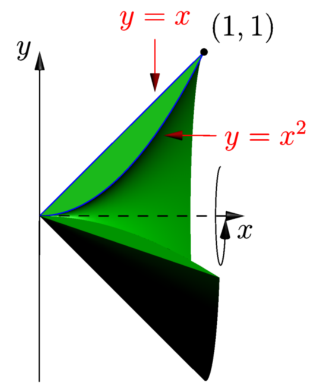
- We could also rotate the region bounded by the curves $y = \sqrt{x}$, $y = 0$, and $x = 1$ about the y axis. If we slice the y axis, each cross section is just an annulus of outer radius $r_{\text{out}}(y) = 1$ and inner radius $r_{\text{in}}(y) = x = y^2$, with area $A(y) = \pi {r_{\text{out}}}^2 - \pi {r_{\text{in}}}^2 = \pi(1 - y^4)$. The volume of the resulting object is then

$$V = \int_0^1 A(y)\,dy = \pi \int_0^1 (1 - y^4)\,dy = \pi \left[y - \frac{y^5}{5} \right]_0^1 = \frac{4\pi}{5}.$$



Problem 8.1: Explain why the volumes calculated in the previous two examples add up to π, the volume of a cylinder of unit radius and unit height.

- If we rotate the region bounded by $f(x) = x$ and $g(x) = x^2$ between $x = 0$ and $x = 1$ about the x axis,



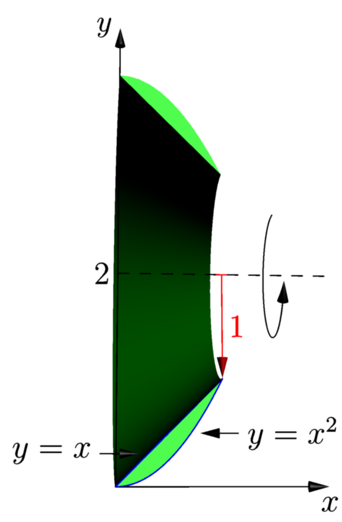
we need to find the area of the annular region with outer radius $r_{\text{out}}(x) = f(x) = x$ and inner radius $r_{\text{in}}(x) = g(x) = x^2$:

$$A(x) = \pi {r_{\text{out}}}^2 - \pi {r_{\text{in}}}^2 = \pi\left[x^2 - (x^2)^2\right] = \pi(x^2 - x^4).$$

Thus

$$V = \int_0^1 [\pi(x^2 - x^4)]\,dx = \pi\left[\frac{x^3}{3} - \frac{x^5}{5}\right]_0^1 = \pi\left(\frac{1}{3} - \frac{1}{5}\right) = \frac{2\pi}{15}.$$

- We could also rotate the same area about the line $y = 2$ instead of $y = 0$.



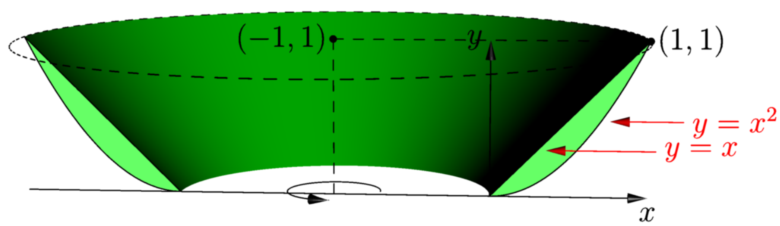
Now

$$A(x) = \pi {r_{\text{out}}}^2 - \pi {r_{\text{in}}}^2 = \pi\left(2 - x^2\right)^2 - \pi(2 - x)^2 = \pi(x^4 - 5x^2 + 4x).$$

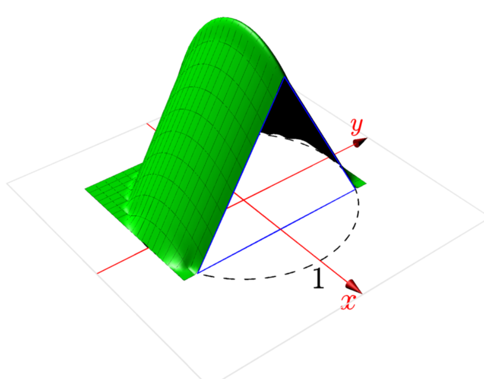
The generated volume is then

$$V = \int_0^1 \mathrm{A}(x)\,dx = \pi \int_0^1 (x^4 - 5x^2 + 4x)\,dx = \pi \left[\frac{x^5}{5} - \frac{5x^3}{3} + \frac{4x^2}{2}\right]_0^1$$
$$= \pi\left(\frac{1}{5} - \frac{5}{3} + 2\right) = \frac{8\pi}{15}.$$

Problem 8.2: Find the volume generated by rotating the region bounded by $f(x) = x$ and $g(x) = x^2$ between $x = 0$ and $x = 1$ about the line $x = -1$.



- Consider the three-dimensional object formed by erecting an equilateral triangle, with altitude perpendicular to the x–y plane, on every chord $x =$ const of the circle $x^2 + y^2 = 1$.



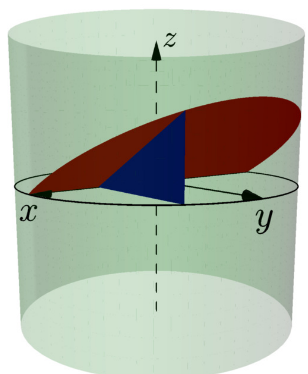
To find the volume of this object, we only need to find the cross-sectional area $A(x)$ of each equilateral triangle obtained by slicing the object along the planes $x =$ const. The length of the base of this triangle, which has endpoints $(x, -y)$ and (x, y), where $x^2 + y^2 = 1$, is $2y$. Pythagoras' Theorem tell us that the altitude h of this equilateral triangle is $\sqrt{(2y)^2 - y^2} = \sqrt{3}y$. Hence

$$A(x) = \frac{1}{2}(2y)h = \sqrt{3}y^2 = \sqrt{3}(1 - x^2).$$

The volume of the object is then easily computed:

$$V = \int_{-1}^{1} A(x)\,dx = \int_{-1}^{1} \sqrt{3}(1-x^2)\,dx = 2\int_{0}^{1} \sqrt{3}(1-x^2)\,dx = 2\sqrt{3}\left[x - \frac{x^3}{3}\right]_0^1 = \frac{4\sqrt{3}}{3}.$$

- Consider the volume of one of the two wedge-shaped regions bounded by the cylinder $x^2 + y^2 = 16$ and the plane containing the x axis and oriented at an angle of $30^\circ = \pi/6$ to the x–y plane. If we slice this object in the x direction, we obtain triangular cross sections with base length y and altitude $y\tan(\pi/6) = y/\sqrt{3}$. Thus



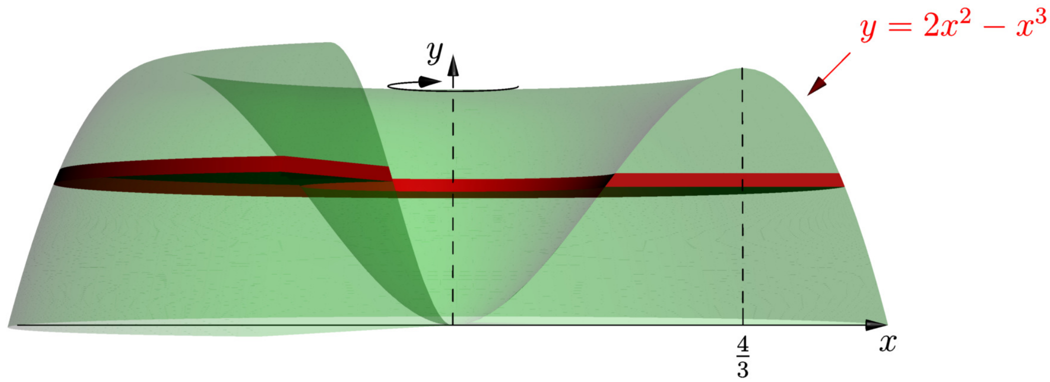
$$A(x) = \frac{1}{2}y\left(\frac{y}{\sqrt{3}}\right) = \frac{16 - x^2}{2\sqrt{3}},$$

so that

$$V = \int_{-4}^{4} \frac{16 - x^2}{2\sqrt{3}}\,dx = 2\int_{0}^{4} \frac{16 - x^2}{2\sqrt{3}}\,dx = \frac{1}{\sqrt{3}}\left[16x - \frac{x^3}{3}\right]_0^4 = \frac{128}{3\sqrt{3}}.$$

8.D Volume by Shells

Suppose we wish to rotate the area under the curve $y = f(x) = 2x^2 - x^3$ about the y axis. The method of cross sections requires that we slice the y axis and express all quantities as functions of y. Finding the radii r_{in} and r_{out} amounts to inverting the equation $y = 2x^2 - x^3$ to find two distinct values of x for every y within the limits of integration.



In general, performing this kind of inversion can be a difficult problem. In this example, f has roots only at $x = 0$ and $x = 2$ and $f(x) > 0$ on $(0, 2)$. We can easily see that the maximum value of f must occur at $4/3$, since $f'(x) = 4x - 3x^2 = x(4 - 3x)$. However, it is much more difficult (although in this case not impossible) to find for each y the two values r_{in} and r_{out} such that $f(r_{\text{in}}) = f(r_{\text{out}}) = y$.

For such cases, there is an easier alternative, the *method of cylindrical shells*, where one computes the volume using Riemann sums of volumes of cylindrical shells:

1. Partition an axis that is perpendicular to the rotation axis. For each subinterval of the partition, compute the volume of the cylindrical shell generated by revolving the portion of the curve that lies in that subinterval around the rotation axis.

2. To find the total volume, add up the volumes of all shells and take the limit as the width of the subintervals goes to 0.

We readily see that the volume of a cylindrical shell of inner radius r_1 and outer radius r_2 and height h is given by

$$\pi r_2^2 h - \pi r_1^2 h = \pi h(r_2^2 - r_1^2) = \pi h(r_2 + r_1)(r_2 - r_1) = (2\pi r)h\Delta_r,$$

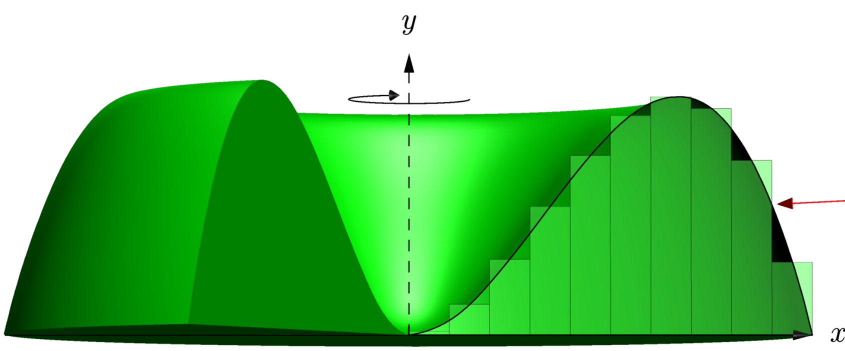
where $r \doteq (r_1 + r_2)/2$ is the mean radius and $\Delta_r \doteq r_2 - r_1$ is the width of the subinterval.

When the area under the curve $y = f(x)$ is rotated about the y axis, we can use a uniform partition to form a Riemann sum for the volume by approximating the height h of the cylindrical shell on each subinterval $[x_{i-1}, x_i]$ by the value of the function f at the mean radius $\bar{x}_i = (x_{i-1} + x_i)/2$. Then

$$V = \lim_{n\to\infty} \sum_{i=1}^{n} 2\pi \bar{x}_i f(\bar{x}_i)(x_i - x_{i-1}) = 2\pi \int_a^b x f(x)\, dx.$$

Note here that $0 \le a \le b$.

- We can compute the volume formed by rotating the region under $y = f(x) = 2x^2 - x^3$ about the y axis very easily now, using the method of cylindrical shells:



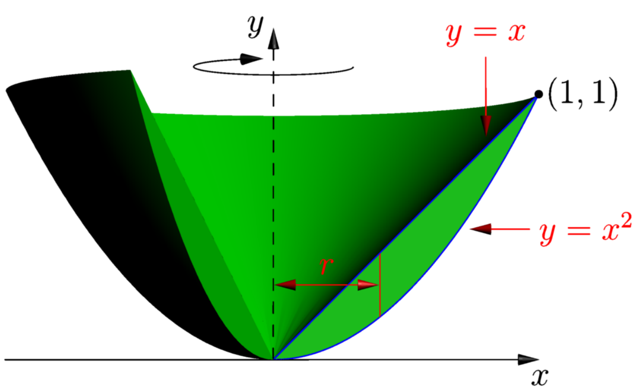
$$V = \int_0^2 2\pi x(2x^2 - x^3)\,dx = 2\pi \int_0^2 (2x^3 - x^4)\,dx$$
$$= 2\pi \left[\frac{2x^4}{4} - \frac{x^5}{5}\right]_0^2 = 2\pi\left(8 - \frac{32}{5}\right) = \frac{16\pi}{5}.$$

In general, the volume of the object generated by rotating the region bounded by the functions $f(x)$ and $g(x)$ about the y axis is given by

$$V = \int_0^1 \underbrace{2\pi x}_{\text{circumference}} \underbrace{|f(x) - g(x)|}_{\text{height}} \underbrace{dx}_{\text{width}},$$

where we have given a geometric interpretation for each factor.

- When the region bounded by the functions $f(x) = x$ and $g(x) = x^2$ and the lines $x = 0$ and $x = 1$ is rotated about the y axis,



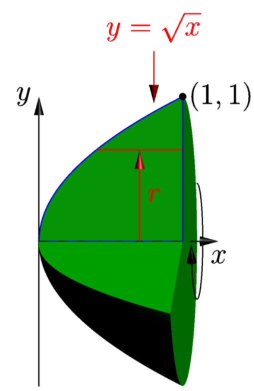
the volume generated is

$$V = 2\pi \int_0^1 x(x - x^2)\, dx = 2\pi \left[\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1 = 2\pi \left(\frac{1}{3} - \frac{1}{4}\right) = \frac{\pi}{6}.$$

Alternatively, we could have obtained the same answer with the method of cross sections by slicing the y axis. Since $r_{\mathrm{out}} = \sqrt{y}$ and $r_{\mathrm{in}} = y$, we see that

$$V = \pi \int_0^1 ({r_{\mathrm{out}}}^2 - {r_{\mathrm{in}}}^2)\, dy = \pi \int_0^1 \left[(\sqrt{y})^2 - y^2\right] dy = \pi \left[\frac{y^2}{2} - \frac{y^3}{3}\right]_0^1 = \frac{\pi}{6}.$$

- We can of course also rotate a region about the x axis. The region bounded by $y = \sqrt{x}$, $y = 0$, $x = 0$, and $x = 1$



would generate the volume

$$V = \int \underbrace{2\pi y}_{\text{circumference}} \underbrace{(1 - x)}_{\text{height}} \underbrace{dy}_{\text{width}} = 2\pi \int_0^1 y(1 - y^2)\, dy$$
$$= 2\pi \left[\frac{y^2}{2} - \frac{y^4}{4}\right]_0^1 = 2\pi \left(\frac{1}{2} - \frac{1}{4}\right) = \frac{\pi}{2},$$

in agreement with the result we previously obtained using the method of cross sections.

8.E Work

Mechanical *work* W is defined as the product of force F times distance D:

$$\underset{\text{(work=force} \times \text{distance)}}{W = FD,} \qquad \text{where} \qquad \underset{\text{(force=mass} \times \text{acceleration)}}{F = ma,}$$

according to Newton's Second Law. Suppose you lift 1 kg vertically 1 m, against the force of gravity (where the acceleration a due to gravity is $g = 9.8$ m/s^2). The force required is

$$F = 1\text{kg} \times 9.8\frac{\text{m}}{\text{s}^2} = 9.8\frac{\text{kg}\cdot\text{m}}{\text{s}^2} = 9.8\,\text{N (Newtons)}.$$

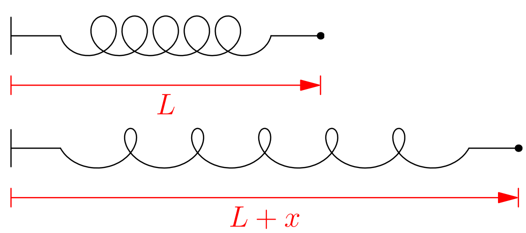
The work required is $W = FD = 9.8\ \text{N} \times 1\,\text{m} = 9.8\,\text{N}\cdot\text{m} = 9.8\,\text{J}$ (Joules). Often, the force F is not constant and we must integrate the force over the entire distance D:

$$W = \int_0^D F(x)\,dx.$$

In general, the work required to move an object from $x = a$ to $x = b$ is

$$W = \int_a^b F(x)\,dx.$$

For example, the force F required to stretch a spring is proportional to the extension x relative to its unstretched (equilibrium) position L. That is, $F(x) = kx$, where k is a constant.



Sometimes, the value of k is not given explicitly, but its value can be calculated from specified force and extension values.

- The force required to stretch a spring from its equilibrium length of 10 cm to 15 cm is 40 N. Calculate the additional work required to stretch the spring an additional 3 centimeters, to a final length of 18 cm.

  Converting all distances to SI (m, k, s) units, we find

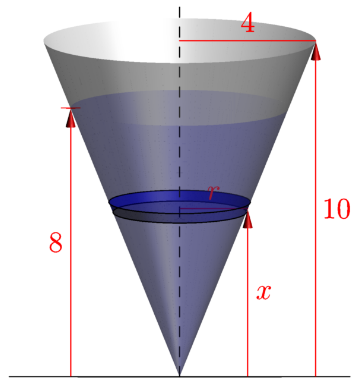
$$40\text{N} = k\,0.05\,\text{m} \Rightarrow k = \frac{800\,\text{N}}{\text{m}}.$$

Thus, the work required to stretch the spring from 0.05 m to 0.08 m beyond its equilibrium length is

$$W = \int_{0.05}^{0.08} kx\,dx = \frac{1}{2}kx^2\Big|_{0.05}^{0.08} = \frac{1}{2}\left(800\ \frac{\text{N}}{\text{m}}\right)\left(0.0064\,m^2 - 0.0025\,m^2\right) = 1.56\ \text{J}.$$

Sometimes the force varies due to a nonconstant volume.

- Suppose, for example, that we want to calculate the work required to pump all of the water out of an inverted cone (apex down) of height 10 m and radius 4 m, if the initial water level is 2 m below the top surface of the cone. We could pump out the water layer by layer, using the method of cross sections.



If we use the variable x to represent height above the apex of the cone, the initial water level is at 8 m. The area of each cross section of radius r is $A(x) = \pi r^2$ and the value of r can be determined in terms of the height x of the cross section by similar triangles:

$$\frac{r}{x} = \frac{4}{10} = \frac{2}{5}.$$

Thus $A(x) = \pi(2x/5)^2$, and the volume of water that we need to pump out is

$$V = \int_0^8 \underbrace{A(x)\,dx}_{\text{volume element}} = \frac{4}{25}\pi \int_0^8 x^2\,dx.$$

The mass M of this volume of water can then be expressed in terms of the density of water, ρ:

$$M = \int_0^8 \underbrace{\rho A(x)\,dx}_{\text{mass element}} = \rho V.$$

It is often a good approximation to assume that the density of water is constant (independent of depth), with the value $\rho = 1000\ \text{kg/m}^3$.

The total force needed to lift all of the water is given by Newton's Second Law:

$$F = \rho g \int_0^8 A(x)\,dx.$$

However, the distance in meters that we need to raise each layer of water located at height x to the upper surface, located at height 10, varies with x as $10 - x$. Thus, the work required to pump all of the water out is given by

$$
\begin{aligned}
W &= \rho g \int_0^8 (10-x)A(x)\,dx = \frac{4}{25}\pi\rho g \int_0^8 (10-x)x^2\,dx = \frac{4}{25}\pi\rho g \int_0^8 (10x^2 - x^3)\,dx \\
&= \frac{4}{25}\pi\rho g \left[\frac{10x^3}{3} - \frac{x^4}{4}\right]_0^8 = \frac{4}{25}\pi\left(9.8\frac{\mathrm{m}}{\mathrm{s}^2}\right)\left(1000\frac{\mathrm{kg}}{\mathrm{m}^3}\right)\left(\frac{10}{3} - \frac{8}{4}\right)8^3\,\mathrm{m}^4 \\
&= 3.36 \times 10^6\,\mathrm{J}.
\end{aligned}
$$

8.F Hydrostatic Force

Integration is often useful for calculating hydrostatic forces, such as the outward force on one wall of a swimming pool. The hydrostatic force F on a horizontal area A submerged under water to a depth D is given by the weight of water (of volume $V = AD$) above the area: $F = \rho V g = \rho A D g$. The *pressure* P, or force per unit area, arising from this weight of water is given by $P = F/A = \rho g D$. Being independent of the area, the pressure provides a useful means of computing the hydrostatic force on an area that is not at a constant depth.

To compute the hydrostatic force on a wall, divide the wall up into infinitesimal horizontal strips at depth $D(y)$ of width $w(y)$ and height dy, where y is a variable in the vertical direction. It is often convenient to set up the y axis so that $y = 0$ corresponds to the surface and with y increasing downwards (with increasing depth) so that $D(y) = y$. The force on each horizontal infinitesimal strip at depth $D(y)$ is then given by the product of the pressure $\rho g D(y)$ at that depth times the area $w(y)\,dy$ of the strip. The total force on the wall is then given by the integral

$$
F = \int \rho g D(y) w(y)\,dy.
$$

- To compute the hydrostatic force on a semicircular wall of radius a which forms the end of a swimming pool filled with water, between depth $y = 0$ and $y = a$, we note that $D(y) = y$ and compute

  $$
  F = \rho g \int_0^a y w(y)\,dy,
  $$

  assuming that the density ρ and gravitational acceleration g are both constant.

If (x, y) is a point on the semicircular edge of the wall, then $x^2 + y^2 = a^2$. Hence $w(y) = 2x = 2\sqrt{a^2 - y^2}$. Thus

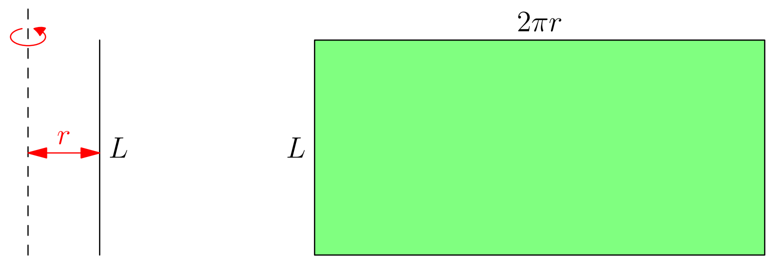
$$
F = \rho g \int_0^a 2y\sqrt{a^2 - y^2}\,dy = \rho g \left[-\frac{2}{3}(a^2 - y^2)^{3/2}\right]_0^a = \frac{2}{3}\rho g a^3.
$$

For example, for $a = 1$ m we find $F = 6.67 \times 10^3$ N.

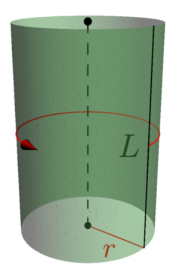
8.G Surfaces of Revolution

We have already discussed methods for finding the volume of an object that results when we rotate a smooth curve about an axis. We now consider how to find the surface area of such an object.

- If we revolve a line segment L about an axis parallel to itself, we obtain a cylinder. If we cut this cylinder along L and unfold it, we see immediately that its surface area is given by the product of its circumference $2\pi r$ and length L: $A = 2\pi r L$.

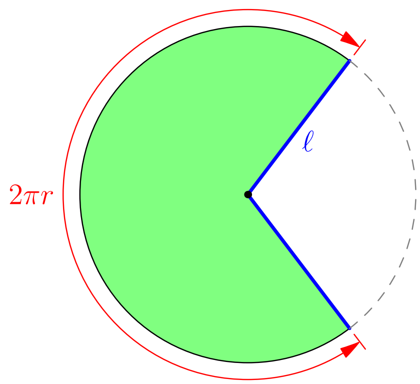


In three dimensions, the green shaded region can be wrapped into a cylinder:

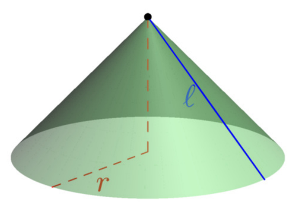


- Similarly, the surface area of a cone of slant height ℓ and radius r is given by

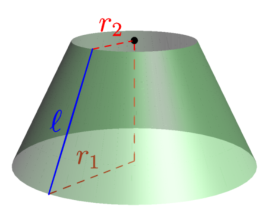
$$A = \underbrace{\left(\frac{2\pi r}{2\pi \ell}\right)}_{\substack{\text{fraction}\\ \text{of circle}}} \underbrace{\pi \ell^2}_{\substack{\text{area of}\\ \text{circle of}\\ \text{radius } \ell}} = \pi r \ell.$$

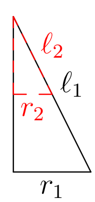


In three dimensions, the two blue lines in the above figure can be joined by wrapping the green shaded region into a cone:



- A *conical band* (*frustum*) is obtained by removing from a large cone of radius r_1 and slant height ℓ_1 a smaller cone of radius $r_2 < r_1$ and slant height ℓ_2 with the same axis of symmetry,





such that (by similar triangles)

$$\frac{\ell_1}{r_1} = \frac{\ell_2}{r_2}.$$

The surface area of a conical band may be computed as the difference of the respective surface areas A_1 and A_2 of the large and small cones. We may express this area as

$$A = A_1 - A_2 = \pi r_1 \ell_1 - \pi r_2 \ell_2 = \pi(r_1\ell_1 - r_2\ell_2) = \pi(r_1 + r_2)(\ell_1 - \ell_2) \doteq 2\pi r \ell$$

since $r_1\ell_2 - r_2\ell_1 = 0$, where $r \doteq (r_1 + r_2)/2$ is the mean radius and $\ell \doteq \ell_1 - \ell_2$ is the length of the sloped edge of the band.

Remark: Thus, the surface area generated by rotating a straight line segment of length ℓ about an axis a mean distance r away is just $2\pi r\ell$.

We can now calculate the surface area of the object formed by rotating any smooth curve about an axis.

Definition: The *surface area* of the object formed by rotating the smooth curve $(x(t), y(t))$ on $[a, b]$ is the unique differentiable function $A(t)$ that satisfies $A(a) = 0$ and the property that

$$\lim_{h\to 0^+} \frac{A(t+h) - A(t)}{2\pi\left(\dfrac{r(t) + r(t+h)}{2}\right)\sqrt{[x(t+h) - x(t)]^2 + [y(t+h) - y(t)]^2}} = 1,$$

for all $t \in [a, b)$, where $r(t)$ is the distance of $(x(t), y(t))$ from the axis of rotation.

Hence

$$A'(t) = 2\pi r(t)\sqrt{[x'(t)]^2 + [y'(t)]^2},$$

so that

$$A(b) = A(b) - A(a) = 2\pi \int_a^b r(t)\sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt.$$

Remark: An easy way to remember this result is to integrate the product of the circumference $2\pi r$ (associated with a complete revolution of the curve about the axis) and the infinitesimal arc length $ds = \sqrt{dx^2 + dy^2}$.

- The area generated by revolving the curve $y = f(x)$ for $x \in [a, b]$ about the x axis is

$$2\pi \int |y|\ ds = 2\pi \int_a^b |f(x)|\ \sqrt{1 + [f'(x)]^2}\, dx.$$

- The area generated by revolving the curve $y = f(x)$ for $x \in [a, b]$ about the y axis is

$$2\pi \int |x|\ ds = 2\pi \int_a^b |x|\ \sqrt{1 + [f'(x)]^2}\, dx.$$

- The surface area of a sphere of radius a can be computed by revolving the curve $y = \sqrt{a^2 - x^2}$ for $x \in [-a, a]$ about the x axis. Since $dy/dx = -x/\sqrt{a^2 - x^2}$, the surface area is seen to be

$$2\pi \int y\, ds = 2\pi \int_{-a}^{a} y\sqrt{1 + \frac{x^2}{a^2 - x^2}}\, dx = 2\pi \int_{-a}^{a} \sqrt{a^2 - x^2}\left(\frac{a}{\sqrt{a^2 - x^2}}\right) dx = 4\pi a^2.$$

- Alternatively, the surface area of a sphere of radius a can be computed using the parametric representation $(a\cos t, a\sin t)$ of a half circle, with $t \in [0, \pi]$. If we rotate this curve about the x axis, the surface area is seen to be

$$2\pi \int y\, ds = 2\pi \int_0^{\pi} a\sin t\sqrt{a^2\sin^2 t + a^2\cos^2 t}\, dt = 2\pi a^2 \int_0^{\pi} \sin t\, dt = 2\pi a^2[-\cos t]_0^{\pi} = 4\pi a^2.$$

- The surface area generated by rotating the section of the parabola $y = x^2$ from $(1, 1)$ to $(2, 4)$ about the y axis can be computed from the formula $2\pi \int x\, ds = 2\pi \int x\sqrt{dx^2 + dy^2}$:

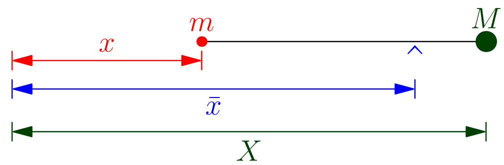
$$\begin{aligned} 2\pi \int_1^2 x\sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx &= 2\pi \int_1^2 x\sqrt{1 + 4x^2}\, dx \\ &= 2\pi \left[\frac{2}{3}\left(1 + 4x^2\right)^{\frac{3}{2}}\frac{1}{8}\right]_1^2 = \frac{\pi}{6}\left(17^{\frac{3}{2}} - 5^{\frac{3}{2}}\right). \end{aligned}$$

8.H Centroids and Pappus's Theorems

In mechanics, one often needs to find the point at which an object will balance under its own weight.

Q. At what position should the fulcrum (pivot point) of a (massless) lever be placed so that a mass m at position x and a mass M at position X are in balance?

A. Archimedes noticed that the masses will be in balance if the product of each mass times its respective distance from the fulcrum is the same. Such products are known as *moments*.



Let $\bar{x}$ denote the position of the fulcrum. Then

$$\begin{aligned} m(\bar{x}-x) &= M(X-\bar{x}) \\ \Rightarrow (M+m)\bar{x} &= MX+mx \\ \Rightarrow \bar{x} &= \frac{MX+mx}{M+m}. \end{aligned}$$

The position $\bar{x}$ is known as the *center of mass* of this system of two masses.

In general, if you have n objects with masses m_i located at x_i, for $i = 1, 2, \ldots, n$, they will balance at the point

$$\bar{x} = \frac{\displaystyle\sum_{i=1}^{n} m_i x_i}{\displaystyle\sum_{i=1}^{n} m_i}.$$

The expression $\sum_{i=1}^{n} m_i x_i$ in the numerator is known as the *first moment* of the system. We recognize the expression in the denominator as the total mass of the system.

Remark: In three dimensions, the center of mass $(\bar{x}, \bar{y}, \bar{z})$ of a system of n particles of mass m_i located at (x_i, y_i, z_i) is given by

$$\bar{x} = \frac{\displaystyle\sum_{i=1}^{n} m_i x_i}{\displaystyle\sum_{i=1}^{n} m_i}, \qquad \bar{y} = \frac{\displaystyle\sum_{i=1}^{n} m_i y_i}{\displaystyle\sum_{i=1}^{n} m_i}, \qquad \bar{z} = \frac{\displaystyle\sum_{i=1}^{n} m_i z_i}{\displaystyle\sum_{i=1}^{n} m_i}.$$

In we define the total mass $M = \sum_{i=1}^{n} m_i$, we see that the moments of the whole system about each axis are the same as those of a single particle of mass M located at the center of mass $(\bar{x}, \bar{y}, \bar{z})$.

For a continuum one-dimensional mass distribution, such as a thin wire, $m_i \to dm \doteq \rho\, dx$, where ρ is the *mass density*, and

$$\bar{x} = \frac{\displaystyle\int x\rho\, dx}{\displaystyle\int \rho\, dx} = \frac{1}{M}\int x\rho\, dx,$$

where $M = \int \rho\, dx$ is the total mass of the system. Given a *uniform* mass distribution over the interval $[a, b]$, for which ρ is constant, the center of mass becomes the *centroid*:

$$\bar{x} = \frac{\displaystyle\int_a^b x\, dx}{\displaystyle\int_a^b dx} = \frac{\frac{1}{2}(b^2-a^2)}{b-a} = \frac{a+b}{2}.$$

If a wire of length $L = \int ds$ is bent to form a smooth curve $(x(t), y(t))$, then $dm = \rho\,dx \to \rho\,ds$, so the coordinates $(\bar{x}, \bar{y})$ of the center of mass are

$$\bar{x} = \frac{\displaystyle\int \rho x\,ds}{\displaystyle\int \rho\,ds}, \qquad \bar{y} = \frac{\displaystyle\int \rho y\,ds}{\displaystyle\int \rho\,ds}.$$

If the wire is of uniform density ($\rho =$ constant), we obtain the coordinates of its centroid:

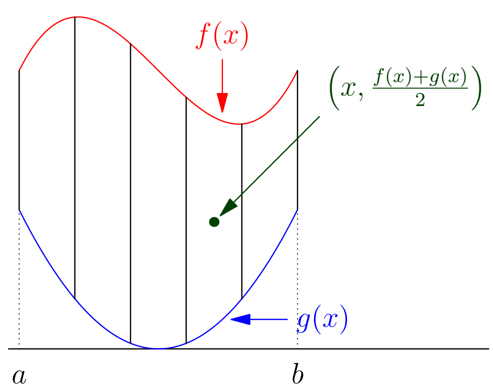
$$\bar{x} = L^{-1}\int x\,ds, \qquad \bar{y} = L^{-1}\int y\,ds.$$

Problem 8.3: A piece of wire of length L is bent into a semicircle. Compute the coordinates of its centroid.

In two dimensions, the centroid (balance point) of a plane area of constant density (e.g, a uniform plate) can be calculated by splitting the region into infinitesimal strips of width dx, height $\ell(x)$, and mass $dm = \rho\ell(x)\,dx$. As we have seen above, the centroid of each infinitesimal strip is located at its midpoint, say $(x, \tilde{y})$. The coordinates of the centroid $(\bar{x}, \bar{y})$ are thus given by

$$\bar{x} = \frac{\int x\ell(x)dx}{\int \ell(x)dx}, \qquad \bar{y} = \frac{\int \tilde{y}\ell(x)dx}{\int \ell(x)dx}.$$

- To find the centroid $(\bar{x}, \bar{y})$ of the region bounded by the continuous functions $y = f(x)$ and $y = g(x)$ between $x = a$ and $x = b$,



we note that $\ell(x) = |f(x) - g(x)|$ and $\tilde{y} = (f(x) + g(x))/2$. Thus

$$\bar{x} = \frac{1}{A}\int_a^b x\,|f(x) - g(x)|\;dx,$$

$$\bar{y} = \frac{1}{A}\int_a^b \left(\frac{f(x)+g(x)}{2}\right)|f(x)-g(x)|\,dx,$$

where $A = \int_a^b |f(x)-g(x)|\,dx$. In the case where $f(x) \geq g(x)$, the expression for $\bar{y}$ simplifies to

$$\bar{y} = \frac{1}{2A}\int_a^b [f^2(x) - g^2(x)]\,dx.$$

Problem 8.4: Prove the *symmetry principle* for centroids: the centroid of an object remains fixed if we flip an object about a line of symmetry.

- The centroid of the semicircular region bounded by $f(x) = \sqrt{a^2 - x^2}$ and $g(x) = 0$, with area $A = \pi a^2/2$, is $(0, \bar{y})$, where

$$\bar{y} = \frac{1}{2A}\int_{-a}^a (a^2 - x^2)\,dx = \frac{2}{\pi a^2}\int_0^a (a^2 - x^2)\,dx = \frac{2}{\pi a^2}\left[a^2 x - \frac{x^3}{3}\right]_0^a = \frac{4a}{3\pi}.$$

The following theorem sometimes provides an easier way of computing centroids.

Theorem 8.1 (Pappus's Theorems): *Let $\mathcal{L}$ be a line in a plane.*

(i) If a curve lying entirely on one side of $\mathcal{L}$ is rotated about $\mathcal{L}$, the area of the surface generated is the length of the curve times the distance travelled by the centroid.

(ii) If a region lying entirely on one side of $\mathcal{L}$ is rotated about $\mathcal{L}$, the volume of the solid generated is the area of the region times the distance travelled by the centroid.

Proof:

(i) If we rotate a curve of length $L = \int ds$ by an angle α about $\mathcal{L}$ (instead of a full rotation by an angle 2π), the surface area generated is

$$A = \int \alpha x\,ds = \alpha \int x\,ds = \alpha\bar{x}\int ds = (\alpha\bar{x})L,$$

where x represents the distance of each length element from $\mathcal{L}$ and $\bar{x} = L^{-1}\int x\,ds$ is the component of the centroid of the curve perpendicular to $\mathcal{L}$. Note that $\alpha\bar{x}$ is the distance travelled by the centroid.

(ii) Let $\ell(x)$ be the length of an infinitesimal strip of the region, parallel to $\mathcal{L}$ and at a distance x from $\mathcal{L}$. Using the method of shells, we find that the volume generated by rotating the region of area $A = \int \ell(x)\,dx$ by an angle α about $\mathcal{L}$ is

$$V = \int \alpha x\ell(x)\,dx = \alpha\int x\ell(x)\,dx = \alpha\bar{x}\int \ell(x)\,dx = (\alpha\bar{x})A,$$

where $\bar{x} = A^{-1}\int x\ell(x)\,dx$ is the centroid of the region.

- To find the centroid $(0, \bar{y})$ of the semicircular region $\{(x, y) : x^2 + y^2 \leq a^2, y \geq 0\}$, we may use Pappus's Second Theorem and the volume $\frac{4}{3}\pi a^3$ of a sphere:
$$\frac{4}{3}\pi a^3 = \frac{1}{2}\pi a^2(2\pi\bar{y}) \Rightarrow \bar{y} = \frac{4a}{3\pi},$$
in agreement with our previous result obtained by direct integration.

- By Pappus's Second Theorem, the volume of a torus of major radius R and minor radius a is $(2\pi R)(\pi a^2) = 2\pi^2 R a^2$.

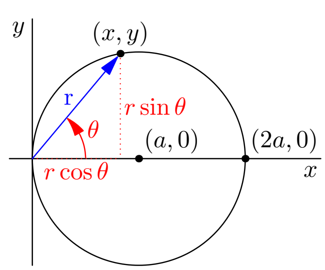
8.I Polar Coordinates

Polar coordinates (r, θ) are related to the usual *Cartesian coordinates* (x, y) by
$$\begin{aligned} x &= r\cos\theta, \\ y &= r\sin\theta. \end{aligned}$$

Remark: Polar coordinates are not unique:
$$\begin{aligned} x &= r\cos(\theta + 2m\pi), \\ y &= r\sin(\theta + 2m\pi), \end{aligned}$$
specify the same point $\forall m \in \mathbb{Z}$. Also, the points $(r, \theta + \pi)$ and $(-r, \theta)$ are identical and $(0, \theta)$ denotes the origin for all θ.

- Describe the circle $(x - a)^2 + y^2 = a^2$ in polar coordinates.
$$\begin{aligned} x^2 + y^2 - 2ax &= 0 \\ \Rightarrow r^2 - 2ra\cos\theta &= 0 \\ \Rightarrow r(r - 2a\cos\theta) &= 0 \\ \Rightarrow r = 0 \text{ or } r &= 2a\cos\theta. \end{aligned}$$

Remark: Thus, a point on the circle $(x - a)^2 + y^2 = a^2$ is either the origin $(r = 0)$ or else it lies on the curve $r = 2a\cos\theta$. In fact, since the origin is already contained in the second solution $r = 2a\cos\theta$ (at $\theta = \pi/2$), this equation alone generates the entire curve.



Remark: Notice that θ varies from 0 to 2π, the point (r, θ) moves **twice** around the circle (this corresponds to an elementary result from geometry that the angle subtended by an arc measured at the center of a circle is twice that measured on the circumference). Therefore, in order to compute the arc length of the circle $r(\theta) = 2a\cos\theta$, we should only integrate from 0 to π.

Q. How can we, using polar coordinates, find the arc length of a curve $r = r(\theta)$ for $\theta \in [a, b]$?

A. Use the fact that $x(\theta) = r(\theta)\cos\theta$ and $y(\theta) = r(\theta)\sin\theta$:

$$\begin{aligned}\int ds = \int \sqrt{dx^2 + dy^2} &= \int_a^b \sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2}\, d\theta \\ &= \int_a^b \sqrt{[r'(\theta)\cos\theta - r(\theta)\sin\theta]^2 + [r'(\theta)\sin\theta + r(\theta)\cos\theta]^2}\, d\theta \\ &= \int_a^b \sqrt{r'^2 + r^2}\, d\theta.\end{aligned}$$

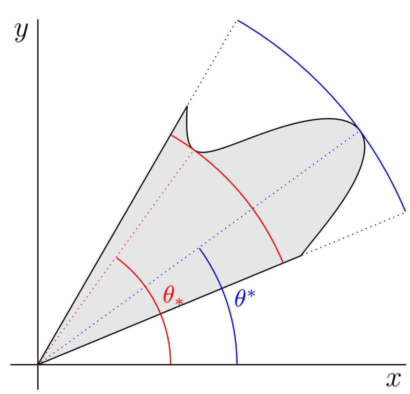
- The circumference of the circle $r = 2a\cos\theta$ (which has radius a), is

$$\int_0^\pi \sqrt{4a^2\sin^2\theta + 4a^2\cos^2\theta}\, d\theta = \pi(2a) = 2\pi a.$$

As noted above, we only integrate from $\theta = 0$ to $\theta = \pi$.

Q. Can we also compute the area of a region bounded by a continuous curve, say $r = f(\theta) \geq 0$ for $\theta \in [a, b]$, in polar coordinates?

A. Yes. Let P be a partition of $[a, b]$. If f is continuous, there exists points θ_* and θ^* where f takes on its minimum and maximum values, respectively, in each subinterval of P.



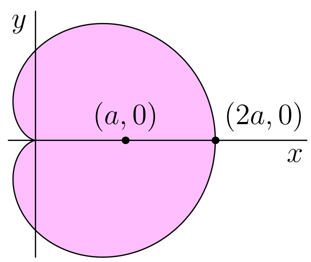
The area contribution Δ_A from each subinterval of width Δ_θ must lie between the areas $r^2\Delta_\theta/2$ bounded by the circular arcs $r = f(\theta_*)$ and $r = f(\theta^*)$:

$$f^2(\theta_*)\frac{\Delta_\theta}{2} \le \Delta_A \le f^2(\theta^*)\frac{\Delta_\theta}{2}$$
$$\Rightarrow \lim_{\Delta_\theta\to 0}\frac{f^2(\theta_*)}{2} \le \lim_{\Delta_\theta\to 0}\frac{\Delta_A}{\Delta_\theta} \le \lim_{\Delta_\theta\to 0}\frac{f^2(\theta^*)}{2}.$$

Since $\lim\limits_{\Delta_\theta\to 0} f^2(\theta_*) = \lim\limits_{\Delta_\theta\to 0} f^2(\theta^*) = f^2(\theta)$, we see from the Squeeze Principle for Functions that

$$\frac{dA}{d\theta} = \frac{1}{2}f^2(\theta) \Rightarrow A = \frac{1}{2}\int f^2(\theta)\,d\theta.$$

- The area enclosed by the *cardioid* $r = a(1 + \cos\theta)$



is given by

$$A = \frac{1}{2}\int_0^{2\pi} a^2(1+\cos\theta)^2\,d\theta = \frac{a^2}{2}\int_0^{2\pi}(1+2\cos\theta+\cos^2\theta)\,d\theta = \frac{a^2}{2}(2\pi+\pi) = \frac{3\pi a^2}{2}.$$

Chapter 9

Improper Integrals and Infinite Series

Until now, we have only defined the Riemann integral for bounded functions on closed intervals. In this chapter, we discuss situations where these restrictions may be somewhat relaxed.

9.A Improper Integrals

First, we extend the notion of integration to certain bounded functions on infinite intervals.

Definition: Let f be a function that is integrable on every closed subinterval $[a, T]$ of $[a, \infty)$. We define the *improper integral*

$$\int_a^\infty f(x)\,dx \doteq \lim_{T\to\infty} \int_a^T f(x)\,dx.$$

If this limit exists and is finite we say that $\int_a^\infty f(x)\,dx$ *converges*; otherwise we say that $\int_a^\infty f(x)\,dx$ *diverges*.

- For which values of p does $\int_1^\infty x^{-p}\,dx$ exist? To answer this question, we first compute the definite integral

$$\int_1^T \frac{dx}{x^p} = \begin{cases} \left[\dfrac{x^{1-p}}{1-p}\right]_1^T & \text{if } p \neq 1, \\ [\log|x|]_1^T & \text{if } p = 1. \end{cases}$$

But $\lim\limits_{T\to\infty} T^{1-p}$ exists only when $p \geq 1$. Also, $\lim\limits_{T\to\infty} \log T = \infty$. Thus

$$\int_1^\infty \frac{dx}{x^p} = \lim_{T\to\infty} \int_1^T \frac{dx}{x^p} = \begin{cases} \dfrac{1}{p-1} & \text{if } p > 1, \\ \infty & \text{if } p \leq 1. \end{cases}$$

Problem 9.1: Show that $\int_0^\infty \sin x\,dx$ diverges.

Problem 9.2: Let $f : \mathbb{R} \to \mathbb{R}$ be integrable on every closed interval. Show that

$$\int_a^\infty f(x)\,dx \in \mathcal{C} \Rightarrow \int_b^\infty f(x)\,dx \in \mathcal{C}$$

for any real numbers a and b.

Definition: Let f be a function that is integrable on every closed subinterval $[T, a]$ of $(-\infty, a]$. Define

$$\int_{-\infty}^a f(x)\,dx \doteq \lim_{T\to-\infty} \int_T^a f(x)\,dx.$$

Problem 9.3: Evaluate $\displaystyle\int_{-\infty}^0 xe^x\,dx$.

Q. Sometimes an explicit form for the antiderivative of an integrable function f is unavailable. Are there other ways to determine whether the improper integral $\int_a^\infty f(x)\,dx$ converges?

A. Yes. The following theorem, which is essentially a function analog of Theorem 2.3, will help us develop convergence tests for improper integrals.

Theorem 9.1 (Increasing Functions: Bounded $\iff$ Asymptotic Limit Exists): *Let f be a monotonic increasing function on $[a, \infty)$. Then f is bounded on $[a, \infty)$ $\iff$ $\lim\limits_{x\to\infty} f$ exists.*

Proof:

"$\Rightarrow$" The set $\mathcal{S} = \{f(x) : x \geq a\}$ is non-empty and bounded, so we may define $L = \sup \mathcal{S}$. Given $\epsilon > 0$, we know that $L - \epsilon$ is not a supremum of $\mathcal{S}$, so there exists $X \geq a$ such that $f(X) > L - \epsilon$. Then

$$x > X \Rightarrow L - \epsilon < f(X) \leq f(x) \leq L.$$

That is, $\lim\limits_{x\to\infty} f(x) = L$.

"$\Leftarrow$" Let $\epsilon = 1$. There exists $X \in \mathbb{R}$ such that

$$x > X \Rightarrow L - 1 < f(x) < L + 1.$$

For $a \leq x \leq X$, we have $f(a) \leq f(x) \leq f(X)$. Hence f is bounded by $\max(f(X), L + 1)$ and $\min(f(a), L - 1)$.

Corollary 9.1.1 (Improper Integrals of Non-Negative Functions): Let f be a non-negative function that is integrable on $[a, T]$ for all $T \geq a$. If there exists a bound B such that $\int_a^T f \leq B$ for all $T \geq a$, then $\int_a^\infty f$ converges.

Proof: Apply Theorem 9.1 to the increasing function $F(x) = \int_a^x f$, noting that

$$y \geq x \Rightarrow F(y) - F(x) = \int_x^y f \geq 0.$$

Definition: If $\int_a^\infty f$ converges, we say $\int_a^\infty f \in \mathcal{C}$, the set of convergent improper integrals. Otherwise, we say $\int_a^\infty f \in \mathcal{D}$, the set of divergent improper integrals.

Corollary 9.1.2 (Comparison Test): Suppose $0 \leq f(x) \leq g(x)$ and $\int_a^T f$ and $\int_a^T g$ exist for all $T \geq a$. Then

(i) $\int_a^\infty g \in \mathcal{C} \Rightarrow \int_a^\infty f \in \mathcal{C}$;

(ii) $\int_a^\infty f \in \mathcal{D} \Rightarrow \int_a^\infty g \in \mathcal{D}$.

Proof: Note that $0 \leq \int_a^T f \leq \int_a^T g$ and both integrals are monotonic increasing functions of T. Apply Theorem 9.1.

- To decide on whether

$$\int_1^\infty \frac{1}{1+x^3}\,dx$$

converges, we could first find $\displaystyle\int_1^T \frac{dx}{1+x^3}$ and then check that the limit as $T \to \infty$ exists. However, it is much easier to use Corollary 9.1.2 (i), noting that

$$0 \leq \frac{1}{1+x^3} \leq \frac{1}{x^2}$$

for all $x \geq 1$. That is,

$$\int_1^\infty \frac{1}{x^2}\,dx \in \mathcal{C} \Rightarrow \int_1^\infty \frac{1}{1+x^3}\,dx \in \mathcal{C}.$$

- We may use the previous result to establish that

$$\int_0^\infty \frac{1}{1+x^3}\,dx$$

exists, even though $1/x^2$ is not bounded (and hence $\int_0^\infty x^{-2}\,dx$ does not exist). This is seen by writing

$$\begin{aligned}\int_0^\infty \frac{1}{1+x^3}\,dx &= \lim_{T\to\infty}\int_0^T \frac{1}{1+x^3}\,dx = \lim_{T\to\infty}\left(\int_0^1 \frac{1}{1+x^3}\,dx + \int_1^T \frac{1}{1+x^3}\,dx\right)\\ &= \int_0^1 \frac{1}{1+x^3}\,dx + \int_1^\infty \frac{1}{1+x^3}\,dx.\end{aligned}$$

- To show that
$$\int_1^\infty e^{-x^2}\,dx$$
converges we note on $[1,\infty)$ that $x \le x^2$ so that $-x^2 \le -x$ and hence
$$0 \le e^{-x^2} \le e^{-x}.$$
On noting that
$$\int_1^\infty e^{-x}\,dx = \lim_{T\to\infty}\left[-e^{-x}\right]_1^T = \lim_{T\to\infty}\left(e^{-1}-e^{-T}\right) = \frac{1}{e},$$
we make use of the Comparison Test:
$$\int_1^\infty e^{-x}\,dx \text{ converges } \Rightarrow \int_1^\infty e^{-x^2}\,dx \text{ converges.}$$

- We may use the previous result to establish that
$$\int_0^\infty e^{-x^2}\,dx$$
converges:
$$\begin{aligned}\int_0^\infty e^{-x^2}\,dx = \lim_{T\to\infty}\int_0^T e^{-x^2}\,dx &= \lim_{T\to\infty}\left(\int_0^1 e^{-x^2}\,dx + \int_1^T e^{-x^2}\,dx\right)\\ &= \int_0^1 e^{-x^2}\,dx + \int_1^\infty e^{-x^2}\,dx.\end{aligned}$$

Problem 9.4: Use the fact that $\displaystyle\int_1^\infty \frac{1}{e^x}\,dx$ converges to show that $\displaystyle\int_1^\infty \frac{1}{x+e^x}\,dx$ converges.

Remark: A few useful improper integrals that can be used with the Comparison Test are listed in Table 9.1.

Corollary 9.1.3 (Limit Comparison Test): Let f and g be positive integrable functions satisfying
$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = L.$$

(i) For $0 < L < \infty$ we have $\int_a^\infty g \in \mathcal{C} \iff \int_a^\infty f \in \mathcal{C}$.

(ii) When $L = 0$ we can only say $\int_a^\infty g \in \mathcal{C} \Rightarrow \int_a^\infty f \in \mathcal{C}$.

Convergent	Divergent
$\displaystyle\int_1^\infty \frac{1}{x^p}\,dx \quad (p>1)$	$\displaystyle\int_1^\infty \frac{1}{x^p}\,dx \quad (p\le 1)$
$\displaystyle\int_0^\infty e^{-\alpha x}\,dx \quad (\alpha>0)$	$\displaystyle\int_0^\infty e^{-\alpha x}\,dx \quad (\alpha\le 0)$
$\displaystyle\int_{0^+}^1 \frac{1}{x^p}\,dx \quad (p<1)$	$\displaystyle\int_{0^+}^1 \frac{1}{x^p}\,dx \quad (p\ge 1)$
$\displaystyle\int_{0^+}^1 \log x\,dx$	$\displaystyle\int_0^{\pi/2^-} \tan x\,dx$

Table 9.1: Useful integrals for Comparison Test.

Proof:

(i) For x sufficiently large, we have (taking $\epsilon = L/2 > 0$),

$$\frac{L}{2}\,g(x) \le f(x) \le \frac{3L}{2}\,g(x).$$

From Corollary 9.1.2 we then deduce

$$\int_a^\infty g \in \mathcal{C} \iff \int_a^\infty f \in \mathcal{C}.$$

(ii) Exercise.

- Since

$$\lim_{x\to\infty} \frac{x^3-1}{x^3} = 1,$$

we see immediately that

$$\int_2^\infty \frac{1}{x^3}\,dx \in \mathcal{C} \Rightarrow \int_2^\infty \frac{1}{x^3-1}\,dx \in \mathcal{C}.$$

Remark: When $L = 0$, it is possible for $\int_1^\infty f \in \mathcal{C}$ but $\int_1^\infty g \in \mathcal{D}$. For example, consider $f(x) = 1/x^2$ and $g(x) = 1/x$, for which $\lim\limits_{x\to\infty} f(x)/g(x) = \lim\limits_{x\to\infty} 1/x = 0$.

Remark: Integration may thus be extended to bounded functions of x that converge to zero sufficiently fast as $x \to \infty$. We now see that it is even possible to extend our notion of improper Riemann integration to certain unbounded functions.

Definition: If f is integrable on $[a,t]$ for all $t \in (a,b)$ we define

$$\int_a^{b^-} f = \lim_{t\to b^-} \int_a^t f.$$

We say that $\int_a^{b^-} f$ *converges* if the limit exists; otherwise it *diverges*. Similarly, we define

$$\int_{a^+}^{b} f = \lim_{t\to a^+} \int_t^b f$$

if f is integrable on $[t,b]$ for all $t \in (a,b)$.

- Let

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{x}} & \text{if } 0 < x \le 1, \\ 0 & \text{if } x = 0. \end{cases}$$

We know that $\int_0^1 f$ does not exist, since f is not bounded. However, the improper integral $\int_{0^+}^1 f$ does exist:

$$\int_{0^+}^1 f = \lim_{t\to 0^+} \int_t^1 f = \lim_{t\to 0^+} \int_t^1 \frac{1}{\sqrt{x}}\,dx = \lim_{t\to 0^+} \left[2x^{\frac{1}{2}}\right]_t^1 = \lim_{t\to 0^+} \left(2 - 2t^{\frac{1}{2}}\right) = 2.$$

Remark: If f is Riemann integrable on $[a,b]$, we know from Corollary 5.5.3 that

$$\int_a^{b^-} f = \int_a^b f = \int_{a^+}^b f.$$

- For which values of p is $\int_{0^+}^{\infty} x^{-p}\,dx$ convergent?

Since

$$\int_{0^+}^1 \frac{1}{x^p}\,dx \in \mathcal{D} \text{ for } p \ge 1, \qquad \int_1^{\infty} \frac{1}{x^p}\,dx \in \mathcal{D} \text{ for } p \le 1,$$

we see that

$$\int_{0^+}^{\infty} \frac{1}{x^p}\,dx \in \mathcal{D} \quad \forall p.$$

Theorem 9.2 (Cauchy Criterion for Improper Integrals): *Let f be a function.*

(i) Suppose $\int_a^t f$ exists for all $t \in (a,b)$. Then $\int_a^{b-} f \in \mathcal{C} \iff \forall \epsilon > 0,\ \exists \delta > 0$ such that

$$x, y \in (b-\delta, b) \Rightarrow \left|\int_x^y f\right| < \epsilon;$$

(ii) Suppose $\int_a^T f$ exists for all $T > a$. Then $\int_a^\infty f \in \mathcal{C} \iff \forall \epsilon > 0,\ \exists T$ such that

$$T_2 \geq T_1 \geq T \Rightarrow \left| \int_{T_1}^{T_2} f \right| < \epsilon.$$

Proof:

(i) Apply the Cauchy Criterion, Corollary 3.1.2, to $F(t) = \int_a^t f$.

(ii) Exercise.

Definition: Let f be a function that is integrable on every finite interval $[c, d]$ of $\mathbb{R}$. If the improper integrals

$$\int_{-\infty}^{a} f(x)\, dx \qquad \text{and} \qquad \int_{a}^{\infty} f(x)\, dx$$

both converge for some $a \in \mathbb{R}$, then we say that the improper interval

$$\int_{-\infty}^{\infty} f(x)\, dx \doteq \int_{-\infty}^{a} f(x)\, dx + \int_{a}^{\infty} f(x)\, dx$$

converges.

Problem 9.5: Show that if $\int_{-\infty}^{\infty} f(x)\, dx$ exists for one $a \in \mathbb{R}$, it will exist for all $a \in \mathbb{R}$ and its value will not depend on the choice of a.

Remark: We cannot simplify this definition to

$$\int_{-\infty}^{\infty} f(x)\, dx = \lim_{T\to\infty} \int_{-T}^{T} f(x)\, dx.$$

For example, while

$$\lim_{T\to\infty} \int_{-T}^{T} x\, dx = \lim_{T\to\infty} 0 = 0,$$

the improper integrals

$$\int_{-\infty}^{a} x\, dx \qquad \text{and} \qquad \int_{a}^{\infty} x\, dx$$

do not converge for any $a \in \mathbb{R}$. That is, $\int_{-\infty}^{\infty} x$ diverges.

Remark: However, if $\int_{-\infty}^{\infty} f \; \exists$ then

$$\lim_{T\to\infty} \int_{-T}^{T} f \; \exists = \int_{-\infty}^{\infty} f$$

since, by the properties of limits,

$$\int_{-\infty}^{\infty} f = \lim_{T\to\infty} \int_{-T}^{a} f + \lim_{T\to\infty} \int_{a}^{T} f = \lim_{T\to\infty} \left[\int_{-T}^{a} f + \int_{a}^{T} f \right] = \lim_{T\to\infty} \int_{-T}^{T} f.$$

Problem 9.6: Evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx$.

Definition: Let f be defined and continuous everywhere on an interval $[a,b]$ except possibly at a point $c \in [a,b]$. If f is unbounded on $[a,b]$ we know that the Riemann integral of f on $[a,b]$ does not exist. Nevertheless, it is sometimes convenient to define the *improper integral*

$$\int_{a}^{b} f \doteq \lim_{t\to c^-} \int_{a}^{t} f + \lim_{t\to c^+} \int_{t}^{b} f.$$

If both limits exist, we say that the improper Riemann integral $\int_a^b f$ converges.

Problem 9.7: Do the following improper integrals converge or diverge? Evaluate those that converge.

(a)
$$\int_{-2}^{2} \frac{1}{(x-1)^{2/3}}\,dx.$$

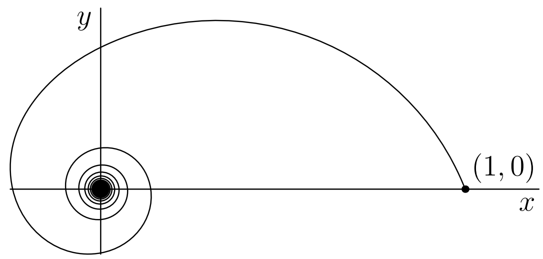
(b)
$$\int_{-2}^{2} \frac{1}{(x-1)^{4/3}}\,dx.$$

Problem 9.8: Use the fact that $\displaystyle\int_{0^+}^{1} \frac{1}{\sqrt{x}}\,dx$ converges to show that $\displaystyle\int_{0^+}^{1} \frac{e^{-x^2}}{\sqrt{x}}\,dx$ converges.

Problem 9.9: Use the fact that $\displaystyle\int_{0^+}^{1} \frac{1}{x}\,dx$ diverges to show that $\displaystyle\int_{0^+}^{1} \frac{1}{x\sin^2 x}\,dx$ diverges.

Problem 9.10: In polar coordinates, $(x, y) = (r\cos\theta, r\sin\theta)$, consider the curve $r(\theta) = \dfrac{1}{1+\theta}$ for $\theta \in [0, \infty)$.

(a) Sketch this curve on an x–y graph.



(b) Express the arc length of this curve as an improper integral.

Since $r'(\theta) = -1/(1+\theta)^2$,

$$\int_0^\infty \sqrt{\frac{1}{(1+\theta)^4} + \frac{1}{(1+\theta)^2}}\, d\theta = \int_1^\infty \sqrt{\frac{1}{u^4} + \frac{1}{u^2}}\, du = \int_1^\infty \frac{1}{u}\sqrt{\frac{1}{u^2} + 1}\, du.$$

(c) Does this curve have finite length? Justify your answer.

No, the integral diverges:

$$0 \le \frac{1}{u} \le \frac{1}{u}\sqrt{\frac{1}{u^2} + 1}$$

and $\displaystyle\int_1^\infty \frac{1}{u}\, du$ diverges, so we know from the Comparison Test that

$$\int_1^\infty \frac{1}{u}\sqrt{\frac{1}{u^2} + 1}\, du$$

also diverges.

9.B Infinite Series

Definition: Let $S_n = \sum_{k=1}^n a_k$. If $\lim\limits_{n\to\infty} S_n$ exists and equals a real number S, we say that the *infinite series*

$$\sum_{k=1}^\infty a_k$$

converges, with sum S, and write $\sum_{k=1}^\infty a_k \in \mathcal{C}$. Otherwise, we say $\sum_{k=1}^\infty a_k$ is *divergent* and write $\sum_{k=1}^\infty a_k \in \mathcal{D}$.

Definition: The finite sum $S_n = \sum_{k=1}^n a_k$ is a *partial sum* of the series $\sum_{k=1}^\infty a_k$.

Problem 9.11: Prove that the *geometric series* $\sum_{k=0}^{\infty} r^k$ converges if and only if $|r| < 1$, with sum $1/(1-r)$. Hint: Show that

$$S_n = \sum_{k=0}^{n} r^k = \frac{1-r^{n+1}}{1-r}$$

by considering the telescoping sum $rS_n - S_n$.

- Consider

$$0.\bar{4} = 0.444\ldots = \lim_{n\to\infty} \sum_{k=1}^{n} 4\left(\frac{1}{10^k}\right) = \left(\frac{4}{10}\right) \lim_{n\to\infty} \sum_{k=0}^{n} \left(\frac{1}{10}\right)^k = \left(\frac{4}{10}\right) \frac{1}{1-\frac{1}{10}} = \frac{4}{9}.$$

Theorem 9.3 (Cauchy Criterion for Infinite Series): *The infinite series* $\sum_{k=1}^{\infty} a_k$ *converges if and only if for each* $\epsilon > 0$*, there exists* $N \in \mathbb{N}$ *such that*

$$m > n > N \Rightarrow \left| \sum_{k=n}^{m} a_k \right| < \epsilon.$$

Proof: Apply the Cauchy Criterion to $\{S_n\}_{n=1}^{\infty}$, noting that

$$\left| \sum_{k=n}^{m} a_k \right| = |S_m - S_{n-1}|.$$

- Thus

$$\sum_{k=1}^{\infty} (-1)^k \in \mathcal{D} \text{ since } |S_{n+1} - S_n| = |a_{n+1}| = 1 \quad \forall n.$$

- Recall that the *harmonic series* is divergent:

$$\sum_{k=1}^{\infty} \frac{1}{k} \in \mathcal{D} \text{ since } |S_{2n} - S_n| = \left| \sum_{k=n+1}^{2n} \frac{1}{k} \right| = \underbrace{\frac{1}{n+1} + \ldots + \frac{1}{2n}}_{n \text{ terms}} > n\left(\frac{1}{2n}\right) = \frac{1}{2}.$$

- However,

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$$

converges to the value 1 since

$$S_n = \sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=2}^{n+1} \frac{1}{k} = 1 - \frac{1}{n+1}.$$

Problem 9.12: Let $\alpha > 0$. Evaluate

$$\sum_{k=0}^{\infty} \frac{1}{(\alpha+k)(\alpha+k+1)}.$$

We can compute the partial sums using partial fraction decomposition:

$$\begin{aligned}\sum_{k=0}^{n} \frac{1}{(\alpha+k)(\alpha+k+1)} &= \sum_{k=0}^{n} \frac{1}{\alpha+k} - \sum_{k=0}^{n} \frac{1}{\alpha+k+1} \\ &= \sum_{k=0}^{n} \frac{1}{\alpha+k} - \sum_{k=1}^{n+1} \frac{1}{\alpha+k} \\ &= \frac{1}{\alpha} - \frac{1}{\alpha+n+1}.\end{aligned}$$

As $n \to \infty$, the sum converges to $1/\alpha$.

Theorem 9.4 (Divergence Test): *If* $\displaystyle\sum_{k=1}^{\infty} a_k \in \mathcal{C}$ *then* $\displaystyle\lim_{n\to\infty} a_n = 0$.

Proof: In terms of the convergent sequence of partial sums $\{S_n\}_{n=1}^{\infty}$ we may express

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1}) = \lim_{n\to\infty} S_{n-1} - \lim_{n\to\infty} S_n = 0.$$

Remark: The contrapositive of Theorem 9.4 states

$$\lim_{n\to\infty} a_n \neq 0 \Rightarrow \sum_{k=1}^{\infty} a_k \in \mathcal{D}.$$

However,

$$\lim_{n\to\infty} a_n = 0 \not\Rightarrow \sum_{k=1}^{\infty} a_k \in \mathcal{C}.$$

For example, $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k} \in \mathcal{D}$ even though $\displaystyle\lim_{k\to\infty} \frac{1}{k} = 0$.

Theorem 9.5 (Non-Negative Terms: Convergence $\iff$ Bounded Partial Sums): *If* $a_k \geq 0$ *and* $S_n = \sum_{k=1}^{n} a_k$ *then* $\sum_{k=1}^{\infty} a_k \in \mathcal{C} \iff \{S_n\}_{n=1}^{\infty}$ *is a bounded sequence.*

Proof: Since $S_{n+1} = S_n + a_{n+1} \geq S_n$ we know that $\{S_n\}_{n=1}^{\infty}$ is an increasing sequence. It then follows from Theorem 2.3 that $\{S_n\}_{n=1}^{\infty}$ is convergent $\iff$ $\{S_n\}_{n=1}^{\infty}$ bounded.

Corollary 9.5.1 (Comparison Test): If $0 \le a_k \le b_k$ for $k \in \mathbb{N}$ then

(i) $\displaystyle\sum_{k=1}^{\infty} b_k \in \mathcal{C} \Rightarrow \sum_{k=1}^{\infty} a_k \in \mathcal{C}$;

(ii) $\displaystyle\sum_{k=1}^{\infty} a_k \in \mathcal{D} \Rightarrow \sum_{k=1}^{\infty} b_k \in \mathcal{D}$.

Proof:

(i) Let $S_n = \sum_{k=1}^{n} a_k$ and $T_n = \sum_{k=1}^{n} b_k$. Since $0 \le S_n \le T_n$,

$$\sum_{k=1}^{\infty} b_k \in \mathcal{C} \Rightarrow \{T_n\}_{n=1}^{\infty} \text{ bounded } \Rightarrow \{S_n\}_{n=1}^{\infty} \text{ bounded } \Rightarrow \sum_{k=1}^{\infty} a_k \in \mathcal{C}.$$

(ii) This is just the contrapositive of (i).

Remark: The condition "$0 \le a_k$" in Corollary 9.5.1 cannot be dropped. Consider the counterexample given by $a_k = -1$, $b_k = 0$.

Corollary 9.5.2 (Limit Comparison Test): Suppose $a_k \ge 0$ and $b_k > 0$ for all $k \in \mathbb{N}$ and $\lim\limits_{k\to\infty} a_k/b_k = L$. Then

(i) if $0 < L < \infty$: $\displaystyle\sum_{k=1}^{\infty} a_k \in \mathcal{C} \iff \sum_{k=1}^{\infty} b_k \in \mathcal{C}$;

(ii) if $L = 0$: $\displaystyle\sum_{k=1}^{\infty} b_k \in \mathcal{C} \Rightarrow \sum_{k=1}^{\infty} a_k \in \mathcal{C}$.

Proof:

(i) This follows from Corollary 9.5.1 since for all sufficiently large k,

$$0 < \frac{L}{2} < \frac{a_k}{b_k} < \frac{3L}{2} \Rightarrow 0 < \left(\frac{L}{2}\right) b_k < a_k < \left(\frac{3L}{2}\right) b_k.$$

(ii) If $L = 0$ then for sufficiently large k, $0 \le a_k/b_k < \epsilon = 1 \Rightarrow 0 \le a_k < b_k$. Apply Corollary 9.5.1.

Remark: When $L = 0$ it is possible that $\sum_{k=1}^{\infty} a_k \in \mathcal{C}$ but $\sum_{k=1}^{\infty} b_k \in \mathcal{D}$. Consider

$$a_k = \frac{1}{k(k+1)}, \qquad b_k = \frac{1}{k}, \qquad \lim_{k\to\infty} \frac{a_k}{b_k} = \lim_{k\to\infty} \frac{1}{k+1} = 0.$$

- Since
$$\lim_{k\to\infty} \frac{k^2}{k(k+1)} = 1,$$
we see that
$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} \in \mathcal{C} \Rightarrow \sum_{k=1}^{\infty} \frac{1}{k^2} \in \mathcal{C}.$$

- Since
$$\lim_{k\to\infty} \frac{2^k-1}{2^k} = 1,$$
we see that
$$\sum_{k=1}^{\infty} \frac{1}{2^k} \in \mathcal{C} \Rightarrow \sum_{k=1}^{\infty} \frac{1}{2^k-1} \in \mathcal{C}.$$

Corollary 9.5.3 (Ratio Comparison Test): If $a_k > 0$ and $b_k > 0$ and
$$\frac{a_{k+1}}{a_k} \le \frac{b_{k+1}}{b_k}$$
for all $k \ge N$, then

(i) $\displaystyle\sum_{k=1}^{\infty} b_k \in \mathcal{C} \Rightarrow \sum_{k=1}^{\infty} a_k \in \mathcal{C}$;

(ii) $\displaystyle\sum_{k=1}^{\infty} a_k \in \mathcal{D} \Rightarrow \sum_{k=1}^{\infty} b_k \in \mathcal{D}$.

Proof: For $k \ge N$ we have
$$\frac{a_{k+1}}{a_k} \le \frac{b_{k+1}}{b_k} \Rightarrow \frac{a_{k+1}}{b_{k+1}} \le \frac{a_k}{b_k} \Rightarrow \frac{a_k}{b_k} \le \frac{a_N}{b_N} \doteq M.$$
Thus $0 < a_k \le M b_k$ for all $k \ge N$ and the result follows from Corollary 9.5.1.

Corollary 9.5.4 (Ratio Test): Suppose $a_k > 0$.

(i) If $\exists$ a number $x \in (0,1)$ such that $\frac{a_{k+1}}{a_k} \le x$ for all $k \ge N$, then $\sum_{k=1}^{\infty} a_k \in \mathcal{C}$.

(ii) If $\exists$ a number $x \ge 1$ such that $\frac{a_{k+1}}{a_k} \ge x$ for all $k \ge N$, then $\sum_{k=1}^{\infty} a_k \in \mathcal{D}$.

Proof: Let $b_k = x^k$. Then $\frac{b_{k+1}}{b_k} = x$ and recall that
$$\begin{cases} \displaystyle\sum_{k=1}^{\infty} b_k \in \mathcal{C} & \text{if } |x| < 1, \\ \displaystyle\sum_{k=1}^{\infty} b_k \in \mathcal{D} & \text{if } |x| \ge 1. \end{cases}$$
Apply Corollary 9.5.3.

Corollary 9.5.5 (Limit Ratio Test): Suppose $a_k > 0$ for all $k \in \mathbb{N}$ and

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = r.$$

Then

(i) $0 \le r < 1 \Rightarrow \sum_{k=1}^{\infty} a_k \in \mathcal{C}$,

(ii) $r > 1 \Rightarrow \sum_{k=1}^{\infty} a_k \in \mathcal{D}$,

(iii) $r = 1 \Rightarrow ?$

Proof: (i) Choose $\epsilon = (1-r)/2$. Then for sufficiently large k, we know that

$$\frac{a_{k+1}}{a_k} < r + \epsilon = \frac{1+r}{2} < 1.$$

Apply the Ratio Test.
(ii) Exercise.

Problem 9.13: Find examples corresponding to each of the three cases in Corollary 9.5.5.

The next theorem, illustrated in Fig. 9.1, sheds some light on why $\sum_{k=1}^{\infty} \frac{1}{k^p}$ and $\int_1^{\infty} \frac{dx}{x^p}$ both diverge for $p = 1$ and converge for $p = 2$.

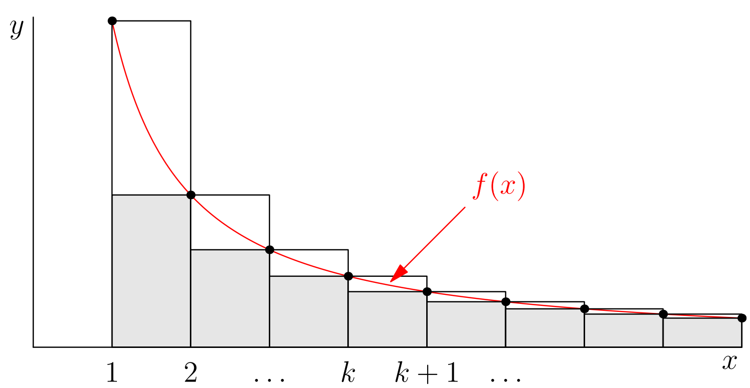


Figure 9.1: The Integral Test

Theorem 9.6 (Integral Test): *Let f be decreasing and non-negative on $[1, \infty)$. Then*

$$\sum_{k=1}^{\infty} f(k) \in \mathcal{C} \iff \int_1^{\infty} f \in \mathcal{C}.$$

Proof: Since f is monotonic we know from Theorem 5.7 that it is integrable on $[k, k+1]$. For $x \in [k, k+1]$ we have $f(k) \geq f(x) \geq f(k+1)$. Corollary 5.5.1 implies

$$f(k) \cdot 1 \geq \int_k^{k+1} f(x)\, dx \geq f(k+1) \cdot 1.$$

We then sum this result $k = 1$ to $k = n$ to obtain

$$S_n \doteq \sum_{k=1}^{n} f(k) \geq \int_1^{n+1} f \geq \sum_{k=1}^{n} f(k+1) = \sum_{k=2}^{n+1} f(k) = S_{n+1} - f(1).$$

"$\Rightarrow$" By Corollary 9.1.1,

$$\lim_{n\to\infty} S_n \, \exists \Rightarrow \lim_{n\to\infty} \int_1^{n+1} f \, \exists \Rightarrow \lim_{T\to\infty} \int_1^{T} f \, \exists$$

since $\int_1^T f$ is an increasing function of T.

"$\Leftarrow$"

$$\lim_{T\to\infty} \int_1^{T} f \, \exists \Rightarrow \lim_{n\to\infty} \int_1^{n+1} f \, \exists \Rightarrow \{S_{n+1}\}_{n=1}^{\infty} \text{ bounded} \Rightarrow \{S_n\}_{n=1}^{\infty} \text{ bounded}.$$

But $f(x) \geq 0 \Rightarrow \{S_n\}_{n=1}^{\infty}$ is increasing $\Rightarrow \lim\limits_{n\to\infty} S_n \, \exists$ by Theorem 2.3.

Problem 9.14: Use the Integral Test to show that

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

converges if $p > 1$.

Consider the improper integral

$$\int_1^{\infty} \frac{1}{x^p}\, dx,$$

which converges only when $p > 1$. By the Integral Test, we therefore know that

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

converges only when $p > 1$.

Problem 9.15: Use the Integral Test to show that

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

diverges.

We first consider the improper integral

$$\int_2^{\infty} \frac{1}{x \log x}\, dx$$

on letting $u = \log x$, the integral becomes

$$\int_{\log 2}^{\infty} \frac{1}{u}\, du,$$

which diverges. The Integral Test then tells us that

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

diverges as well.

Remark: While the Integral Test is useful for establishing the convergence of a series, it does not tell us anything about its value. For example, $\displaystyle\int_1^{\infty} \frac{1}{x^2}\, dx = 1$ but it can be shown that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} = \pi^2/6$. Often a closed-form expression for a series is unavailable and one must resort to numerical computation of the partial sums up to a certain value of n. The following related theorem can be used to estimate the error in such approximations.

Corollary 9.6.1 (Remainder Estimate): Let f be decreasing and non-negative on $[1, \infty)$. Then the *remainder* $\displaystyle\sum_{k=n+1}^{\infty} f(k)$ of $\displaystyle\sum_{k=1}^{\infty} f(k)$ that results on truncating the series after n terms satisfies

$$\int_{n+1}^{\infty} f \leq \sum_{k=n+1}^{\infty} f(k) \leq \int_n^{\infty} f.$$

Proof: In the proof of the Integral Test we saw that

$$f(k+1) \leq \int_k^{k+1} f$$

On summing from $k = n$ to ∞ we thus find that

$$\sum_{k=n+1}^{\infty} f(k) = \sum_{k=n}^{\infty} f(k+1) \le \sum_{k=n}^{\infty} \int_k^{k+1} f = \int_n^{\infty} f.$$

We also saw that

$$\int_k^{k+1} f \le f(k).$$

On summing from $k = n+1$ to ∞ we obtain

$$\int_{n+1}^{\infty} f = \sum_{k=n+1}^{\infty} \int_k^{k+1} f \le \sum_{k=n+1}^{\infty} f(k).$$

- The partial sum $S_{10} = \displaystyle\sum_{k=1}^{10} \frac{1}{k^2} \approx 1.5498$ underestimates $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$ by a remainder that lies between

  $$\int_{11}^{\infty} \frac{1}{x^2}\, dx = \frac{1}{11}$$

  and

  $$\int_{10}^{\infty} \frac{1}{x^2}\, dx = \frac{1}{10}.$$

  Indeed, we see that the difference between the exact value $\sum_{k=1}^{\infty} \frac{1}{k^2} = \pi^2/6$ and S_{10} is approximately 0.095, which indeed lies between 1/11 and 1/10.

Q. What happens when some of the terms a_k are negative?

A. In these situations, the following concept is sometimes helpful.

Definition: A series $\sum_{k=1}^{\infty} a_k$ is *absolutely convergent* if $\sum_{k=1}^{\infty} |a_k|$ is convergent. In this case, we write

$$\sum_{k=1}^{\infty} a_k \in \mathrm{Abs}\,\mathcal{C}.$$

- The sequence $\displaystyle\sum_{k=1}^{\infty} \frac{\sin k}{k^2}$ is absolutely convergent by the Comparison Test since $|\sin k| \le 1$ and $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$ is convergent. That is, $\displaystyle\sum_{k=1}^{\infty} \frac{|\sin k|}{k^2}$ is convergent.

The following theorem establishes that the original series $\sum_{k=1}^{\infty} \frac{\sin k}{k^2}$ is itself convergent.

Theorem 9.7 (Absolute Convergence): *An absolutely convergent series is convergent.*

Proof: By the Triangle Inequality we know that

$$\left| \sum_{k=n}^{m} a_k \right| \leq \sum_{k=n}^{m} |a_k| \, .$$

The result then follows by applying the Cauchy Criterion, Theorem 9.3.

Remark: The converse of Theorem 9.7 need not be true: the alternating harmonic series $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ is convergent but not absolutely convergent: the harmonic series $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges.

Definition: A series is *conditionally convergent* if is convergent but not absolutely convergent.

- The alternating harmonic series $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ is conditionally convergent.

Problem 9.16: Show that the alternating series $\sum_{k=1}^{\infty} \frac{(-1)^k}{\sqrt{k}}$ is conditionally convergent.

Problem 9.17: Show that the alternating series $\sum_{k=1}^{\infty} \frac{(-1)^k \log k}{\sqrt{k}}$ is conditionally convergent.

The following *summation by parts* formula is useful for developing tests of conditional convergence.

Lemma 9.1 (Abel's Lemma):

$$\sum_{k=n}^{m} f_k(g_{k+1} - g_k) = f_{m+1}g_{m+1} - f_n g_n - \sum_{k=n}^{m} (f_{k+1} - f_k) g_{k+1}$$

Proof:

$$\begin{aligned}\sum_{k=n}^{m} f_k(g_{k+1}-g_k) &= \sum_{k=n}^{m} f_k g_{k+1} - \sum_{k=n}^{m} f_k g_k = \sum_{k=n}^{m} f_k g_{k+1} + f_{m+1}g_{m+1} - \sum_{k=n+1}^{m+1} f_k g_k - f_n g_n \\ &= f_{m+1}g_{m+1} - f_n g_n + \sum_{k=n}^{m} f_k g_{k+1} - \sum_{k=n}^{m} f_{k+1} g_{k+1} \\ &= f_{m+1}g_{m+1} - f_n g_n - \sum_{k=n}^{m} (f_{k+1}-f_k) g_{k+1}.\end{aligned}$$

Definition: An *alternating series* is of the form

$$\sum_{k=1}^{\infty} (-1)^k a_k,$$

where $a_k \geq 0$ for $k \in \mathbb{N}$.

Definition: A sequence $\{a_k\}_{k=1}^{\infty}$ is of *bounded variation* if $\sum_{k=1}^{\infty} |a_{k+1}-a_k| \in \mathcal{C}$.

Problem 9.18: Show that a monotonic bounded sequence is always of bounded variation.

Remark: By Theorem 9.7, if a sequence $\{a_k\}_{k=1}^{\infty}$ is of bounded variation, then $\sum_{k=1}^{\infty}(a_{k+1}-a_k) \in \mathcal{C}$. It follows that $\lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = a_1 + \lim_{n\to\infty} \sum_{k=1}^{n}(a_{k+1}-a_k)$ exists.

Theorem 9.8 (Dirichlet Test): *Let* $S_n = \sum_{k=1}^{n} b_k$. *Suppose*

(i) $\{S_n\}_{n=1}^{\infty}$ *is a bounded sequence.*

(ii) $\lim_{n\to\infty} a_n = 0$ *and* $\{a_k\}_{k=1}^{\infty}$ *is of bounded variation.*

Then $\sum_{k=1}^{\infty} a_k b_k \in \mathcal{C}$.

Proof: We know from (i) that there exists a number $B > 0$ such that $|S_n| \leq B$ for all $n \in \mathbb{N}$. Given $\epsilon > 0$, we know from (ii) and the Cauchy Criterion for Infinite Series that there exists $N > 0$ such that

$$n > N \Rightarrow |a_n| < \frac{\epsilon}{3B}$$

and

$$m > n > N \Rightarrow \sum_{k=n}^{m} |a_{k+1} - a_k| < \frac{\epsilon}{3B}.$$

Using Abel's Lemma we then find that

$$\begin{aligned}
\left|\sum_{k=n}^{m} a_k b_k\right| &= \left|\sum_{k=n}^{m} a_k(S_k - S_{k-1})\right| \\
&= \left|a_{m+1}S_m - a_n S_{n-1} - \sum_{k=n}^{m}(a_{k+1} - a_k)S_k\right| \\
&\le |a_{m+1}|\, B + |a_n|\, B + B\sum_{k=n}^{m} |a_{k+1} - a_k| \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

On applying the Cauchy Criterion for Infinite Series again, we see that $\sum_{k=1}^{\infty} a_k b_k$ converges.

Corollary 9.8.1 (Leibniz Alternating Series Test): The alternating series $\sum_{k=1}^{\infty}(-1)^k a_k$ is convergent if the sequence $\{a_k\}_{k=1}^{\infty}$ decreases monotonically to 0.

Proof: Apply the Dirichlet Test with $b_k = (-1)^k$, using the fact that the bounded monotone sequence a_k is of bounded variation.

- Applying the Leibniz Alternating Series Test with $a_k = \dfrac{1}{k}$ and $b_k = (-1)^k$ shows that $\sum_{k=1}^{\infty} \dfrac{(-1)^k}{k} \in \mathcal{C}$.

- The alternating series $\sum_{k=1}^{\infty}(-1)^k \dfrac{k^2}{k^3+1}$ converges since the function $f(x) = \dfrac{x^2}{x^3+1}$ decreases monotonically to zero on $[2,\infty)$.

Problem 9.19: Prove the Leibniz Remainder Estimate: if a_k is a monotone decreasing sequence that converges to 0, the alternating sum $S = \sum_{k=1}^{\infty}(-1)^k a_k$ can be approximated by the nth partial sum $S_n = \sum_{k=1}^{n}(-1)^k a_k$ with an error of at most a_{n+1}:

$$|S_n - S| \le |S_{n+1} - S_n| = |a_{n+1}|\,.$$

That is,

$$\left|\sum_{k=n+1}^{\infty}(-1)^k a_k\right| \le a_{n+1}.$$

Let $S_n = \sum_{k=1}^{n}(-1)^k a_k$ and $S = \sum_{k=1}^{\infty}(-1)^k a_k$. Since $\{S_{2n-1}\}_{n=1}^{\infty}$ is increasing and $\{S_{2n}\}_{n=1}^{\infty}$ is decreasing, we see that $S_{2n-1} \le S \le S_{2n}$ for all $n \ge 1$. Since $a_n - a_{n+1} \ge 0$, we then find that

$$0 \le S - S_{2n-1} = \sum_{k=2n}^{\infty}(-1)^k a_k = a_{2n} - (a_{2n+1} - a_{2n+2}) + \ldots \le a_{2n}$$

and

$$0 \le S_{2n} - S = -\sum_{k=2n+1}^{\infty}(-1)^k a_k = a_{2n+1} - (a_{2n+2} - a_{2n+3}) + \ldots \le a_{2n+1}.$$

For either odd or even n the error in approximating S by S_n is seen to be less than the magnitude of the first neglected term:

$$\left|\sum_{k=n+1}^{\infty}(-1)^k a_k\right| = |S - S_n| \le a_{n+1}.$$

Remark: If the magnitude of the terms of an alternating Taylor series decreases monotonically to zero, it is much easier to use the Leibniz Remainder Estimate rather than explicitly estimating the remainder using Taylor's Theorem: the error is simply less than the magnitude of the very next term!

- Evaluate $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n}{n!}$ to three decimal places.

Problem 9.20:

(a) Suppose for a sequence of positive numbers $\{a_k\}_{k=1}^{\infty}$ that $\sqrt[k]{a_k} \le q$ for all sufficiently large k. If $0 < q < 1$, use the Comparison Test to show that $\displaystyle\sum_{k=1}^{\infty} a_k$ converges.

For sufficiently large k we are given that

$$0 \le a_k \le q^k.$$

Since the geometric series $\displaystyle\sum_{k=1}^{\infty} q^k$ converges for $0 < q < 1$, we know that $\displaystyle\sum_{k=1}^{\infty} a_k$ converges.

(b) Prove the *Root Test*: Suppose $\lambda = \lim_{k\to\infty} \sqrt[k]{a_k}$ exists. If $\lambda < 1$, show that $\sum_{k=1}^{\infty} a_k$ converges.

This would imply for sufficiently large k that

$$\sqrt[k]{a_k} < \frac{1+\lambda}{2} < 1.$$

(e.g. $\epsilon = (1-\lambda)/2 \Rightarrow \lambda + \epsilon = (1+\lambda)/2$) and hence part (a) implies that $\sum_{k=1}^{\infty} a_k$ converges.

(c) If $\lim_{k\to\infty} \sqrt[k]{a_k} > 1$, show that $\sum_{k=1}^{\infty} a_k$ diverges.

This would imply for sufficiently large k that $a_k > 1$, which violates the convergence condition $\lim_{k\to\infty} a_k = 0$ (Theorem 9.4).

(d) Does $\sum_{k=2}^{\infty} \frac{1}{(\log k)^k}$ converge or diverge?

Since $\lim_{k\to\infty} \frac{1}{\log k} = 0 < 1$, we know from the Root Test that the series converges.

Problem 9.21: In this problem we show that the function

$$f(x) = \sum_{n=0}^{\infty} 4^{-n} \langle 4^n x \rangle ,$$

where $\langle x \rangle$ denotes the distance of x to a nearest integer, is continuous for all real x but differentiable nowhere.

(a) Sketch the graph of $\langle x \rangle$.

(b) Find the minimum and maximum values of $\{\langle x \rangle : x \in \mathbb{R}\}$.

(c) Use the Comparison Test to show that series expression given for $f(x)$ converges at all x.

(d) Establish for all real x and a the overestimate $|\langle x \rangle - \langle a \rangle| \le 1$.

(e) Show that $|\langle x \rangle - \langle a \rangle| \le \min(|x-a|, 1)$.

Hint: let X be a nearest integer to x and A be a nearest integer to a. First establish the result in the case where $X = A$. Then consider $y = x + A - X$.

(f) If $|x-a| < 4^{-m}$ for $m \in \mathbb{N}$, use the fact that

$$\begin{aligned} |f(x) - f(a)| &\le \sum_{n=0}^{\infty} 4^{-n} |\langle 4^n x \rangle - \langle 4^n a \rangle| \le \sum_{n=0}^{\infty} 4^{-n} \min(|4^n x - 4^n a|, 1) \\ &< \sum_{n=0}^{\infty} 4^{-n} \min(4^{n-m}, 1) = \sum_{n=0}^{\infty} \min(4^{-m}, 4^{-n}) = \sum_{n=0}^{m} 4^{-m} + \sum_{n=m+1}^{\infty} 4^{-n} \end{aligned}$$

to show that f is continuous at any $a \in \mathbb{R}$.

(g) Show that $f(x+1) = f(x)$ for all $x \in \mathbb{R}$.
(h) Let us represent $a \in [0,1)$ with the base-4 digits $a_i \in \{0,1,2,3\}$ for $i \in \mathbb{N}$:

$$a = 0.a_1a_2a_3\ldots,$$

avoiding, for uniqueness, infinite patterns of repeating threes. Let

$$h_m = \begin{cases} 4^{-m} & \text{if } a_m = 0 \text{ or } 1; \\ -4^{-m} & \text{if } a_m = 2 \text{ or } 3. \end{cases}$$

If $n \geq m$ show that

$$\langle 4^n(a+h_m)\rangle - \langle 4^n(a)\rangle = 0.$$

(i) We now compute

$$\frac{f(a+h_m) - f(a)}{h_m}$$

and show that the result diverges as $m \to \infty$. This will require evaluating $\langle 4^n(a+h_m)\rangle - \langle 4^n a\rangle$, which can be done as follows.

If $n < m$ express $4^n a = N + x_n$ where $x_n = 0.a_{n+1}a_{n+2}\ldots a_m\ldots$. We know that $4^n h_m = 4^{n-m} \in (0,1/4]$ when $a_m = 0$ or 2 and $4^n h_m = -4^{n-m} \in [-1/4,0)$ when $a_m = 1$ or 3. The digits $0,1,2,3$ in the mth place of x_n thus get mapped by $x_n \to x_n + 4^n h_m$ to the digits $1,0,3,2$ respectively. This means that x_n and $x_n + 4^n h_m$ either both belong to $[0,1/2)$ or both belong to $[1/2,1)$, so that

$$\langle 4^n(a+h_m)\rangle - \langle 4^n a\rangle = \langle x_n + 4^n h_m\rangle - \langle x_n\rangle = \begin{cases} 4^n h_m & \text{if } a_m = 0 \text{ or } 2, \\ \\ -4^n h_m & \text{if } a_m = 1 \text{ or } 3. \end{cases}$$

Show using this result together with part (h) that

$$\lim_{m\to\infty} \frac{f(a+h_m) - f(a)}{h_m}$$

does not exist at any real a, even though f is continuous everywhere.

9.C Power Series

Let f be a function with derivatives of all orders at a point $a \in \mathbb{R}$. Recall that the Taylor expansion of f at x about a to order n is given by

$$f(x) = f(a) + f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + \ldots + f^{(n-1)}(a)\frac{(x-a)^{n-1}}{(n-1)!} + R_n(x).$$

Q. Is it always true that $\lim_{n\to\infty} R_n(x) = 0$, so that the *Taylor series*

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

converges to $f(x)$?

A. No, although this statement holds for polynomials, e^x, $\sin x$, $\cos x$, and many other familiar functions, the Taylor series need not in general converge. Even when it does converge, the limit is not necessarily $f(x)$.

Definition: A *power series* about the *expansion point* a is an infinite series of the form $\sum_{k=0}^{\infty} c_k(x-a)^k$, where the coefficients c_k are independent of x. In the context of power series, we interpret x^0 to mean 1 for all x, even at $x = 0$.

Remark: It is often convenient to shift the variable x by a, so that the power series takes the form $\sum_{k=0}^{\infty} c_k x^k$.

Remark: A power series always converges at its expansion point.

- We have seen for $x \in (-1, 1)$ that the geometric series

$$\sum_{k=0}^{\infty} x^k$$

converges to the function $\frac{1}{1-x}$.

- If we apply the Limit Ratio Test to the series

$$\sum_{k=0}^{\infty} k!\, x^k,$$

we see for $x \neq 0$ that

$$\lim_{k\to\infty} \frac{\left|(k+1)!\, x^{k+1}\right|}{|k!\, x^k|} = \lim_{k\to\infty} (k+1)\,|x| = \infty.$$

The series thus converges only at $x = 0$.

- In contrast, the Limit Ratio Test tells us that the series

$$\sum_{k=0}^{\infty} \frac{x^k}{k!}$$

converges for all real x:

$$\lim_{k\to\infty} \frac{|x|^{k+1}}{(k+1)!} \cdot \frac{k!}{|x|^k} = \lim_{k\to\infty} \frac{|x|}{(k+1)} = 0 < 1.$$

Problem 9.22: For what values of x does the series $\sum_{k=1}^{\infty} \frac{(x-1)^k}{k}$ converge?

The Limit Ratio Test tells us that the series converges if

$$\lim_{k\to\infty} \frac{|x-1|^{k+1}}{(k+1)} \cdot \frac{k}{|x-1|^k} = \lim_{k\to\infty} \frac{k}{(k+1)} |x-1| = |x-1| < 1,$$

that is, when $-1 < x-1 < 1$ or in other words, $0 < x < 2$. Furthermore, we see that the series converges at $x = 0$ (where it reduces to the alternating harmonic series and diverges at $x = 2$ (where it reduces to the harmonic series). The Limit Ratio Test tell us that the series diverges for $|x| > 2$. The interval of convergence is thus $[0, 2)$.

Theorem 9.9 (Radius of Convergence): *For each power series* $\sum_{k=0}^{\infty} c_k x^k$ *there exists a number* R*, called the* radius of convergence*, with* $0 \le R \le \infty$*, such that*

$$\sum_{k=0}^{\infty} c_k x^k \in \begin{cases} \mathrm{Abs}\,\mathcal{C} & \text{if } |x| < R, \\ \mathcal{D} & \text{if } |x| > R, \\ ? & \text{if } |x| = R. \end{cases}$$

Proof: Suppose $\sum_{k=0}^{\infty} c_k x_0^k \in \mathcal{C}$ for some real value x_0. Theorem 9.4 implies $\lim\limits_{k\to\infty} c_k x_0^k = 0$; in particular, there exists a positive number M such that $\left|c_k x_0^k\right| \le M$ for all k. Now

$$\left|c_k x^k\right| = \left|c_k x_0^k\right| \left|\frac{x}{x_0}\right|^k \le M \left|\frac{x}{x_0}\right|^k,$$

and we know for $|x| < |x_0|$ that the geometric sum $\sum_{k=0}^{\infty} M \left|\frac{x}{x_0}\right|^k$ converges. We then deduce from Corollary 9.5.1 that $\sum_{k=0}^{\infty} \left|c_k x^k\right| \in \mathcal{C}$ whenever $|x| < |x_0|$. That is,

$$\sum_{k=0}^{\infty} c_k x_0^k \in \mathcal{C} \Rightarrow \sum_{k=0}^{\infty} c_k x^k \in \mathrm{Abs}\,\mathcal{C} \quad \text{whenever } |x| < |x_0|.$$

Consider now the set $S = \{r \ge 0 : \sum_{k=0}^{\infty} c_k r^k \in \mathcal{C}\}$. Notice that $0 \in S$. If S is unbounded, then $\sum_{k=0}^{\infty} c_k x^k \in \mathrm{Abs}\,\mathcal{C}$ for all real x; that is, the desired result holds with $R = \infty$.

Otherwise, let $R = \sup S \ge 0$. Given $x \in (-R, R)$, there must be at least one element $r \in S$ greater than $|x|$ (otherwise $|x| < R$ would be an upper bound for S). Since $\sum_{k=0}^{\infty} c_k r^k \in \mathcal{C}$, we deduce that $\sum_{k=0}^{\infty} c_k x^k \in \operatorname{Abs}\mathcal{C}$ for $|x| < R$.

Now suppose that $\sum_{k=0}^{\infty} c_k x_0^k \in \mathcal{C}$ for some x_0 with $|x_0| > R \ge 0$. Consider any $r \in (R, |x_0|)$. We know that $\sum_{k=0}^{\infty} c_k r^k \in \operatorname{Abs}\mathcal{C}$ and hence $\sum_{k=0}^{\infty} c_k r^k \in \mathcal{C}$. But then $r \in S$, which contradicts $r > R$. We conclude that $\sum_{k=0}^{\infty} c_k x^k \in \mathcal{D}$ whenever $|x| > R$.

- The Taylor series for $f(x) = e^x$, that is, $\sum_{k=0}^{\infty} x^k/k!$, converges to e^x for all $x \in \mathbb{R}$, by the Limit Ratio Test, since $\lim\limits_{k\to\infty} |x|/(k+1) = 0 < 1$.

- Find the radius of convergence R of
$$\sum_{k=2}^{\infty} \frac{x^k}{\log k}.$$
The ratio of consecutive terms has limit
$$\lim_{k\to\infty} \left| \frac{x^{k+1}}{\log(k+1)} \cdot \frac{\log k}{x^k} \right| = |x| \lim_{k\to\infty} \frac{\log k}{\log(k+1)} = |x| \lim_{u\to\infty} \frac{\frac{1}{u}}{\frac{1}{u+1}} = |x|\,,$$
using L'Hôpital's Rule, where $u \in \mathbb{R}$. The Limit Ratio Test then implies that $R = 1$.

- Consider the Taylor series for the function
$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases}$$
about the point $a = 0$: it is readily shown, using L'Hôpital's Rule, that $f^{(k)}(0) = 0$ for all $k \in \mathbb{N}$. That is, the Taylor series of f converges to zero for all $x \in \mathbb{R}$ (it has an infinite radius of convergence), even though $f(x) \neq 0$ for nonzero x. This example emphasizes that the Taylor series for an infinitely differentiable function f does not necessarily converge to f, even within its radius of convergence!

Remark: To determine the actual interval of convergence, we need to determine R and then test for convergence at $x = a + R$ and $x = a - R$ by other means.

Problem 9.23: Consider the power series $\sum_{k=0}^{\infty} c_k x^k$.

(a) Suppose that $\lim\limits_{k\to\infty} \left| \dfrac{c_{k+1}}{c_k} \right|$ exists. Use the Limit Ratio Test to show that the radius of convergence of the power series is given by
$$R = \frac{1}{\lim\limits_{k\to\infty} \left| \dfrac{c_{k+1}}{c_k} \right|}.$$

If the limit of the ratio of successive terms $\left| \frac{c_{k+1}}{c_k} x \right|$ is less than 1 (i.e. if $|x| < R$) the series $\sum_{k=0}^{\infty} \left| c_k x^k \right|$ converges and if it is bigger than 1 (i.e. if $|x| > R$) the series diverges. Hence R is indeed the radius of convergence.

(b) Suppose that $\lim_{k\to\infty} \sqrt[k]{|c_k|}$ exists. Use the Root Test to show that

$$R = \frac{1}{\lim_{k\to\infty} \sqrt[k]{|c_k|}}$$

is another expression for the radius of convergence.

If $\lim_{k\to\infty} \sqrt[k]{|c_k x^k|}$ is less than 1 (i.e. if $|x| < R$), the series $\sum_{k=0}^{\infty} \left|c_k x^k\right|$ converges, and if it is bigger than 1 (i.e. if $|x| > R$), the series diverges. Hence R is indeed the radius of convergence.

Problem 9.24: Determine the radius of convergence, interval of convergence, and expansion point for the power series

$$\sum_{k=0}^{\infty} \frac{(3x+4)^k}{5^k}.$$

From the ratio test we see that the series converges whenever $\left|\frac{3x+4}{5}\right| < 1$; that is, when $-5 < 3x + 4 < 5$. This corresponds to the interval $(-3, 1/3)$, a radius of convergence of $5/3$, and an expansion point of $-4/3$. Note that the series diverges at both $x = -3$ and $x = 1/3$.

9.D Representation of Functions as Power Series

The closed form sum of a geometric series

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k, \qquad |x| < 1$$

can be used to sum up other power series.

- On substituting $-x^2$ for x, we find

$$\frac{1}{1+x^2} = \sum_{k=0}^{\infty} (-x^2)^k = \sum_{k=0}^{\infty} (-1)^k x^{2k}, \qquad |x| < 1.$$

- On substituting $-x/2$ for x, we find

$$\frac{1}{2+x} = \frac{1}{2}\left(\frac{1}{1+\frac{x}{2}}\right) = \frac{1}{2}\sum_{k=0}^{\infty}\left(-\frac{x}{2}\right)^k = \sum_{k=0}^{\infty} \frac{(-1)^k}{2^{k+1}} x^k, \qquad |x| < 2.$$

- On multiplying $\dfrac{1}{2+x}$ by x^2 we find

$$\frac{x^2}{2+x} = \sum_{k=0}^{\infty} \frac{(-1)^k}{2^{k+1}} x^{k+2}, \qquad |x| < 2.$$

Theorem 9.10 (Derivative and Integral of a Power Series): *The power series*

$$(i) \qquad \sum_{k=0}^{\infty} c_k x^k,$$

$$(ii) \qquad \sum_{k=0}^{\infty} k c_k x^{k-1},$$

$$(iii) \qquad \sum_{k=0}^{\infty} c_k \frac{x^{k+1}}{k+1}$$

all have the same radius of convergence.

Proof: For $k \geq |x|$, note that

$$\left|c_k x^k\right| = |x| \left|c_k x^{k-1}\right| \leq \left|k c_k x^{k-1}\right|.$$

We thus see from the Comparison Test that if (ii) converges absolutely, so does (i). On the other hand, suppose (i) converges absolutely at some $x_0 \neq 0$. The Divergence Test implies that $\left|c_k x_0^k\right| \leq M$ for all k. Thus

$$\left|k c_k x^{k-1}\right| = \left|c_k x_0^k\right| \frac{k}{|x_0|} \left|\frac{x}{x_0}\right|^{k-1} \leq \frac{M}{|x_0|} k \left|\frac{x}{x_0}\right|^{k-1}.$$

We know from the Limit Ratio Test that $\displaystyle\sum_{k=0}^{\infty} k \left|\frac{x}{x_0}\right|^{k-1}$ is convergent for $|x| < |x_0|$. The absolute convergence of (ii) then follows from the Comparison Test. On noting that (i) is just the result of formally differentiating (iii), we see that (i) and (iii) also have the same radius of convergence.

Remark: Suppose that the Taylor series $\displaystyle\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ converges to $f(x)$ for $|x| < R$. Theorem 9.10 tells us that the term-by-term differentiated series

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} k x^{k-1} = \sum_{k=1}^{\infty} \frac{f^{(k)}(0)}{(k-1)!} x^{k-1} = \sum_{k=0}^{\infty} \frac{f^{(k+1)}(0)}{k!} x^k,$$

which we note is just the Taylor series for f', has the same radius of convergence as the Taylor series for f. This means that we may differentiate (or integrate) a power series term-by-term within its radius of convergence: if $\sum_{k=0}^{\infty} c_k x^k$ converges to $f(x)$ for $|x| < R$, then $\sum_{k=1}^{\infty} k c_k x^{k-1}$ converges to $f'(x)$ for $|x| < R$.

- For $|x| < 1$, we may differentiate the geometric series

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$$

term-by-term to find that

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1} = 1 + 2x + 3x^2 + 4x^3 + \ldots, \qquad |x| < 1.$$

- For $|x| < 1$, we may integrate the geometric series

$$\frac{1}{1+x} = \sum_{k=0}^{\infty} (-x)^k$$

term-by-term to find

$$\log(1+x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{k+1}}{k+1} = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots,$$

where we see that the constant of integration vanishes since both sides evaluate to zero when $x = 0$. While both series converge for $|x| < 1$, notice that the Leibniz Alternating Series Test guarantees that the differentiated series also converges at $x = 1$. That is, the interval of convergence of the differentiated series is $(-1, 1]$. On taking the limit as $x \to 1$, we see from the above closed-form expression that the alternating harmonic series converges to $\log 2$.

- For $|x| < 1$, we may integrate the geometric series

$$\frac{1}{1+x^2} = \sum_{k=0}^{\infty} (-x^2)^k$$

term-by-term to find

$$\tan^{-1} x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} + \ldots.$$

Again, the constant of integration is seen to vanish (for the principal branch of the arctangent).

- We can use power series to integrate functions that we cannot integrate by elementary means:

$$\int_0^t e^{-x^2}\,dx = \int_0^t \sum_{k=0}^{\infty} \frac{(-x^2)^k}{k!}\,dx = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)k!} = t - \frac{t^3}{3} + \frac{t^5}{10} + \ldots.$$

Remark: If $\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} b_k x^k$ whenever $|x| < R$, on setting $x = 0$ we see that $a_0 = b_0$, so that $\sum_{k=1}^{\infty} a_k x^k = \sum_{k=1}^{\infty} b_k x^k$. On differentiating each side with respect to x and again setting $x = 0$, we see that $a_1 = b_1$. On repeating this procedure, we deduce that $a_k = b_k$ for $k = 0, 1, 2, \ldots$. That is, the coefficients of a power series are unique, just like the coefficients of a polynomial.

Remark: If $\sum_{k=0}^{\infty} c_k x^k$ converges to $f(x)$ for $|x| < R$, the uniqueness of power series guarantees that $\sum_{k=0}^{\infty} c_k x^k$ is the Taylor series for f; that is $c_k = f^{(k)}(0)/k!$.

Remark: Within their radii of convergence, power series can be added, subtracted, multiplied, divided, differentiated, and integrated just like polynomials.

- For $|x| < 1$ we may expand

$$\begin{aligned}
f(x) &= \frac{e^x}{1+x} \\
&= \left(1 + x + \frac{x^2}{2} + \ldots\right)\left(1 - x + x^2 + \ldots\right) \\
&= \left(1 - x + x^2 + \ldots\right) + x\left(1 - x + x^2 + \ldots\right) + \frac{x^2}{2}\left(1 - x + x^2 + \ldots\right) \\
&= \left(1 - x + x^2 + \ldots\right) + \left(x - x^2 + \ldots\right) + \frac{x^2}{2} + \ldots \\
&= 1 + \frac{x^2}{2} + \ldots.
\end{aligned}$$

From Taylor's theorem, we immediately see that $f(0) = 1$, $f'(0) = 0$, and $f''(0) = 1$.

Problem 9.25: Using long division, show that

$$\tan x = x + \frac{x^3}{3} + \frac{2}{15}x^5 + \ldots.$$

- Another important series is the *Binomial Series*: for $|x| < 1$ and any real number n, the Taylor Series for the function $f(x) = (1+x)^n$ evaluates to (see Problem 4.15)

$$\sum_{k=0}^{\infty} \binom{n}{k} x^k,$$

where

$$\binom{n}{k} = \begin{cases} 1 & \text{if } k = 0, \\ \dfrac{n(n-1)\ldots(n-k+1)}{k!} & \text{if } k \geq 1. \end{cases}$$

The Limit Ratio Test tell us that the series converges when

$$\lim_{k\to\infty} \frac{|n-k|}{k+1}\,|x| = |x| < 1.$$

To see that the series actually converges to f define

$$g(x) = \sum_{k=0}^{\infty} \binom{n}{k} x^k$$

and consider

$$h(x) = (1+x)^{-n} g(x).$$

Note that

$$h'(x) = -n(1+x)^{-n-1} g(x) + (1+x)^{-n} g'(x) = (1+x)^{-n-1}[-ng(x) + (1+x)g'(x)].$$

On using the identity

$$k\binom{n}{k} = k\frac{n(n-1)\ldots(n-k+1)}{k!} = n\frac{(n-1)\ldots(n-k+1)}{(k-1)!} = n\binom{n-1}{k-1},$$

and Pascal's Triangle Law, we find that

$$\begin{aligned}
(1+x)g'(x) &= (1+x)\sum_{k=1}^{\infty} \binom{n}{k} k x^{k-1} \\
&= (1+x)n\sum_{k=1}^{\infty} \binom{n-1}{k-1} x^{k-1} \\
&= n\sum_{k=0}^{\infty} \binom{n-1}{k} x^k + n\sum_{k=1}^{\infty} \binom{n-1}{k-1} x^k \\
&= n + n\sum_{k=1}^{\infty} \left[\binom{n-1}{k} + \binom{n-1}{k-1}\right] x^k \\
&= n + n\sum_{k=1}^{\infty} \binom{n}{k} x^k \\
&= ng(x).
\end{aligned}$$

Thus $h'(x) = 0$ and since $h(0) = g(0) = 1$, we see that $h(x) = 1$ for all $x \in (-1, 1)$. Thus for $|x| < 1$ we find

$$(1+x)^n = \sum_{k=0}^{\infty} \binom{n}{k} x^k = 1 + nx + \frac{n(n-1)}{2}x^2 + \frac{n(n-1)(n-3)}{3!}x^3 + \dots.$$

- For $n = 1/2$ and $|x| < 1$, we find that

$$\begin{aligned}
\sqrt{1+x} &= \sum_{k=0}^{\infty} \binom{1/2}{k} x^k \\
&= 1 + \sum_{k=1}^{\infty} \frac{1/2(1/2-1)(1/2-2)\dots(1/2-k+1)}{k!} x^k \\
&= 1 + \sum_{k=1}^{\infty} \frac{1(1-2)(1-4)\dots(1-2k+2)}{2^k k!} x^k \\
&= 1 + \frac{x}{2} + \sum_{k=2}^{\infty} \frac{(-1)(-3)\dots(3-2k)}{2^k k!} x^k \\
&= 1 + \frac{x}{2} + \sum_{k=2}^{\infty} (-1)^{k-1} \frac{1 \cdot 3 \cdot \dots \cdot (2k-3)}{2^k k!} x^k.
\end{aligned}$$

Appendix A

Complex Numbers

To complete our proof that the method of partial fraction decomposition can be used to integrate any rational function, it will be helpful to first extend our number system beyond the real numbers $\mathbb{R}$.

Recall that z is a root of the polynomial $P(x)$ if $P(z) = 0$.

Q. Do all polynomials have at least one root $z \in \mathbb{R}$?

A. No, consider $P(x) = x^2 + 1$. It has no real roots: $P(x) \geq 1$ for all x.

The *complex numbers* $\mathbb{C}$ are introduced precisely to circumvent this problem. If we replace "$z \in \mathbb{R}$" by "$z \in \mathbb{C}$", we can answer the above question affirmatively.

The complex numbers consist of ordered pairs (x, y) together with the usual component-by-component addition rule (e.g. which one has in a vector space)

$$(x, y) + (u, v) = (x + u, y + v),$$

but with the unusual multiplication rule

$$(x, y) \cdot (u, v) = (xu - yv, xv + yu).$$

Note that this multiplication rule is associative, commutative, and distributive. Since

$$(x, 0) + (u, 0) = (x + u, 0) \quad \text{and} \quad (x, 0) \cdot (u, 0) = (xu, 0),$$

we see that $(x, 0)$ and $(u, 0)$ behave just like the real numbers x and u. In fact, we can map $(x, 0) \in \mathbb{C}$ to $x \in \mathbb{R}$:

$$(x, 0) \equiv x.$$

Hence $\mathbb{R} \subset \mathbb{C}$.

Remark: We see that the complex number $z = (0, 1)$ satisfies the equation $z^2 + 1 = 0$. That is, $(0, 1) \cdot (0, 1) = (-1, 0)$.

Definition: Denote $(0, 1)$ by the letter i. Then any complex number (x, y) can be represented as $(x, 0) + (0, 1)(y, 0) = x + iy$.

Remark: Unlike $\mathbb{R}$, the set $\mathbb{C} = \{(x, y) : x \in \mathbb{R}, y \in \mathbb{R}\}$ is not ordered; there is no notion of positive and negative (greater than or less than) on the complex plane. For example, if i were positive or zero, then $i^2 = -1$ would have to be positive or zero. If i were negative, then $-i$ would be positive, which would imply that $(-i)^2 = i^2 = -1$ is positive. It is thus not possible to divide the complex numbers into three classes of negative, zero, and positive numbers.

Remark: The frequently appearing notation $\sqrt{-1}$ for i is misleading and should be avoided, because the rule $\sqrt{xy} = \sqrt{x}\sqrt{y}$ (which one might anticipate) does not hold for negative x and y, as the following contradiction illustrates:

$$1 = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = i^2 = -1.$$

Furthermore, by definition $\sqrt{x} \geq 0$, but one cannot write $i \geq 0$, since $\mathbb{C}$ is not ordered.

Remark: We may write $(x, 0) = x + i0 = x$ since $i0 = (0, 1) \cdot (0, 0) = (0, 0) = 0$.

Definition: The complex conjugate $\overline{(x, y)}$ of (x, y) is $(x, -y)$. That is,

$$\overline{x + iy} = x - iy.$$

Definition: The *complex modulus* $|z|$ of $z = x + iy$ is given by $\sqrt{x^2 + y^2}$.

Remark: If $z \in \mathbb{R}$ then $|z| = \sqrt{z^2}$ is just the absolute value of z.

We now establish some important properties of the complex conjugate. Let $z = x + iy$ and $w = u + iv$ be elements of $\mathbb{C}$. Then

(i)

$$z\overline{z} = (x, y)(x, -y) = (x^2 + y^2, yx - xy) = (x^2 + y^2, 0) = x^2 + y^2 = |z|^2,$$

(ii)

$$\overline{z + w} = \overline{z} + \overline{w},$$

(iii)

$$\overline{zw} = \overline{z}\,\overline{w}.$$

Problem A.1: Prove properties (ii) and (iii).

Remark: Property (i) provides an easy way to compute *reciprocals* of complex numbers:

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{\bar{z}}{|z|^2}.$$

Remark: Properties (i) and (iii) imply that

$$|zw|^2 = zw\overline{zw} = z\overline{z}w\overline{w} = |z|^2 |w|^2 .$$

Thus $|zw| = |z| |w|$ for all $z, w \in \mathbb{C}$.

Lemma A.1 (Complex Conjugate Roots): *Let P be a polynomial with real coefficients. If z is a root of P, then so is $\bar{z}$.*

Proof: Suppose $P(z) = \sum_{k=0}^{n} a_k z^k = 0$, where each of the coefficients a_k are real. Then

$$P(\overline{z}) = \sum_{k=0}^{n} a_k (\overline{z})^k = \sum_{k=0}^{n} a_k \overline{z^k} = \sum_{k=0}^{n} \overline{a_k}\,\overline{z^k} = \overline{\sum_{k=0}^{n} a_k z^k} = \overline{P(z)} = \overline{0} = 0.$$

Thus, complex roots of real polynomials occur in *conjugate pairs*, z and $\overline{z}$.

Remark: There is a remarkable similarity between the complex multiplication rule

$$(x, y) \cdot (u, v) = (xu - yv, xv + yu)$$

and the trigonometric angle sum formulae. Notice that

$$\begin{aligned}(\cos\theta, \sin\theta) \cdot (\cos\phi, \sin\phi) &= (\cos\theta\cos\phi - \sin\theta\sin\phi, \cos\theta\sin\phi + \sin\theta\cos\phi)\\ &= (\cos(\theta+\phi), \sin(\theta+\phi)).\end{aligned}$$

That is, multiplication of 2 complex numbers on the unit circle $x^2 + y^2 = 1$ corresponds to addition of their angles of inclination to the x axis. In particular, the mapping $f(z) = z^2$ doubles the angle of $z = (x, y)$ and $f(z) = z^n$ multiplies the angle of z by n.

These statements hold even if z lies on a circle of radius $r \neq 1$,

$$(r\cos\theta, r\sin\theta)^n = r^n(\cos n\theta, \sin n\theta);$$

this is known as *deMoivre's Theorem*.

Remark: If we allow θ to vary from 0 to 2π, then when we raise $z = (r\cos\theta, r\sin\theta)$ to the n^{th} power, we move around a circle of radius r^n exactly n times.

We are now ready to prove an important theorem of complex analysis.

Theorem A.1 (Fundamental Theorem of Algebra): *Any non-constant polynomial $P(z)$ with complex coefficients has a root in $\mathbb{C}$.*

Proof: Because $\lim_{|z|\to\infty} |P(z)| = \infty$ and $|P(z)|$ is a continuous function in both x and y, a two-dimensional version of Theorem 3.4 on the set $\{z : |z| \le R\}$, for R sufficiently large, can be used to show that $|P(z)|$ must achieve a minimum value at some point $z_0 = (x_0, y_0) \in \mathbb{C}$.

We now show that $P(z_0) = 0$. Let $n = \deg P(z) \ge 1$. If $z = z_0 + r(\cos\theta, \sin\theta)$, with r sufficiently small, then $P(z)$ will encircle $P(z_0)$ at least once as θ is varied from 0 to 2π. If $P(z_0) \ne 0$, then there exists a $z_1 \in \mathbb{C}$ such that $P(z_1)$ is closer to the origin than $P(z_0)$. This contradicts the fact that $|P(z)|$ has a minimum value at z_0. Hence $P(z_0) = 0$.

Corollary A.1.1 (Polynomial Factorization): Every complex polynomial $P(z)$ of degree $n \ge 0$ has exactly n complex roots z_1, z_2, ..., z_n and can be factorized as $P(z) = A(z - z_1)(z - z_2)\ldots(z - z_n)$, where $A \in \mathbb{C}$.

Proof: Apply Theorem A.1 and Lemma 7.1 recursively n times. (Recall that the degree of the zero polynomial, which has infinitely many roots, is $-\infty$.)

Corollary A.1.2 (Real Polynomial Factorization): Every polynomial with real coefficients can be factorized as

$$P(x) = A(x - a_1)^{n_1}\ldots(x - a_k)^{n_k}(x^2 + \gamma_1 x + \lambda_1)^{m_1}\ldots(x^2 + \gamma_\ell x + \lambda_\ell)^{m_\ell}.$$

Proof: Since the coefficients of $P(x)$ are real, Lemma A.1 $\Rightarrow$ any complex roots must occur in conjugate pairs, say $a + ib$ and $a - ib$, with $b \ne 0$. We can combine these conjugate pairs into irreducible quadratic factors with real coefficients:

$$\begin{aligned}[x - (a + ib)][x - (a - ib)] &= x^2 - (a + ib)x - (a - ib)x + (a + ib)(a - ib)\\ &= x^2 - 2ax + (a^2 + b^2),\end{aligned}$$

with $(2a)^2 - 4(a^2 + b^2) = -4b^2 < 0$. The real roots form the linear factors $(x - a_j)$.

Bibliography

[Fomin *et al.* 1996] D. Fomin, S. Genkin, & I. Itenberg, *Mathematical Circles (Russian Experience)*, volume 7 of *Mathematical World*, American Mathematical Society, Providence, Rhode Island, 1996, translated from Russian by Mark Saul.

[Fridy 1987] J. A. Fridy, *Introductory Analysis: The Theory of Calculus*, Harcourt Brace Jovanovich, San Diego, California, 1987.

[Hille & Salas 1968] E. Hille & S. L. Salas, *First-Year Calculus*, Blaisdell, Waltham, Massachusetts, 1968.

[Kitchen 1968] J. W. Kitchen, Jr., *Calculus of One Variable*, Addison–Wesley, Reading, Massachusetts, 1968.

[Muldowney 1990] J. S. Muldowney, Lecture Notes for Mathematics 117, University of Alberta, 1990.

[Spivak 1994] M. Spivak, *Calculus*, Publish or Perish, Berkeley, 3rd edition, 1994.

Index