Rigorous Calculus

Notes for Math 147 and 148

Kenneth R. Davidson
University of Waterloo

PREFACE

These notes are designed for a rigorous two semester course in calculus based on the completeness of the real numbers. Students are assumed to have had a high school level calculus course already. While we do not assume that students know about proofs, we only provide a rather brief introduction to the notion of proofs in the first chapter before diving in to the business of doing real analysis.

The real numbers are defined as the set of infinite decimals with the identification of a number ending in an infinite string of 9s with another decimal number ending in an infinite string of 0s. This definition has its problems when it comes to defining addition and multiplication. However it is familiar, and is the quickest way to get going. A better method of defining the real numbers is provided in an appendix. Chapter 2 is the meat of the course, and is also the most difficult. We explore what is meant by the completeness of the real numbers in several guises. All of the deeper theorems of calculus, including the Extreme Value Theorem and the Intermediate Value Theorem, rely in an essential way on this material.

Curve sketching is stressed from the beginning. Students can use their knowledge of calculus from their earlier course to help them. To a great extent, the early examples require only a little differentiation. Trigonometric functions, and the logarithm and exponential functions are used throughout, as these provide a wealth of interesting functions. We assume that students are already familiar with trig functions and the basic trig identities such as the addition formula. The natural logarithm is introduced as an area and its properties are derived without the use of integration. The exponential function is the inverse function of the logarithm.

Chapter 4 introduces the key concept of continuity and establishes the Extreme Value Theorem and the Intermediate Value Theorem.

Finally in Chapter 5, we define the derivative formally. We discuss maxima and minima with an emphasis on the Mean Value Theorem. We do not spend much time on the useful topic of max-min problems. We assume that this was well covered in the high school course. We do include some exercises with such problems. We discuss convexity of functions in some detail, relating it to the second derivative and deriving Jensen's inequality.

It is my experience that students have learned L'Hôpital's Rule in high school, of course without the rather tricky proof. I make a big deal about this being unacceptable until it is proven. In particular, if the famous limit $\lim_{x\to 0} \frac{\sin x}{x} = 1$ is ever 'proven' using L'Hôpital's Rule, there will be a severe penalty. This limit is needed to determine the derivative of the sin function, and so using L'Hôpital's Rule is circular. Moreover, I make the point soon after that many limits that can be computed by two or three applications of L'Hôpital's Rule can be established quickly using low order Taylor polynomial approximations.

At the University of Waterloo, all of integration is left for the second semester. To accomodate this, we cover some of the material in Chapter 10 in the first semester. Specifically we discuss Taylor polynomials and a few examples of when the infinite series converges. The general discussion of series, both of numbers and functions, is done in the second semester.

In the second semester, we begin with the theory of integration culminating in the Fundamental Theorem of Calculus which ties the integral (a computation of area) with the derivative. Chapter 7 deals with a variety of computational methods and tricks for calculating integrals. Then in Chapter 8, there is a variety of topics on integration: improper integrals, volumes and arc length, polar coordinates and parameterizations.

Chapter 9 deals with infinite series. In a certain sense, a series is just another way of describing a sequence. However it is a common method, and there are a number of new methods for handling them. Then we turn to sequences and series of functions. As mentioned above, we cover Taylor polynomials in the first semester. We make a systematic study of uniform convergence. Then we look at power series, which have some especially nice properties. We conclude the chapter with a proof of Abel's theorem about convergence at the radius of convergence.

The last chapter is an introduction to differential equations. We only look at first order DEs and second order linear DEs. This will give students a sense of the ideas involved. It is really only a taste.

Finally, the appendix contains a number of interesting enrichment topics.

Kenneth R. Davidson
August, 2021

CONTENTS

CHAPTER 1

The Logic of Proofs

1.1. The Language of Mathematics

In this section, we provide a quick overview of some of the language behind mathematical and logical statements. This is not a thorough treatment, but should suffice to get us going.

Mathematics deals with *statements*, which assert some relationship between certain items or classes of items. Examples are

- x belongs to the set X.
- Y is a subset of X.
- $x = y$ or $x < y$ or $x > y$.
- $x < y$.
- $f : [0, 1] \to \mathbb{R}$ is a monotone increasing function.
- All horses are white.
- Every cat has one tail.

Examples of non-statements are

- $x + y$.
- Don't divide by 0.

Statements must have the property that they are either *True* or *False*. Often a context is specified that limits the variables (if any) in the statements. Certain *self-referential statements* are not allowed, such as

- This statement is false.

Statements can be manipulated or combined with others to make new statements. There are three key words that have precise meanings that may differ from popular usage in everyday speech. These are *not*, *and* and *or*.

If A is a statement, then "not A" or $\neg A$ is the *negation* of A. If A is true, then $\neg A$ is false and vice versa.

If A and B are statements, then "A *and* B" or $A \wedge B$ is true if and only if both A and B are true. If either is false, then $A \wedge B$ is false.

If A and B are statements, then "A *or* B" or $A \vee B$ is true if and only if at least one of A or B is true. If both are false, then $A \vee B$ is false.

We also define *implication*, "A implies B" or $A \Rightarrow B$, means that if A is true, then B is true. The statement $A \Rightarrow B$ is true if A and B are both true, but also if A is false regardless of the truth of B. This is because such a statement means "*If A is true, then B is also true.*" For this reason, you also see this written "*if A, then* B".

Since a statement can be either true or false, one can make *truth tables* which start with the various possibilities for the original statements and calculates the truth or falsity of other statements obtained by combining them. Here is an example. To save space, let MP be the statement $(A \wedge (A \Rightarrow B)) \Rightarrow B$.

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$B \Rightarrow A$	$\neg B \Rightarrow \neg A$	$A \wedge (A \Rightarrow B)$	MP
T	T	F	T	T	T	T	T	T	T
T	F	F	F	T	F	T	F	F	T
F	T	T	F	T	T	F	T	F	T
F	F	T	F	F	T	T	T	F	T

We don't use truth tables to prove theorems, but they can be used to investigate general relationships. The *converse* of the statement $A \Rightarrow B$ is the statement $B \Rightarrow A$. You can see from the truth table that these statements are not equivalent. If $A \Rightarrow B$ is true, the converse may or may not be true depending on the circumstances.

The *contrapositive* of the statement $A \Rightarrow B$ is the statement $\neg B \Rightarrow \neg A$. The truth table shows that these statements are equivalent. You can think this through as follows: Assume that $A \Rightarrow B$. This means that if A is true, then B is true. So if B is false, then A cannot be true, so A is also false. That is the statement $\neg B \Rightarrow \neg A$.

We use the expression "A if and only if B" and write $A \Leftrightarrow B$ to mean that $(A \Rightarrow B) \wedge (B \Rightarrow A)$. It means that either both A and B are true or both are false.

A statement is a *tautology* if it is always true. The statement MP is a tautology. This particular statement is known as *modus ponens*.

A *formula* is a statement usually involving some variables. The truth or falsity of the statement may depend on the values assigned to the variables.

For example, consider the statement $P(x, y)$ that $2x^2 - 2xy + y^2 + 2x - 4y \geqslant 20$. Here we are told that x and y are real numbers. You can check that $P(4, 2)$ is true, but $P(3, 3)$ is false.

1.1.1. Quantifiers. There are two important *quantifiers* that we use in mathematics. The first is the universal quantifier "*for all*" or $\forall$. This has the form $\forall x \in X,\ P(x)$. It states that for every variable x in some specified range X, the statement $P(x)$ is true. If the statement fails for a single value of a variable, then the whole statement is false. Consider

(A) $\forall n \in \mathbb{N}$, $n^2 - 2$ is even.

(B) $\forall n \in \mathbb{N}$, $n^2 + n + 41$ is prime.

In both statements, the range is specified to be positive integers. Statement (A) is true because $n^2 - n = n(n-1)$, and either n is even, and thus so is the product or

n is odd, in which case $n - 1$ is even, and so is the product. Thus $n^2 - n$ is even for all choices of n.

For (B), we look at the value of $n^2 + n + 41$ for $n \geqslant 1$. It starts out well: $43, 47, 53, 61, 71, 83, 97, 113, 131, 151$. So far, all are prime. However when you get to $n = 40$, we get $(40)^2 + 40 + 41 = (41)^2$ which is not prime. So the statement is false. The number $n = 40$ is called a *counterexample* to statement (B).

The second is the existential quantifier "*there exists*" or $\exists$. A statement has the form "$\exists x \in X$ such that $P(x)$". It is true if there is a single x for which $P(x)$ is true. Consider

(C) $\exists x \in \mathbb{R}$ such that $x \sin(\frac{1}{x}) = 1$.

(D) $\exists n \in \mathbb{N}$ such that $1141n^2 + 1$ is a perfect square.

The function $f(x) = x \sin(\frac{1}{x})$ is even, i.e., $f(-x) = f(x)$, so we can consider $x \geqslant 0$. When $0 \leqslant x < 1$, $|x \sin(\frac{1}{x})| \leqslant |x| < 1$. If $x \geqslant 1$, then $\frac{1}{x} \leqslant 1 \in (0, \frac{\pi}{2})$. In this range, we will show in the course that $0 < \sin \frac{1}{x} < \frac{1}{x}$, so $0 < f(x) < 1$. Therefore there is not a single value of x which makes the statement true—so (A) is false.

To properly analyze (D), we would need to know some more number theory. You could check on a computer, say the first trillion numbers, and you would not succeed. It turns out that there are infinitely many integers which make this a perfect square, but the smallest has 26 digits:

$$30,693,385,322,765,657,197,397,208$$

Therefore (D) is true, even though you would never find it by crude methods.

When you negate a statement using $\forall$, it becomes an $\exists$. The reason is that a "for all" statement is contradicted by a single counterexample. For instance $\neg(B)$ is the statement: $\exists n \in \mathbb{N}$ such that $n^2 + n + 41$ is not prime. This is a true statement, since $n = 40$ does the job.

Similarly, when you negate $\exists$, it becomes a $\forall$ statement. This is again because to contradict $\exists x \in X$ such that $P(x)$, you must show that $P(x)$ is false for all $x \in X$. So the negation is $\forall x \in X, \neg P(x)$. For instance, $\neg(C)$ is the statement $\forall x \in \mathbb{R}$, $x \sin(\frac{1}{x}) \neq 1$. We showed that this is a true statement by proving something a bit stronger, that $\left|x \sin(\frac{1}{x})\right| < 1$ for all $x \in \mathbb{R}$.

Things start to get more complicated when we have more quantifiers. This comes up in calculus because the definitions of limit and continuity require multiple quantifiers. Here we just give a couple of elementary examples. Consider

(E) $\forall n \in \mathbb{N}_0\ \exists m \in \mathbb{N}_0$ such that 13 divides $m^2 + n^2$.

(F) $\exists m \in \mathbb{N}_0$ such that $\forall n \in \mathbb{N}_0$, 13 divides $m^2 + n^2$.

For statement (E), we are asked if for each $n \in \mathbb{N}_0$, we can select some $m \in \mathbb{N}_0$ so that $m^2 + n^2$ is a multiple of 13. We are allowed to choose m any way we wish. So let's choose $m = 5n$. Then $m^2 + n^2 = 25n^2 + n^2 = 13(2n^2)$ is divisible by 13. Therefore (E) is true.

Look at the difference when we reverse the order of the quantifiers. This is asking for a single m so that m^2+n^2 is always a multiple of 13. Thus both m^2+0^2 and m^2+1^2 would need to be divisible by 13. But then $(m^2+1)-(m^2+0)=1$ would be divisible by 13. This is absurd. So the statement (F) is false.

1.2. Proofs

1.2.1. Direct proofs. Start with the hypothesis and work through straight to the answer. Here is an example.

1.2.1. DEFINITION. If $x = a_0.x_1x_2x_3\dots$ is an infinite decimal (here a_0 is an integer and $x_i \in \{0,1,\dots,9\}$), we say that the expansion is *eventually periodic* if there are positive integers N and d so that $x_{n+d} = x_n$ for all $n \geqslant N$.

1.2.2. THEOREM. *If $x \in \mathbb{R}$ has a decimal expansion which is eventually periodic, then x is rational.*

PROOF. Multiply x by 10^N and by 10^{N+d}. There are integers b and c so that

$$10^{N+d}x = c.x_{N+d+1}x_{N+d+2}x_{N+d+3}\cdots$$
$$10^{N}x = b.x_{N+1}x_{N+2}x_{N+3}\cdots$$

Hence by subtracting,

$$(10^{N+d}-10^N)x = c-b.$$

Therefore $x = \dfrac{c-b}{10^{N+d}-10^N}$ is rational. ∎

1.2.2. Pigeonhole Principle. If $n+1$ or more objects are divided into n categories, then there are at least two objects in the same category. The name refers to an office mailroom in which each person gets a *pigeonhole* in which to receive letters. There are variants, such as putting $mn+1$ objects into n categories. You can deduce that some category contains at least $m+1$ objects. Usually two is enough.

1.2.3. THEOREM. *If $x \in \mathbb{Q}$, then the decimal expansion of x is eventually periodic.*

PROOF. Since x is rational, we can write $x = \dfrac{p}{q}$ where p, q are integers and $q > 0$. For each $k \geqslant 0$, there is a unique integer $r_k \in \{0,1,\dots,q-1\}$ so that q divides $10^k - r_k$. That is, r_k is the remainder left when dividing q into 10^k. Then $\{r_0, r_1, \dots, r_q\}$ are $q+1$ remainders taking only q possible values. By the pigeonhole principle, there are two values $0 \leqslant i < j \leqslant q$ so that $r_i = r_j$. Since q divides $10^i - r_i$ and $10^j - r_j$, it also divides the difference $10^j - 10^i$, say $10^j - 10^i = qa$. Then

$$x = \frac{p}{q} = \frac{ap}{aq} = 10^{-i}\frac{ap}{10^{j-i} - 1}.$$

Let $d = j - i$ and divide the denominator $10^d - 1$ into ap to get an integer b with a remainder s with $0 \leqslant s < 10^d - 1$. So s has a most d digits, and thus we can write $s = s_1 s_2 \dots s_d$ with $s_i \in \{0, 1, \dots, 9\}$ and we put 0's at the beginning if required to make this a d digit number. Then

$$y = 0.s_1 s_2 \dots s_d s_1 s_2 \dots s_d \dots = \sum_{k=1}^{\infty} 10^{-dk} s = \frac{10^{-d} s}{1 - 10^{-d}} = \frac{s}{10^d - 1}$$

because a periodic decimal is a geometric series. Hence

$$10^i x = b.s_1 s_2 \dots s_d s_1 s_2 \dots s_d \dots = b.\overline{s_1 s_2 \dots s_d}$$

and the decimal expansion of x is just the same with the decimal place shifted i places to the left. So it is eventually periodic. ■

1.2.3. Proof by Contradiction. We saw that the contrapositive of a statement $A \Rightarrow B$ is $\neg B \Rightarrow \neg A$, and has the same truth. So we assume that B is false, and deduce that A is false. This proves the contrapositive, so we are done. We usually think of this as reaching a contradiction to the hypothesis that A is true, which explains the name.

1.2.4. THEOREM. *If $d \in \mathbb{N}$ is a positive integer which is not a perfect square, then $\sqrt{d}$ is irrational.*

PROOF. Assume that $\sqrt{d}$ is rational. Then $A = \{n \in \mathbb{N} : n\sqrt{d} \in \mathbb{N}\}$ is not empty. A non-empty subset of $\mathbb{N}$ has a smallest element, so let a be the smallest element of A. So $a\sqrt{d} \in \mathbb{N}$ but $b\sqrt{d}$ is not an integer for $1 \leqslant b < a$.

Choose the integer $m \in \mathbb{N}$ so that $m^2 < d < (m+1)^2$. Notice that

$$0 < \sqrt{d} - m < \sqrt{(m+1)^2} - m = 1.$$

Therefore $0 < b := a(\sqrt{d} - m) < a$. However $b = a\sqrt{d} - am$ is an integer, so $1 \leqslant b \leqslant a - 1$. Finally $b\sqrt{d} = ad - (a\sqrt{d})m$ is an integer. This means that $b \in A$ and $b < a$. This contradicts the assumption that a was the smallest element, which in turn was a consequence of the fact that A was non-empty. Hence A is empty, and so $\sqrt{d}$ is irrational. ■

1.2.4. Proof by Induction. The *Principle of Induction*: let $P(n)$ be a sequence of statements for $n \geqslant n_0$. Suppose that

(1) $P(n_0)$ is true, and

(2) If $n > n_0$ and $P(k)$ is true for $n_0 \leqslant k < n$, then $P(n)$ is true.

Then $P(n)$ is true for each $n \geqslant n_0$.

You can think of the statements $P(n)$ as dominoes which are lined up in such a way that if the earlier dominoes are knocked over, then one of them will knock over the nth domino. Then you knock over the first one and watch them all fall in succession.

Simple induction works by showing that statement $P(n-1)$ implies $P(n)$. But there is no reason that it cannot be some other earlier statement or more than one earlier statement which are required to establish $P(n)$. We will see two examples. The first is a bit trickier than simple induction in that it depends on two previous statements.

1.2.5. THEOREM. *The Fibonacci sequence is defined recursively by*

$$F(0) = F(1) = 1 \quad \textit{and} \quad F(n+2) = F(n) + F(n+1) \textit{ for } n \geqslant 0.$$

Let $\tau = \dfrac{\sqrt{5}+1}{2}$. *Then* $F(n) = \dfrac{\tau^{n+1} - (-1/\tau)^{n+1}}{\sqrt{5}}$.

PROOF. First observe that

$$\frac{1}{\tau} = \frac{2}{\sqrt{5}+1}\,\frac{\sqrt{5}-1}{\sqrt{5}-1} = \frac{2(\sqrt{5}-1)}{4} = \frac{\sqrt{5}-1}{2}.$$

Let $P(n)$ be the statement $F(n) = \dfrac{\tau^{n+1} - (-1/\tau)^{n+1}}{\sqrt{5}}$ for $n \geqslant 0$. When $n = 0$,

$$\frac{\tau^1 - (-1/\tau)^1}{\sqrt{5}} = \frac{1}{\sqrt{5}}\Big(\frac{\sqrt{5}+1}{2} + \frac{\sqrt{5}-1}{2}\Big) = \frac{\sqrt{5}}{\sqrt{5}} = 1.$$

Hence $P(0)$ is true. Next consider $n = 1$.

$$\frac{\tau^2 - (-1/\tau)^2}{\sqrt{5}} = \frac{1}{\sqrt{5}}\Big(\frac{(\sqrt{5}+1)^2}{4} - \frac{(\sqrt{5}-1)^2}{4}\Big) = \frac{4\sqrt{5}}{4\sqrt{5}} = 1.$$

Thus $P(1)$ is true. This is step one.

We need the identities

$$1 + \tau = 1 + \frac{\sqrt{5}+1}{2} = \frac{\sqrt{5}+3}{2} = \tau^2$$

and

$$1 - 1/\tau = 1 - \frac{\sqrt{5}-1}{2} = \frac{3-\sqrt{5}}{2} = \frac{1}{\tau^2}.$$

Now for the induction step. Suppose $n \geqslant 2$ and that $P(n-2)$ and $P(n-1)$ are true. Then

$$\begin{aligned} F(n) &= F(n-2) + F(n-1) \\ &= \frac{\tau^{n-1} - (-1/\tau)^{n-1}}{\sqrt{5}} + \frac{\tau^{n} - (-1/\tau)^{n}}{\sqrt{5}} \\ &= \frac{1}{\sqrt{5}}\Big(\tau^{n-1}(1+\tau) - (-1/\tau)^{n-1}(1-1/\tau)\Big) \\ &= \frac{1}{\sqrt{5}}\Big(\tau^{n-1}\tau^{2} - (-1/\tau)^{n-1}(1/\tau)^{2}\Big) \\ &= \frac{\tau^{n+1} - (-1/\tau)^{n+1}}{\sqrt{5}} \end{aligned}$$

This shows that $P(n)$ follows from the previous two statements, $P(n-2)$ and $P(n-1)$. By induction, the statements $P(n)$ are true for all $n \geqslant 0$. ■

Here is a second example.

1.2.6. THEOREM. *Every natural number $n \geqslant 2$ is a product of prime numbers.*

PROOF. Let $P(n)$ be the statement that n factors as a product of primes. Start at $n = 2$. Then $P(2)$ holds because 2 is prime (so is the product of one prime). For the induction step, we need to assume $P(k)$ for all $2 \leqslant k < n$. Consider $P(n)$. There are two cases.

Case 1: n is prime. Then $P(n)$ is true,

Case 2: n is not prime, so $n = ab$ where $2 \leqslant a, b < n$. By $P(a)$ and $P(b)$, we can write $a = p_1 \dots p_m$ as a product of primes and $b = q_1 \dots q_n$ as another product of primes. Therefore $n = ab = p_1 \dots p_m q_1 \dots q_n$ is a product of primes, so $P(n)$ is true. By induction, every $n \geqslant 2$ is a product of primes. ■

In principle, this proof is more complex than the first. For example, for $P(48)$, we might factor $48 = 6 \cdot 8$. Then the statement $P(48)$ is deduced from $P(6)$ and $P(8)$. In each case, we don't need to know the explicit factorization, only that a and b are strictly smaller so that we are assured that $P(a)$ and $P(b)$ have already been verified.

Exercises for Chapter 1

1. In Xanadu, people are either Knights, Normals or Villains. Knights always tell the truth, Villains always lie, and Normals can do both. Knights outrank Normals, who outrank Villains.

(a) Three people are known to consist of exactly one Knight, one Normal and one Villain. They say:
Alice: I'm normal.
Bob: That is true.
Charlie: I'm not normal.
Determine which type each person is. Explain.

(b) Two people from Xanadu, Dick and Jane, say:
Dick: I rank below Jane.
Jane: That is not true.
Determine the types of both and decide who is telling the truth. Explain.

2. Three young men accused of stealing cellphones make the following statements:
(1) Ed: "Fred did it, and Ted is innocent."
(2) Fred: "If Ed is guilty, then so is Ted."
(3) Ted: "I'm innocent, but at least one of the others is guilty."

(a) If they are all innocent, who is lying? Explain.
(b) If all these statements are true, who is guilty? Explain.
(c) If the innocent told the truth and the guilty lied, who is guilty? Explain.

3. (a) Show that $|15\sin\theta + 8\cos\theta| \leqslant 17$.
Use trig identities, not calculus, to do this exercise.
HINT: show that there is an angle α with $\sin\alpha = \frac{8}{17}$ and $\cos\alpha = \frac{15}{17}$.
(b) When does equality hold in this inequality?

4. Let a, b, c be positive real numbers greater than 1. Show that

$$\log_a(bc)\ \log_b(ac)\ \log_c(ab) = \log_a(bc) + \log_b(ac) + \log_c(ab) + 2.$$

HINT: : express everything in terms of $A = \log a$, $B = \log b$ and $C = \log c$.

5. (a) Find an expression for

$$f(x) = \big|2x - |2x+1|\big| - \big|x - |x-2|\big| \qquad \text{for} \quad x \in \mathbb{R}$$

which avoids the use of absolute value signs or square roots. You may split the real line into disjoint intervals and have a different algebraic expression on each one.
(b) Graph $f(x)$. In particular, indicate all solutions of $f(x) = 0$.

6. (a) Show that if x is a positive real number, then

$$\frac{-1}{8x^2} < x\left(\sqrt{x^2+1}-x\right) - \frac{1}{2} < \frac{-1}{8(x^2+1)}.$$

HINT: Use $a - b = \frac{a^2-b^2}{a+b}$ to 'rationalize the numerator' twice! *This is superior to working backwards from the answer and bashing it out using algebra.*

(b) Use (a) to show that for $x > 0$,

$$x + \frac{1}{2x} - \frac{1}{8x^3} < \sqrt{x^2+1} < x + \frac{1}{2x} - \frac{1}{8x^3} + \frac{1}{8x^4}.$$

7. The *Fibonacci sequence* is defined by $F(0) = F(1) = 1$ and $F(n+2) = F(n+1) + F(n)$ for all $n \geqslant 0$. Fix an integer $D \geqslant 2$. Consider the remainders $q(n)$ obtained by dividing $F(n)$ by D, with $0 \leqslant q(n) < D$. Prove that this sequence is periodic with some period $d \leqslant D^2$ as follows:

(a) Show there are integers $0 \leqslant i < j \leqslant D^2$ such that $q(i) = q(j)$ and $q(i+1) = q(j+1)$.
HINT: pigeonhole.

(b) Let $d = j - i$. Use induction to show that $q(n+d) = q(n)$ and $q(n+1+d) = q(n+1)$ for all $n \geqslant i$.

(c) Show that $q(n+d) = q(n)$ for all $n \geqslant 0$. HINT: work backwards from $n = i$.

CHAPTER 2

The Real Numbers and Limits

2.1. The real numbers

What are the real numbers? You may think that you know the answer, but it turns out that you need to be very careful about this. It took mathematicians a long time to realize that this was even necessary. There are some sophisticated ways to accomplish this, but we will get by with a more mundane approach. We will have to slough over some fine points in order to get on with doing calculus.

We will define a real number to be an infinite decimal. For convenience, we will (temporarily) write our real numbers in the form

$$x = a_0.a_1a_2a_3\dots \quad \text{where } a_0 \in \mathbb{Z} \text{ and } a_i \in \{0, 1, \dots, 9\} \text{ for } i \geqslant 1.$$

This is a bit peculiar in the sense that we think of x as a (possibly negative) integer a_0 *plus* a positive infinite decimal number $0.a_1a_2a_3\dots$. This will be convenient for us, in that it treats all intervals $[n, n+1]$ the same. Once we start doing our real business, we make the fairly trivial change back to the common usage where negative real numbers are written as the negative of a positive real number.

We immediately run into difficulty with this definition. Are $1.000\dots$ and $0.999\dots$ different real numbers? They are definitely different infinite decimals. However, when we try to distinguish them, we find that they are infinitely close to one another. Indeed, if we have any infinite decimal that ends in an infinite string of 9's, we use the formula for summing a geometric series to get

$$\begin{aligned} x &= a_0.a_1\dots a_n999\dots \\ &= a_0.a_1\dots a_n + 10^{-n}\sum_{k\geqslant 1}\tfrac{9}{10^k} \\ &= a_0.a_1\dots a_n + 10^{-n}. \end{aligned}$$

The rational number on the right hand side has a finite decimal expansion. If I start at the point where $a_n \neq 9$, this is $y = a_0.a_1\dots a_{n-1}(a_n+1)000\dots$ So it make sense to identify x and y as the same real number. We can think of the infinite decimal as a name for the real number, and certain numbers, those ending in an infinite string of 0's or 9's, have two names. We call this an *equivalence relation* where certain names are identified and considered as a single object. (See Appendix A.1.)

The set $\mathbb{R}$ of real numbers is an *ordered field*. It has a *total order*: a relation $<$ on $\mathbb{R}$ such that

(1) For $x, y \in \mathbb{R}$, exactly one of $x < y$, $x = y$ or $y < x$ holds.

(2) For $x, y, z \in \mathbb{R}$, if $x < y$ and $y < z$, then $x < z$.

And $\mathbb{R}$ is a field: there are operations of addition $(x + y)$ and multiplication (xy) and special elements 0, 1 such that for $x, y, z \in \mathbb{R}$,

(3) $x + y = y + x$ (addition is commutative).

(4) $x + (y + z) = (z + y) + z$ (addition is associative).

(5) $x + 0 = x$ (0 is the additive identity).

(6) For $x \in \mathbb{R}$, $\exists y \in \mathbb{R}$ called "$-x$" so that $x + y = 0$ (additive inverse).

(7) $xy = yx$ (multiplication is commutative).

(8) $x(yz) = (xy)z$ (multiplication is associative).

(9) $x1 = x$ (1 is the multiplicative identity).

(10) If $x \neq 0$, $\exists y \in \mathbb{R}$ called "x^{-1}" so that $xy = 1$ (multiplicative inverse).

(11) $(x + y)z = xz + yz$ (distributive law).

Finally there are some axioms that relate the order with the algebraic relations.

(12) If $x < y$, then $x + z < y + z$.

(13) If $0 < x$ and $0 < y$, then $0 < xy$.

It is a lot of work to check all of these properties, and we are not going to do it. We will just discuss a few issues that arise.

Firstly, the order is easy to describe. If $x \neq y$, then choose a decimal expansion for each:

$$x = a_0.a_1a_2a_3\ldots \quad \text{and} \quad y = b_0.b_1b_2b_3\ldots$$

Since they differ, there is a first $n \geqslant 0$ such that $a_i = b_i$ for $0 \leqslant i < n$ and $a_n \neq b_n$. If $a_n < b_n$, we say that $x < y$, and if $a_n > b_n$, we say $y < x$. The rational numbers, and even the numbers with a finite decimal expansion (you can check that these are rational numbers of the form $x = \frac{p}{2^m 5^n}$) are *order dense* in $\mathbb{R}$, i.e., if $x < y$, there is a finite decimal number z such that $x < z < y$. Indeed, $z = b_0.b_1b_2b_3\ldots b_n000\ldots$ works unless $y = z$. If $a_n \leqslant b_n - 2$, $z = b_0.b_1b_2b_3\ldots b_{n-1}(b_n - 1)000\ldots$ works. Lastly, if $a_n = b_n - 1$ and $x = a_0.a_1a_2a_3\ldots a_n999999999a_m\ldots$ with $a_m < 9$, then

$$z = a_0.a_1a_2a_3\ldots a_n999999999(a_m + 1)000\ldots$$

works. Note that in this case, x cannot end in infinitely many 9's because then

$$x = a_0.a_1a_2a_3\ldots a_n999\cdots = a_0.a_1a_2a_3\ldots(a_n + 1)000\cdots = y.$$

Adding and multiplying infinite decimals is delicate. Considering computing $\pi + e$.

$$\begin{aligned} \pi &= 3.141592653589\ldots \\ e &= \underline{2.718281828459\ldots} \\ \pi + e &= 5.85987448204? \end{aligned}$$

We have to add from the left, but it is necessary to look ahead for carries from the right. The red digits are the result of carries. The blue ? indicates that we need more information to decide if the next digit is 8 or 9. You may have to look a long way to be sure of the next digit. However we can say for sure that

$$5.859874482048 < \pi + e < 5.859874482050.$$

So we know the answer is $5.859874482049 \pm 10^{-12}$. In calculus, this information is just as useful as knowing the digit for sure.

Multiplication is even more challenging to define than addition. Again we can bound the product to any desired accuracy by using finite decimal approximations. In Appendix A.2, we will explain a better way to approach this problem.

We will need to use the absolute value frequently. This is defined as

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

An easy but important fact is the *triangle inequality*

$$|x + y| \leqslant |x| + |y|.$$

This is an equality if x and y have the same sign, but is strict when $xy < 0$. It can be rearranged to provide the inequalies

$$|x| \geqslant |x + y| - |y| \quad \text{and} \quad |x \pm y| \geqslant \big||x| - |y|\big|.$$

A frequent use of the absolute value is in describing an interval by $\{x : |x-a| < r\}$. This means that $-r < x - a < r$, which can be rewritten as $a - r < x < a + r$.

2.2. Limits

What does it mean to say that a sequence a_n converges to L? Here are some attempts:

(A) The larger n gets, the closer a_n gets to L.

(B) The larger n gets, the closer a_n gets to L; and it gets arbitrarily close.

(C) Eventually $a_n = L$.

(D) Eventually a_n is close but not equal to L.

(E) Eventually every a_n is as close as we want to L.

One problem with (A) is that it doesn't say how close. So a sequence like $\frac{1}{2}, 1, \frac{2}{3}, 1, \frac{3}{4}, 1, \frac{4}{5}, \dots$ gets closer and closer to π, but also closer and closer to e. Statement (B) tries to fix that by specifying that it has to get very, very close. However it seems to suggest a monotone approach. What about $\frac{1}{2}, \frac{3}{2}, \frac{2}{3}, \frac{5}{4}, \frac{3}{4}, \frac{9}{8} \cdot \frac{4}{5}, \frac{17}{16}, \dots$ This sequence approaches 1 from both sides, and the even terms are getting there a lot faster than the odd terms. This sequence should be considered as convergent—so the sequence can get close, back off a little bit, and get even closer. A sequence can approach the limit from both sides, and doesn't have to get closer at each step. Statement (C) is way too strong. And (D) excludes a sequence like $1, 1, 1, 1, \dots$ or $\frac{1}{2}, 1, \frac{2}{3}, 1, \frac{3}{4}, 1, \dots$.

What about the sequence $\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{4}{5}, \dots$? Does this converge to both 0 and 1, or to neither? What about $0, 1, 0, 1, 0, 1, \dots$? We say that these sequences do not converge because they do not get close to any number L. The odd terms are far from 1, and the even terms are far from 0. All are eventually bounded away from anything else.

Statement (E) seems to paraphrase what we want. The trouble with it is just that the meaning is not precisely articulated. In mathematics, we need to be able to nail it down is quantitative terms. By "*arbitrarily close*", we use a small (but unspecified) positive number $\varepsilon > 0$. By '*eventually*', we mean that there should be some number N so that we are close within ε for *all* $n \geqslant N$. Putting this altogether we get the following formulation.

2.2.1. DEFINITION. $\lim_{n\to\infty} a_n = L$ means: for any $\varepsilon > 0$, there is an N so that for all $n \geqslant N$, we have $|a_n - L| < \varepsilon$. In symbols: $\forall_{\varepsilon>0}\exists_{N\in\mathbb{N}}\forall_{n\geqslant N}|a_n - L| < \varepsilon$. A sequence which has a limit is said to *converge*.

We will refer to this as the ε–N definition of limit. To verify this definition in examples, you should think of ε being given to you, and your job is to find the N which makes the definition work.

2.2.2. EXAMPLE. Let $a_n = \begin{cases} \frac{n}{n+1} & \text{if } n \text{ is odd.} \\ 1 + 2^{-n} & \text{if } n \text{ is even.} \end{cases}$ Let's show that $L = 1$ is the limit. We are given $\varepsilon > 0$, and we can find some N large enough so that $\frac{1}{N} < \varepsilon$. If $n \geqslant N$ is odd, then $|a_n - 1| = \frac{1}{n} \leqslant \frac{1}{N} < \varepsilon$; while if $n \geqslant N$ is even, then $|a_n - 1| = 2^{-n} \leqslant \frac{1}{N} < \varepsilon$. So this choice of N does the job. There is no advantage to choosing N in an optimal way. Thus $\lim_{n\to\infty} a_n = 1$.

2.2.3. EXAMPLE. Let $x_n = \dfrac{2n^3 + n^2 - 137}{5n^3 - n - 1}$ for $n \geqslant 1$. We rewrite this as

$$x_n = \frac{2 + \frac{1}{n} - \frac{137}{n^3}}{5 - \frac{1}{n^2} - \frac{1}{n^3}}.$$

This makes it clear that the numerator tends to 2 while the denominator tends to 5. We need a quantitative estimate for the difference that is tractable and goes to 0.

$$\begin{aligned} |x_n - \tfrac{2}{5}| &= \Big|\frac{2n^3 + n^2 - 137}{5n^3 - n - 1} - \frac{2}{5}\Big| \\ &= \frac{|(10n^3 + 5n^2 - 685) - (10n^3 - 2n - 2)|}{(5n^3 - n - 1)5} \\ &= \frac{|7n^2 - 687|}{25n^3 - 5n - 5} < \frac{7n^2}{24n^3} < \frac{1}{3n}. \end{aligned}$$

The second last inequality is valid provided that $n \geqslant 10$. Now given $\varepsilon > 0$, choose an $N \geqslant \max\{10, \frac{1}{3\varepsilon}\}$. Then $N \geqslant 10$, so that the inequality is valid, and also $\frac{1}{3N} < \varepsilon$. Now if $n \geqslant N$, we have $|x_n - \frac{2}{5}| < \frac{1}{3n} < \varepsilon$. This verifies the definition of limit. Thus $\lim_{n\to\infty} x_n = \frac{2}{5}$.

2.2.4. EXAMPLE. Let $a_n = (-1)^n$ for $n \geqslant 1$; i.e., $-1, 1, -1, 1, -1, 1, \dots$. This sequence does not appear to converge. To establish this, we first need to understand the negation of the limit definition. In order for the definition to fail *for a specific value of* L, we only need to find a single $\varepsilon > 0$ for which it fails, but we then need to show that no choice of N will work. To that end, given any N, we need to find some $n \geqslant N$ so that $|a_n - L| \geqslant \varepsilon$. But we also need to consider all values of L.

So take an arbitrary $L \in \mathbb{R}$. Consider two cases.

Case 1 $L \geqslant 0$. Take $\varepsilon = 1$. Given any N, choose an *odd* $n > N$. Then

$$|a_n - L| = L + 1 \geqslant \varepsilon.$$

So no N works, and thus L is not the limit.

Case 2 $L < 0$. Take $\varepsilon = 1$. Given any N, choose an *even* $n > N$. Then

$$|a_n - L| = |L| + 1 > \varepsilon.$$

So no N works, and thus L is not the limit. Therefore this sequence has no limit.

2.2.5. EXAMPLE. If the limit exists, then it is *unique*. That is, if $\lim_{n\to\infty} a_n = L$ and $\lim_{n\to\infty} a_n = M$, then $L = M$. Indeed, if $M \neq L$, let $\varepsilon = |L - M|/2$. If $\lim_{n\to\infty} a_n = L$, use this ε and find N so that $|a_n - L| < \varepsilon$ for all $n \geqslant N$. Then

$$|a_n - M| \geqslant |L - M| - |L - a_n| > |L - M| - |L - M|/2 = \varepsilon.$$

Hence the sequence does not converge to M.

2.2.6. SQUEEZE THEOREM. *Suppose that* $a_n \leqslant b_n \leqslant c_n$ *for* $n \geqslant 1$, *and that* $\lim_{n\to\infty} a_n = L = \lim_{n\to\infty} c_n$. *Then* $\lim_{n\to\infty} b_n = L$.

PROOF. Let $\varepsilon > 0$. Since $\lim_{n\to\infty} a_n = L$, there is an $N_1 \in \mathbb{N}$ so that for all $n \geqslant N_1$, $|a_n - L| < \varepsilon$. Similarly, since $\lim_{n\to\infty} c_n = L$, there is an $N_2 \in \mathbb{N}$ so that for all $n \geqslant N_2$, $|c_n - L| < \varepsilon$. Let $N = \max\{N_1, N_2\}$. If $n \geqslant N$, then

$$L - \varepsilon < a_n \leqslant b_n \leqslant c_n < L + \varepsilon.$$

Therefore $|b_n - L| < \varepsilon$. So $\lim_{n\to\infty} b_n = L$. ■

Here are some more examples.

2.2.7. EXAMPLE. Consider $b_n = \left(1 + \frac{1}{n^2}\right)^n$ for $n \geqslant 1$. By the binomial theorem,

$$b_n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^{2k}} = 1 + \sum_{k=1}^{n} \frac{n(n-1)\dots(n+1-k)}{n^k} \frac{1}{k!n^k}$$

Therefore for $n \geqslant 2$,

$$1 \leqslant b_n \leqslant 1 + \sum_{k=1}^{\infty} \frac{1}{n^k} = 1 + \frac{\frac{1}{n}}{1 - \frac{1}{n}} = 1 + \frac{1}{n-1}.$$

This is crude but sufficient for our purpose. Take $b_n = 1$ and $c_n = 1 + \frac{1}{n-1}$ for $n \geqslant 2$. They both converge to 1. So by the Squeeze Theorem, $\lim_{n\to\infty} b_n = 1$.

2.2.8. EXAMPLE. Let $a_1 = \frac{1}{2}$, $a_2 = \frac{1}{2+\frac{1}{2}}$, $a_3 = \frac{1}{2+\frac{1}{2+\frac{1}{2}}}$, $a_4 = \frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2}}}}$,

$$a_5 = \frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2}}}}}, \qquad a_6 = \frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2}}}}}}, \qquad a_7 = \frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2}}}}}}}, \quad \dots$$

The limit is called a *continued fraction*. What we need is a formula for a_n. The natural way to do this is to find a recursion formula which defines a_{n+1} in terms of a_n. Here we have $a_1 = \frac{1}{2}$ and $a_{n+1} = \dfrac{1}{2 + a_n}$ for $n \geqslant 1$.

First suppose that the limit L exists. Then we can compute

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

Therefore $L^2 + 2L - 1 = 0$. Thus $L = -1 \pm \sqrt{2}$. However $L \geqslant 0$, so we get $L = \sqrt{2} - 1$. To check that this really is the limit, we must verify the definition.

Note that $L = \frac{1}{2+L}$.

$$a_{n+1} - L = \frac{1}{2+a_n} - \frac{1}{2+L} = \frac{L - a_n}{(2+a_n)(2+L)}.$$

This shows that $a_n - L$ alternates in sign, and we can estimate

$$|a_{n+1} - L| = \frac{|a_n - L|}{(2+a_n)(2+L)} < \frac{|a_n - L|}{4}.$$

Now

$$|a_1 - L| = |\frac{1}{2} - (\sqrt{2} - 1)| = \frac{3}{2} - \sqrt{2} < \frac{1}{10}.$$

Therefore

$$|a_{n+1} - L| < \frac{|a_n - L|}{4} < \frac{|a_{n-1} - L|}{4^2} < \frac{|a_1 - L|}{4^n} < \frac{1}{10}4^{-n}.$$

We can now show that $\lim_{n\to\infty} a_n = \sqrt{2} - 1$.

Suppose that $\varepsilon = \frac{1}{2}10^{-100}$ because we want 100 decimals accuracy. We would require n so that $\frac{1}{10}4^{-n} < \frac{1}{2}10^{-100}$, or $10^{99} < 2^{2n-1}$. Use the simple fact that $2^{10} = 1024 > 10^3$. Then $10^{99} < 2^{330}$, so that $n = 166$ works.

2.2.9. Example. Consider $\lim_{n\to\infty} \dfrac{5n^{100} + 3\cdot 2^n + 7n!}{3n^{100} + 2^n + 5n!}$. The question here is which term dominates as $n \to \infty$, n^{100}, 2^n or $n!$? First, polynomials grow more slowly than exponentials. The way to see this is to consider the ratio $b_n = \dfrac{n^{100}}{2^n}$. Observe that

$$\frac{b_{n+1}}{b_n} = \Big(\frac{n+1}{n}\Big)^{100}\frac{1}{2} \longrightarrow \frac{1}{2}.$$

So eventually this ratio is decreasing almost by a factor of 2 each time, and thus $b_n \to 0$. This shows that 2^n grows faster than n^{100}. Let's deal with $n!$ in a similar fashion. Set $c_n = \dfrac{n!}{2^n}$. Then

$$\frac{c_{n+1}}{c_n} = \frac{(n+1)!}{n!}\frac{2^n}{2^{n+1}} = \frac{n+1}{2} \longrightarrow +\infty.$$

This shows that $c_n \to \infty$, and hence $n!$ grows more quickly that 2^n. Therefore

$$\lim_{n\to\infty} \frac{5n^{100} + 3\cdot 2^n + 7n!}{3n^{100} + 2^n + 5n!} = \lim_{n\to\infty} \frac{5n^{100}/n! + 3\cdot 2^n/n! + 7}{3n^{100}/n! + 2^n/n! + 5} = \frac{7}{5}.$$

We have already been using some rules of manipulation of limits which appear to be true. We record these basic operations.

2.2.10. Proposition. *Suppose that* $\lim_{n\to\infty} a_n = L$ *and* $\lim_{n\to\infty} b_n = M$*, and let* $r \in \mathbb{R}$*. Then*

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

(2) $\lim_{n\to\infty} ra_n = rL$.

(3) $\lim_{n\to\infty} a_n b_n = LM$.

(4) *If $M \neq 0$, then there is an N_0 so that $b_n \neq 0$ for $n \geqslant N_0$, and* $\lim_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{L}{M}$.

PROOF. I will prove (4) and leave the others as exercises. First take $\varepsilon = |M|/2$. Find an N_0 so that for $n \geqslant N_0$, $|b_n - M| < |M|/2$. Then

$$|b_n| \geqslant |M| - |M|/2 = |M|/2.$$

In particular, $b_n \neq 0$. Calculate

$$\begin{aligned}\frac{L}{M} - \frac{a_n}{b_n} = \frac{Lb_n - Ma_n}{Mb_n} &= \frac{Lb_n - LM + LM - Ma_n}{Mb_n} \\ &= \frac{L(b_n - M) + M(L - a_n)}{Mb_n}.\end{aligned}$$

Therefore for $n \geqslant N_0$,

$$\begin{aligned}\Big|\frac{L}{M} - \frac{a_n}{b_n}\Big| &\leqslant \frac{|L|\,|b_n - M| + |M|\,|L - a_n|}{|M|\,|M|/2} \\ &= \frac{2|L|}{M^2}|b_n - M| + \frac{2}{|M|}|a_n - L|.\end{aligned}$$

Now let $\varepsilon > 0$ be given. Use the two limits to find N_1 so that if $n \geqslant N_1$, then $|a_n - L| < \dfrac{\varepsilon|M|}{4}$; and choose N_2 so that if $n \geqslant N_2$, then $|b_n - M| < \dfrac{\varepsilon M^2}{4|L| + 1}$. Define $N = \max\{N_0, N_1, N_2\}$. If $n \geqslant N$, then

$$\begin{aligned}\Big|\frac{L}{M} - \frac{a_n}{b_n}\Big| &< \frac{2|L|}{M^2}|b_n - M| + \frac{2}{|M|}|a_n - L| \\ &< \frac{2|L|}{M^2}\,\frac{\varepsilon M^2}{4|L|+1} + \frac{2}{|M|}\,\frac{\varepsilon|M|}{4} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Therefore $\lim_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{L}{M}$. ■

2.2.11. PROPOSITION. *Every convergent sequence is bounded.*

PROOF. Suppose that $\lim_{n\to\infty} a_n = L$. Take $\varepsilon = 1$ and find N so that if $n \geqslant N$, then $|a_n - L| < 1$. Then

$$M = \max\{|a_1|, |a_2, \dots, |a_{N-1}|, |L| + 1\}$$

is a bound for $\{a_n : n \geqslant 1\}$. ■

2.3. Least upper bound Principle

2.3.1. DEFINITION. If $\varnothing \neq S \subset \mathbb{R}$, say S is *bounded above* if it has an *upper bound* $R \in \mathbb{R}$, meaning that $s \leqslant R$ for all $s \in S$. A number M is the *least upper bound* or *supremum* of S if M is an upper bound for S and whenever R is an upper bound, then $M \leqslant R$. We write $M = \sup S$.

Similarly S is *bounded below* if there is some $R \in \mathbb{R}$ so that $R \leqslant s$ for all $s \in S$. The *greatest lower bound* or *infimum* of S is the lower bound L which is largest among all lower bounds. We write $L = \inf S$.

If S has no upper bound, we write $\sup S = +\infty$; and if it has no lower bound, we write $\inf S = -\infty$.

2.3.2. EXAMPLES.

(1) If our universe is $\mathbb{Q}$, the set of rational numbers, then some sets do not have a supremum. Consider $A = \{x \in \mathbb{Q} : x^2 < 2\}$. It is not hard to show that $\sup A = \sqrt{2}$. However since $\sqrt{2}$ is not rational, the upper bounds in $\mathbb{Q}$ are all strictly bigger than $\sqrt{2}$. Then because $\mathbb{Q}$ is order dense in $\mathbb{R}$, if $r \in \mathbb{Q}$ is an upper bound, we can find another $s \in \mathbb{Q}$ such that $\sqrt{2} < s < r$. Hence there is no best choice. This is an important difference between $\mathbb{Q}$ and $\mathbb{R}$.

(2) $A = \{1, -e, 6, \sqrt{91}, -3.5, \pi\}$. Then $\sup A = \sqrt{91}$ and $\inf A = -3.5$.

(3) $B = \{2, 4, 6, 8, \dots\} = 2\mathbb{N}$. Then $\sup B = +\infty$ and $\inf B = 2$.

(4) $C - \{(-1)^n \frac{n}{n+1} : n \geqslant 1\}$. Then $\sup C = 1$ and $\inf C = -1$. Neither ± 1 belongs to C.

2.3.3. EXAMPLE.

$D = \{\sin n : n \in \mathbb{N}\}$. Then 1 is an upper bound and -1 is a lower bound. To figure out the sup and inf, we use the pigeonhole principle. Let $\varepsilon > 0$.

The angle n (*always in radians because this is calculus!!*) only matters modulo integer multiples of 2π. So for each n, let $\theta(n) = n - \lfloor \frac{n}{2\pi} \rfloor 2\pi \in [0, 2\pi)$. No two are the same because if 2π divides $m - n$, say $m - n = 2\pi k$, then $\pi = \frac{m-n}{2k}$ is rational. But π is irrational, so this doesn't happen. (See Appendix A.5.)

Divide $[0, 2\pi)$ into N intervals of length less than ε. We have to take $N > 2\pi/\varepsilon$. Now $\{\theta(n) : n \geqslant 1\}$ is an infinite sequence. There are infinitely many $\theta(n)$'s in N intervals. Hence there are two in one interval, say $|\theta(n) - \theta(m)| < \varepsilon$ for $1 \leqslant n < m$. then

$$\theta(m-n) = \begin{cases} \theta(m-n) & \in (0, \varepsilon) & \text{if } \theta(m) > \theta(n) \\ 2\pi + \theta(m-n) & \in (2\pi - \varepsilon, 2\pi) & \text{if } \theta(m) < \theta(n) \end{cases}.$$

Now $\sin \frac{\pi}{2} = 1$. Suppose that $\theta(m - n) = \alpha \in (0, \varepsilon)$. Let $k = \lfloor \frac{\pi}{2\alpha} \rfloor$ and look at $\theta(km - kn)$. This belongs to the interval $\left(\frac{\pi}{2} - \varepsilon, \frac{\pi}{2}\right)$, say $\theta(km - kn) = \frac{\pi}{2} - \beta$ for $0 < \beta < \varepsilon$. If, on the other hand, if $\theta(m - n) = 2\pi - \alpha \in (2\pi - \varepsilon, 2\pi)$, let $k = \lceil \frac{3\pi}{2\alpha} \rceil$ and look at $\theta(km - kn)$. Again, $\theta(km - kn) = \frac{\pi}{2} - \beta$ for $0 < \beta < \varepsilon$. Therefore

$$\sin(km - kn) = \sin(\tfrac{\pi}{2} - \beta) = \cos \beta > 1 - \beta^2 > 1 - \varepsilon^2.$$

(We will prove this estimate for $\cos x$ later in the course.) Since $\varepsilon > 0$ was arbitrary, we obtain that $\sup D = 1$. Similarly, $\inf D = -1$.

Now we will establish a very important property of the real numbers.

2.3.4. LEAST UPPER BOUND PRINCIPLE. *Every non-empty set $S \subset \mathbb{R}$ which is bounded above has a least upper bound. Every non-empty set $S \subset \mathbb{R}$ which is bounded below has a greatest lower bound.*

PROOF. We prove the second statement.

Assume that S has a lower bound M, and we can take M to be an integer. Let $s \in S$, and let $k = \lceil s - M \rceil$. Consider $M, M + 1, M + 2, \ldots, M + k$. Since $M + k \geqslant s$, there is a largest integer in this list, say a_0 so that a_0 is a lower bound for S and $a_0 + 1$ is not. Choose $s_0 \in S$ so that $s_0 < a_0 + 1$. This is a 'witness' to the fact that $a_0 + 1$ is not a lower bound. Now consider the numbers $a_0.0, a_0.1, \ldots, a_0.9$. Pick the largest value $a_1 \in \{0, 1, \ldots, 9\}$ so that $a_0.a_1$ is a lower bound. Select a witness $s_1 \in S$ so that $s_1 < a_0.a_1 + \frac{1}{10}$.

Repeat this procedure recursively. Suppose that $a_0.a_1a_2 \ldots a_n$ is a lower bound for S, and there is an $s_n \in S$ so that $s_n < a_0.a_1a_2 \ldots a_n + 10^{-n}$. Consider $a_0.a_1a_2 \ldots a_n0, \ldots, a_0.a_1a_2 \ldots a_n9$ and pick the largest $a_{n+1} \in \{0, 1, \ldots, 9\}$ so that $a_0.a_1a_2 \ldots a_na_{n+1}$ is a lower bound. Then select a witness $s_{n+1} \in S$ so that $s_{n+1} < a_0.a_1a_2 \ldots a_na_{n+1} + 10^{-n-1}$ to show that this is not a lower bound.

Let $L = a_0.a_1a_2a_3 \ldots$. If $s \in S$, then $s \geqslant a_0.a_1a_2 \ldots a_n$ for all $n \geqslant 0$. Therefore $s \geqslant L$. If $L < b$, then there is a finite decimal so that

$$L < c = c_0.c_1c_2 \ldots c_n < b.$$

Thus

$$a_0.a_1a_2 \ldots a_n \leqslant L \leqslant s_n < a_0.a_1a_2 \ldots a_n + 10^{-n} \leqslant c < b.$$

The witness s_n shows that b is not a lower bound. Therefore L is the greatest lower bound.

For the first part, we observe that $\sup S = -\sup(-S)$. ∎

2.4. Monotone Sequences

2.4.1. DEFINITION. Say that $(a_n)_{n\geqslant 1}$ is a *increasing sequence* if $a_n \leqslant a_{n+1}$ for all $n \geqslant 1$; and say that it is a *strictly increasing sequence* if $a_n < a_{n+1}$ for all $n \geqslant 1$. Sometimes we say that (a_n) is *monotone increasing* for emphasis. Similarly we define *decreasing sequence* and *strictly decreasing sequence*.

2.4.2. MONOTONE CONVERGENCE THEOREM. *If (a_n) is an increasing sequence which is bounded above, then (a_n) converges, i.e., $\lim_{n\to\infty} a_n = L$ exists. If (a_n) is a decreasing sequence which is bounded below, then (a_n) converges.*

PROOF. Suppose M is an upper bound for $S = \{a_n : n \geqslant 1\}$. Let $L = \sup S$, which exists by the Least Upper Bound Principle. We claim that $\lim_{n\to\infty} a_n = L$.

Let $\varepsilon > 0$. Then $L - \varepsilon$ is not a lower bound for S. Hence there is an N so that $L - \varepsilon < a_N$. For $n \geqslant N$, $L - \varepsilon < a_N \leqslant a_n \leqslant L$. Therefore $|L - a_n| < \varepsilon$. Thus $\lim_{n\to\infty} a_n = L$.

If (a_n) is decreasing, then $(-a_n)$ is increasing with limit $\sup\{-a_n : n \geqslant 1\}$. Therefore (a_n) has limit $L = \inf\{a_n : n \geqslant 1\}$. ■

2.4.3. EXAMPLE. Let $a_1 = 1$ and $a_{n+1} = \sqrt{2 + \sqrt{a_n}}$ for $n \geqslant 1$.

Claim: a_n is increasing. Indeed, $a_2 = \sqrt{3} > a_1$. Proceed by induction. If $a_n > a_{n-1}$, then

$$a_{n+1} = \sqrt{2 + \sqrt{a_n}} > \sqrt{2 + \sqrt{a_{n-1}}} = a_n.$$

Hence by induction, $a_{n+1} > a_n$ for all $n \geqslant 1$.

Claim: $a_n \leqslant 2$ for all $n \geqslant 1$. Again this is true for $a_1 = 1$. If $a_n < 2$, then $a_{n+1} = \sqrt{2 + \sqrt{a_n}} < \sqrt{2 + \sqrt{2}} < 2$.

Therefore (a_n) is monotone increasing and bounded above. By the Monotone Convergence Theorem (MCT), $L = \lim_{n\to\infty} a_n$ exists. Hence

$$L = \lim_{n\to\infty} a_{n+1} == \lim_{n\to\infty} \sqrt{2 + \sqrt{a_n}} = \sqrt{2 + \sqrt{L}}$$

So $L^2 = 2 + \sqrt{L}$; whence $(L^2 - 2)^2 = L$, or

$$0 = L^4 - 4L^2 - L + 4 = (L - 1)(L^3 + L^2 - 3L - 4).$$

Now $L \geqslant a_2 = \sqrt{3}$, so $L \neq 1$. Thus L is a root of the cubic $p(x) = x^3 + x^2 - 3x - 4$.

Now $p'(x) = 3x^2 + 2x - 3 = 3(x^2 - 1) + 2x > 0$ on $[1, 2]$. Thus p is strictly increasing and $p(1) = -5$ and $p(2) = 2$. The curve must cross the axis, and it has exactly one root between 1 and 2. There is a formula for cubics, though it is not very enlightening:

$$L = \frac{1}{3}\left(\sqrt[3]{\frac{79+\sqrt{2241}}{2}} + \sqrt[3]{\frac{79-\sqrt{2241}}{2}} - 1\right) \approx 1.831177.$$

2.4.4. EXAMPLE. Let $a_1 = 2$ and $a_{n+1} = \dfrac{1+a_n^2}{2}$ for $n \geqslant 1$. First we show that (a_n) is increasing. Now $a_2 = \frac{5}{2} > 2 = a_1$. If $a_{n-1} < a_n$, then

$$a_n = \frac{1+a_{n-1}^2}{2} < \frac{1+a_n^2}{2} = a_{n+1}.$$

Thus the sequence is increasing by induction.

Suppose that $L = \lim_{n\to\infty} a_n$. Then

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

Therefore $0 = L^2 - 2L + 1 = (L-1)^2$; so that $L = 1$. But this is absurd, because $a_n \geqslant 2$. What went wrong?

The problem is that this sequence isn't bounded, and thus does not converge. In fact, $a_n > n$ for $n \geqslant 1$. We have seen this for $n = 1, 2$; and $a_3 = \frac{29}{8} > 3$. Suppose that $n \geqslant 3$ and $a_n > n$. Then $a_{n+1} = \dfrac{1+a_n^2}{2} > \dfrac{1+n^2}{2}$; and

$$\frac{1+n^2}{2} - (n+1) = \frac{1}{2}(n^2 - 2n - 1) > \frac{1}{2}n(n-3) \geqslant 0.$$

By induction, $a_n > n$ for all n, and the sequence is unbounded.

It is convenient to be able to describe the divergence to infinity precisely.

2.4.5. DEFINITION. If $(a_n)_{n\geqslant 1}$ is a sequence, then $\lim_{n\to\infty} a_n = +\infty$ means that for all $R > 0$, there is an integer N so that for $n \geqslant N$, $a_n > R$. And $\lim_{n\to\infty} a_n = -\infty$ is defined analogously.

The notion of being close to a limit value L is replaced by eventually being greater than any large number R. In our example above, given R, we just choose N so that $N \geqslant R$ and for $n \geqslant N$, $a_n > n > R$.

2.5. Subsequences

2.5.1. DEFINITION. A *subsequence* of $(a_n)_{n\geqslant 1}$ is a sequence $(a_{n_i})_{i\geqslant 1}$ where $n_1 < n_2 < n_3 < \dots$.

The following proposition is elementary.

2.5.2. PROPOSITION. *Suppose that* $\lim_{n\to\infty} a_n = L$ *and that* $(a_{n_i})_{i\geqslant 1}$ *is a subsequence of* $(a_n)_{n\geqslant 1}$*. Then* $\lim_{i\to\infty} a_{n_i} = L$*.*

PROOF. Let $\varepsilon > 0$ and select an N so that $|a_n - L| < \varepsilon$ for $n \geqslant N$. Then if $i \geqslant N$, then $n_i \geqslant i \geqslant N$, so $|a_{n_i} - L| < \varepsilon$. Thus $\lim_{i\to\infty} a_{n_i} = L$. ■

We showed in Proposition 2.2.11 the elementary fact that convergent sequences are bounded. The following result is deep, and deals with a partial converse. It has a very interesting proof.

2.5.3. BOLZANO-WEIERSTRASS THEOREM. *Every bounded sequence has a convergent subsequence.*

PROOF. Suppose that $|a_n| \leqslant B$ for $n \geqslant 1$. Let $I_0 = [-B, B]$, and split this interval into two halves, $I_{0-} = [-B, 0]$ and $I_{0+} = [0, B]$. One (and possibly both) of these two intervals must contain infinitely many terms of the sequence. Let $I_1 = [b_1, c_1]$ be such an interval, and pick n_1 so that $a_{n_1} \in I_1$. Split I_1 into two halves $I_{1-} = \left[b_1, \frac{b_1+c_1}{2}\right]$ and $I_{1+} = \left[\frac{b_1+c_1}{2}, c_1\right]$. Again, at least one of these intervals, say $I_2 = [b_2, c_2]$, contains infinitely many terms of the sequence. Pick an $n_2 > n_1$ so that $a_{n_2} \in I_2$.

We repeat this procedure recursively. Suppose that we have a nested sequence of intervals $I_1 \supset I_2 \supset \cdots \supset I_m = [b_m, c_m]$ so that the length $|I_k| = 2^{-k}|I_0|$ for $1 \leqslant k \leqslant m$ so that I_m contains infinitely many terms of the sequence, and that we have selected $n_1 < n_2 < \cdots < n_m$ so that $a_{n_k} \in I_k$ for $1 \leqslant k \leqslant m$. Split I_m into two halves $I_{m-} = \left[b_m, \frac{b_m+c_m}{2}\right]$ and $I_{m+} = \left[\frac{b_m+c_m}{2}, c_m\right]$. Pick one, say $I_{m+1} = [b_{m+1}, c_{m+1}]$, which contains infinitely many elements of the sequence. Then select $n_{m+1} > n_m$ so that $a_{n_{m+1}} \in I_{m+1}$.

Our subsequence is $(a_{n_i})_{i\geqslant 1}$. Observe that because of the construction of the nested intervals I_m, we have that

$$b_1 \leqslant b_2 \leqslant \cdots \leqslant b_m \leqslant a_{n_m} \leqslant c_m \leqslant \cdots \leqslant c_2 \leqslant c_1$$

and $c_m - b_m = 2^{-m}|I_0| = 2^{1-m}B$. The sequence (b_m) is increasing and bounded above by c_1. By the MCT, $\lim_{m\to\infty} b_m = L$ exists. Also (c_m) is a decreasing sequence bounded below by b_1, and thus by MCT, $\lim_{m\to\infty} c_m = M$ exists. Moreover

$$M - L = \lim_{m\to\infty} c_m - b_m = 0.$$

so $M = L$. Finally $\lim_{i\to\infty} a_{n_i} = L$ follows from the Squeeze Theorem. ■

2.5.4. EXAMPLE. In Example 2.3.3, we considered the sequence $(\sin n)$. What are the possible limits of subsequences of this? The way the argument worked is that we found terms in this sequence approaching 1 by approximating the angle $\frac{\pi}{2}$ by $\theta(n_i)$ for various integers n_i. There was nothing special about $\frac{\pi}{2}$ except that it was where $\sin x$ takes the value 1. We can do the same thing for any angle. Thus given $L \in [-1, 1]$, we can find positive integers m_i (not necessarily increasing) so that $\sin m_i$ converges to L.

To select an increasing sequence, look back at the construction. The number

$$km - kn \geqslant k = \left\lfloor \frac{\pi}{2\alpha} \right\rfloor > \left\lfloor \frac{\pi}{2\varepsilon} \right\rfloor.$$

This is large when ε is very small. So what we do is to choose the sequence recursively, and if $n_1 < \cdots < n_k$ are defined, we select ε so small that the lower bound is greater than n_k. Then the terms that we choose will be increasing. Thus every value in $[-1, 1]$ is a limit of a subsequence of $(\sin n)$.

2.6. Completeness

Given a sequence (a_n), is it possible to decide if it will converge without identifying a limit? The answer is crucial to explaining why the real line has no "holes" in it, while the rational numbers has many. A sequence like Example 2.2.9 is a sequence of rational numbers with an irrational limit, $\sqrt{2} - 1$. How do we know that there aren't sequences of real numbers converging to something in a larger universe? The answer is the notion of completeness, which relies on a good answer to the question just posed.

2.6.1. PROPOSITION. *If* $\lim_{n\to\infty} a_n = L$ *and* $\varepsilon > 0$*, there is an integer* N *so that for all* $N \leqslant m \leqslant n$*,* $|a_n - a_m| < \varepsilon$*.*

PROOF. Use $\frac{\varepsilon}{2}$ is the definition of limit, and find N so that for $n \geqslant N$, we have $|a_n - L| < \frac{\varepsilon}{2}$. Then if $N \leqslant m \leqslant n$,

$$|a_n - a_m| \leqslant |a_n - L| + |L - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \qquad \blacksquare$$

2.6.2. DEFINITION. A sequence (a_n) is a *Cauchy sequence* if for every $\varepsilon > 0$, there is an integer N so that for all $N \leqslant m \leqslant n$, $|a_n - a_m| < \varepsilon$.

We can use essentially the same proof as for Proposition 2.2.11 to show that Cauchy sequences are bounded.

2.6.3. PROPOSITION. *Every Cauchy sequence is bounded.*

PROOF. Suppose that (a_n) is a Cauchy sequence. Take $\varepsilon = 1$ and find N so that for all $N \leqslant m \leqslant n$, $|a_n - a_m| < \varepsilon$. Let

$$M = \max\{|a_1|, |a_2, \dots, |a_{N-1}|, |a_N| + 1\}.$$

If $n \geqslant N$, then $|a_n| \leqslant |a_N| + |a_n - a_N| < |a_N| + 1$. ■

Like convergent sequences, Cauchy sequences don't have multiple limit points.

2.6.4. LEMMA. *Let (a_n) be a Cauchy sequence. If there is a convergent subsequence $\lim_{i\to\infty} a_{n_i} = L$, then $\lim_{n\to\infty} a_n = L$.*

PROOF. Let $\varepsilon > 0$ be given. Using $\frac{\varepsilon}{2}$, select N so that $|a_n - a_m| < \frac{\varepsilon}{2}$ for all $N \leqslant m \leqslant n$. Using the limit, find an integer I so that if $i \geqslant I$, then $|a_{n_i} - L| < \frac{\varepsilon}{2}$. Select some $i_0 \geqslant I$ so that $n_{i_0} \geqslant N$. Then if $n \geqslant N$,

$$|a_n - L| \leqslant |a_n - a_{n_{i_0}}| + |a_{n_{i_0}} - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

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

2.6.5. DEFINITION. A set $S \subseteq \mathbb{R}$ is *complete* if every Cauchy sequence in S converges to a point in S.

2.6.6. COMPLETENESS THEOREM. *$\mathbb{R}$ is complete.*

PROOF. Let (a_n) be a Cauchy sequence. By Proposition 2.6.3, it is bounded. By the Bolzano-Weierstrass Theorem, there is a convergent subsequence (a_{n_i}). Then by Lemma 2.6.4, the whole sequence converges. Therefore $\mathbb{R}$ is complete. ■

2.6.7. EXAMPLE. Let $a_0 = 0$ and $a_n = \dfrac{2}{3a_n + 5}$ for $n \geqslant 0$. The first few terms are $0, \frac{2}{5}, \frac{10}{31}, \frac{62}{185}, \frac{370}{1111}, \dots$ which is approximately

$$0, 0.4, 0.32258, 0.335135, 0.3330333, \dots$$

We will show that $(a_n)_{n\geqslant 1}$ is a Cauchy sequence. First

$$a_{n+1} - a_n = \frac{2}{3a_n + 5} - \frac{2}{3a_{n-1} + 5} = \frac{6(a_{n-1} - a_n)}{(3a_n + 5)(3a_{n-1} + 5)}.$$

Since it is clear that $a_n \geqslant 0$ for all n,

$$|a_{n+1} - a_n| < \frac{6|a_n - a_{n-1}|}{25} < \frac{|a_n - a_{n-1}|}{4}.$$

Therefore,

$$|a_{n+1} - a_n| < \frac{|a_{n-1} - a_{n-2}|}{4^2} < \frac{|a_{n-2} - a_{n-3}|}{4^3} < \frac{|a_1 - a_0|}{4^n}.$$

Now, since $a_1 - a_0 = \frac{2}{5}$, if $N \leqslant m < n$,

$$|a_n - a_m| = \Big|\sum_{i=m}^{n-1} a_{i+1} - a_i\Big| \leqslant \sum_{i=m}^{n-1} |a_{i+1} - a_i|$$
$$< \frac{2}{5}\sum_{i=m}^{n-1} 4^{-i} < \frac{2}{5}4^{-m}\frac{1}{1-\frac{1}{4}} = \frac{8}{15}4^{-m} \leqslant \frac{8}{15}4^{-N}.$$

Given $\varepsilon > 0$, we choose N so large that $\frac{8}{15}4^{-N} < \varepsilon$, and we see that (a_n) is Cauchy. By the Completeness Theorem, $L = \lim_{n\to\infty} a_n$ exists. Therefore,

$$L = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{2}{3a_n + 5} = \frac{2}{3L+5}.$$

Thus $0 = 3L^2 + 5L - 2 = (3L-1)(L+2)$. So $L \in \{\frac{1}{3}, -2\}$ and we know that $L \geqslant 0$. Therefore $\lim_{n\to\infty} a_n = \frac{1}{3}$.

We have established a number of results which all say something rather similar. In particular, we started by establishing the Least Upper Bound Principle. This was used to derive the Monotone Convergence Theorem. Then, we used MCT to prove the Bolzano-Weierstrass Theorem. And finally we used the Bolzano-Weierstrass Theorem to prove Completeness of $\mathbb{R}$. **Let's go full circle**, and use the Completeness Theorem to prove the Least Upper Bound Principle. This will show that each of these results is an equivalent formulation of completeness.

Suppose that $S \subset \mathbb{R}$ is non-empty and bounded above. To get started suppose that $s_0 \in S$ and $s \leqslant M$ for all $s \in S$. Let $L = (s_0 + M)/2$. If L is not an upper bound for S, pick $s_1 \in S$ with $s_1 > L$ and set $M_1 = M$. Otherwise, if L is an upper bound, set $M_1 = L$ and $s_1 = s_0$. Either way, $M_1 - s_1 \leqslant \frac{1}{2}(M - s_0)$. Repeat this procedure recursively to construct an increasing sequence s_n in S and a decreasing sequence M_n of upper bounds for S so that $\lim_{n\to\infty} M_n - s_n = 0$. The sequence (s_n) is Cauchy because if $\varepsilon > 0$, select N so that $M_N - s_N < \varepsilon$. Then if $N \leqslant m \leqslant n$, then

$$s_N \leqslant s_m \leqslant s_n \leqslant M_N < s_N + \varepsilon.$$

Therefore $|s_n - s_m| < \varepsilon$. By the Completeness Theorem, $L = \lim_{n\to\infty} s_n$ exists. Moreover $\lim_{n\to\infty} M_n = \lim_{n\to\infty} s_n + (M_n - s_n) = L$. Thus no number smaller than L is an upper bound, but L is an upper bound; and so it is the supremum of S.

2.7. Some Topology

It is convenient to introduce some notation that is used to describe two special classes of sets of real numbers.

2.7.1. DEFINITION. A subset U of $\mathbb{R}$ is *open* if for each $x \in U$, there is an $r > 0$ so that $(x - r, x + r) \subset U$. In particular, $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ is an *open interval*.

A subset A of $\mathbb{R}$ is *closed* if it contains all of its *limit points*; i.e., if $a_n \in A$ and $\lim_{n\to\infty} a_n = b$, then $b \in A$. In particular, $[a, b] = \{x \in \mathbb{R} : a \leqslant x \leqslant b\}$ is a *closed interval*.

2.7.2. EXAMPLES.

(1) An open interval (a, b) is open. If $x \in (a, b)$, then $r = \min\{x - a, b - x\} > 0$ and $(x - r, x + r) \subset (a, b)$.

(2) A closed interval $[a, b]$ is closed. For if $x_n \in [a, b]$ and $\lim_{n\to\infty} x_n = y$, then since $a \leqslant x_n \leqslant b$, we get $a \leqslant y \leqslant b$.

(3) A union of open sets if open: if U_n are open, then $U = \bigcup_{n\geqslant 1} U_n$ is open. For if $x \in U$, there is some n so that $x \in U_n$. Hence there is some $r > 0$ so that $(x - r, x + r) \subset U_n \subset U$.

(4) The intersection of two open sets is open: if U and V are open, and $x \in U \cap V$, then there are $r_1, r_2 > 0$ so that $(x - r_1, x + r_1) \subset U$ and $(x - r_2, x + r_2) \subset V$. Take $r = \min\{r_1, r_2\}$ and note that $(x - r, x + r) \subset U \cap V$.

2.7.3. PROPOSITION. *A set U is open if and only if $U^c = \mathbb{R}\backslash U$ is closed.*

PROOF. Suppose that U is open, and $a_n \in U^c$ and $\lim_{n\to\infty} a_n = b$. If $b \in U$, then for some $r > 0$, $(b - r, b + r) \subset U$. But then there is an N so that $|a_n - b| < r$ for all $n \geqslant N$, and they all belong to U, which is false. Thus $b \in U^c$; whence U^c is closed.

Conversely, suppose that A is a closed set. Let $x \in A^c$. If $A \cap (x-r, x+r) \neq \varnothing$ for all $r > 0$, then taking $r = \frac{1}{n}$, we can pick $a_n \in A$ so that $|x - a_n| < \frac{1}{n}$. Then $\lim_{n\to\infty} a_n = x$. Since A is closed, $x \in A$, a contradiction. Thus for some $r > 0$, $(x - r, x + r) \subset A^c$. So A^c is open. ∎

2.7.4. PROPOSITION. *A subset $S \subset \mathbb{R}$ is complete if and only if it is closed.*

PROOF. Suppose that S is complete, and $s_n \in S$ so that the sequence $(s_n)_{n\geqslant 1}$ converges in $\mathbb{R}$, say $\lim_{n\to\infty} s_n = b$. Then (s_n) is a Cauchy sequence in S. Therefore the limit b belongs to S. Thus S is closed.

Conversely, suppose that S is closed. Let (s_n) be a Cauchy sequence in S. Since $\mathbb{R}$ is complete, $\lim_{n\to\infty} s_n = b$ exists in $\mathbb{R}$. Then because S is closed, $b \in S$. Thus S is complete. ∎

Exercises for Chapter 2

1. For the following sequences, determine the limit if it exists, or prove that it does not converge using the definition of limit

(a) $\dfrac{(-1)^n\sqrt{n}\cos n}{n^2+1}$.

(b) $\sin\left(\dfrac{n\pi}{3}\right)$.

(c) $\dfrac{2^{100+5n}}{e^{4n-10}}, \qquad n \geqslant 1.$

(d) $(-1)^{\frac{n(n-1)(n-2)(n-3)}{8}}, \qquad n \geqslant 1.$

2. (a) Let $a_n = \sqrt{n^2+3n-3} - n$ for $n \geqslant 1$. Find the limit.
(b) Using $\varepsilon = \frac{1}{2}10^{-20}$, find an N that works in the limit definition.

3. Suppose that $\lim\limits_{n\to\infty} a_n = L$. Show that $\lim\limits_{n\to\infty} \dfrac{a_1+a_2+\cdots+a_n}{n} = L$.

4. Let $x_0 = 3$ and $x_{n+1} = (x_n + \frac{8}{x_n})/2$ for $n \geqslant 0$.
(a) Assume a limit exists and figure out what L must be.
(b) Set $\varepsilon_n = x_n - L$. Show that $0 < \varepsilon_{n+1} < \varepsilon_n^2/5$.
(c) Hence show that the limit exists.

5. Let $x_0 = 0$ and $x_{n+1} = \sqrt{15-2x_n}$ for $n \geqslant 0$.
(a) Figure out what the limit L should be.
(b) Show by induction that $2 \leqslant x_n \leqslant 4$ for $n \geqslant 1$.
(c) Define $\varepsilon_n = x_n - L$. Find a formula for ε_{n+1} in terms of ε_n. Show that this alternates in sign. Hence prove that $\lim\limits_{n\to\infty} x_n = L$.

6. Let a_0 and a_1 be positive numbers, and set $a_{n+2} = \sqrt{a_{n+1}} + \sqrt{a_n}$ for $n \geqslant 0$.
(a) Show that here is some N so that $a_n \geqslant 1$ for all $n \geqslant N$.
(b) Let $\varepsilon_n = |a_n - 4|$. Show that $\varepsilon_{n+2} \leqslant (\varepsilon_{n+1} + \varepsilon_n)/3$ for $n \geqslant N$.
(c) Hence prove that this sequence converges.

7. Give a careful ε–N proof that: if (a_n) and (b_n) are sequences of real numbers so that $\lim\limits_{n\to\infty} a_n = L$ and $\lim\limits_{n\to\infty} b_n = M$, then $\lim\limits_{n\to\infty} a_n b_n = LM$.

8. Let $a_n = \sqrt[n]{3^n + 5^n}$ for $n \geqslant 1$.
(a) Prove that this sequence is monotone decreasing and bounded below. What can you conclude?
(b) Evaluate $\lim\limits_{n\to\infty} a_n$.

9. A *series* $\sum\limits_{n=1}^{\infty} a_n$ is said to *converge* if the sequence $s_n = \sum\limits_{k=1}^{n} a_k$, for $n \geqslant 1$, converges. Suppose that $a_n \geqslant 0$ for all $n \geqslant 1$. Prove that the following are all equivalent:

(i) $\sum_{n=1}^{\infty} a_n$ converges.
(ii) the sequence $s_n = \sum_{k=1}^{n} a_k$ is bounded above.
(iii) for all $\varepsilon > 0$, there is an N so that $\sum_{k=N+1}^{m} a_k < \varepsilon$ for all $m > N$.

10. Suppose that $(x_n)_{n=1}^{\infty}$ is a sequence such that $\sum_{n=1}^{\infty} |x_{n+1} - x_n|$ converges. Prove that (x_n) is a Cauchy sequence.

11. Let $S = \{x \in \mathbb{R} : x \neq 0,\ 0 < \sin(\frac{1}{x}) < \frac{1}{2}\}$. Find $\sup S$ and $\inf S$.

12. Suppose that $(a_n)_{n=1}^{\infty}$ is a sequence such that

$$a_{2n-1} \leqslant a_{2n+1} \leqslant a_{2n+2} \leqslant a_{2n} \quad \text{for all} \quad n \geqslant 1$$

Prove that the sequence converges if and only if $\lim_{n\to\infty} a_n - a_{n+1} = 0$.

13. (a) Find $\lim_{h\to 0} \dfrac{\sqrt[3]{1+h} - 1}{h}$. HINT: rationalize the numerator.
(b) (i) Show that if $h > -1$ and $h \neq 0$, then $\sqrt[3]{1+h} < 1 + \frac{h}{3}$.
(ii) For any $0 < \varepsilon < \frac{1}{8}$, show that if $0 < h < 4\varepsilon$, then

$$1 + (\frac{1}{3} - \varepsilon)h < \sqrt[3]{1+h}.$$

(iii) For any $0 < \varepsilon < \frac{1}{8}$, show that if $-4\varepsilon < h < 0$, then

$$1 + (\frac{1}{3} + \varepsilon)h < \sqrt[3]{1+h}.$$

(c) Use (b) to provide a different proof of (a).

14. Let $x_0 = 0$ and $x_{n+1} = \sqrt{5 - 2x_n}$ for $n \geqslant 0$.
(a) Compute $x_1, \dots, x_{10}$ on your calculator or computer.
(b) Prove that the even terms are increasing and bounded above by all the odd terms, which are decreasing. HINT: $f(x) = \sqrt{5 - 2x}$ is decreasing.
(c) Get a bound for $|x_{n+1} - x_n|$ in terms of $|x_n - x_{n-1}|$ for $n \geqslant 4$.
(d) Prove that $\lim_{n\to\infty} x_n$ exists, and evaluate this limit.

15. Find a sequence of rational numbers (a_n) so that a real number L is a limit of a subsequence of (a_n) if and only if $e \leqslant |L| \leqslant \pi$.

CHAPTER 3

Functions

We briefly review some the of terminology regarding functions. A *function* is a map $f : X \to Y$ from a set X into a set Y that assigns exactly one value $y = f(x)$ for each $x \in X$. The *domain* of f is the set on which it is defined, and the target space Y is the *codomain*. The *range* of f is the set $\{y = f(x) : x \in X\}$.

A function is *one-to-one* or *injective* if $x_1 \neq x_2 \in X$ implies $f(x_1) \neq f(x_2)$. A function is *onto* or *surjective* if for each $y \in Y$, there is some $x \in X$ so that $f(x) = y$. A function is *one-to-one and onto* or *bijective* if it is both injective and surjective.

When the range is a field like $\mathbb{R}$, say $f, g : X \to \mathbb{R}$, we can add, subtract and multiply functions: $(f \pm g)(x) = f(x) \pm g(x)$ and $(fg)(x) = f(x)g(x)$. We can also multiply them by a real scalar, say $t \in \mathbb{R}$: $(tf)(x) = tf(x)$. If $g(x) \neq 0$ for all $x \in X$, we can divide: $(f/g)(x) = \dfrac{f(x)}{g(x)}$.

If $f : X \to Y$ and $g : Y \to Z$, the *composition* $g \circ f : X \to Z$ is defined by $g \circ f(x) = g(f(x))$. If $f : X \to Y$ is a bijection, then there is a unique function $g : Y \to X$ so that $g \circ f(x) = \mathrm{id}_X(x) = x$ for all $x \in X$, namely, for each $y \in Y$, there is a unique $x \in X$ so that $f(x) = y$; and we set $g(y) = x$. This is called the *inverse function* of f, and is denoted f^{-1}. We also have $f \circ f^{-1} = \mathrm{id}_Y$, so that f is also the inverse of f^{-1}. We will discuss inverse functions in more detail later.

3.1. Limits of functions

We want to extend our definition of limits of sequences to limits of functions.

3.1.1. DEFINITION. Suppose that a real valued function f is defined on an interval $(a - d, a + d)\backslash\{a\}$ for some $d > 0$. The *limit of a function* f as $x \to a$ is L, written $\lim_{x\to a} f(x) = L$, means that for all $\varepsilon > 0$, there is a $\delta > 0$ so that if $0 < |x - a| < \delta$, then $|f(x) - L| < \varepsilon$.

In symbols, $\forall_{\varepsilon>0}\exists_{\delta>0}\forall_{0<|x-a|<\delta}\ |f(x) - L| < \varepsilon$.

3.1.2. REMARKS. **(1)** f does not need to be defined at $x = a$, and in any case, the value $f(a)$ has no bearing on the value of the limit.

(2) When f is defined on $(a, a+d)$, we can talk about the *limit from the right*. We write $\lim_{x\to a^+} f(x) = L$ to mean: for all $\varepsilon > 0$, there is a $\delta > 0$ so that if $a < x < a+\delta$, then $|f(x) - L| < \varepsilon$. Similarly, we define *limit from the left*, $\lim_{x\to a^-} f(x) = L$, if for all $\varepsilon > 0$, there is a $\delta > 0$ so that if $a - \delta < x < a$, then $|f(x) - L| < \varepsilon$.

3.1.3. EXAMPLE. $\lim_{x\to 3} \sqrt{x} = \sqrt{3}$. To prove this, estimate the difference

$$|\sqrt{x} - \sqrt{3}| = \left|\frac{(x-\sqrt{3})(x+\sqrt{3})}{x+\sqrt{3}}\right| = \frac{|x-3|}{\sqrt{x}+\sqrt{3}}.$$

Make an initial choice to control the denominator: if $|x-3| < 1$, then $2 < x < 4$ and so

$$\frac{1}{\sqrt{x}+\sqrt{3}} < \frac{1}{\sqrt{2}+\sqrt{3}} < \frac{1}{3}.$$

Therefore $|\sqrt{x} - \sqrt{3}| < \dfrac{|x-3|}{3}$. Given $\varepsilon > 0$, pick $\delta = \min\{3\varepsilon, 1\}$. Then if $0 < |x-3| < \delta$, we get $|\sqrt{x} - \sqrt{3}| < \dfrac{|x-3|}{3} < \varepsilon$. The first inequality requires $\delta \leqslant 1$ and the second requires $\delta \leqslant 3\varepsilon$. Hence $\lim_{x\to 3} \sqrt{x} = \sqrt{3}$.

3.1.4. EXAMPLE. Let $f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$. Then $\lim_{x\to 0} f(x)$ does not exist. Take any value L, and let $\varepsilon = \frac{1}{2}$.

Case 1. Suppose that $L \leqslant \frac{1}{2}$. For any $\delta > 0$, pick $n \in \mathbb{N}$ so that $0 < \frac{1}{n} < \delta$. Then $|f(\frac{1}{n}) - L| = |1 - L| \geqslant \frac{1}{2} = \varepsilon$. So L is not the limit.

Case 2. Suppose that $L \geqslant \frac{1}{2}$. Given $\delta > 0$, choose $n \in \mathbb{N}$ so that $0 < \frac{\pi}{n} < \delta$. Then $|f(\frac{\pi}{n}) - L| = |L| \geqslant \frac{1}{2} = \varepsilon$. Thus L is not the limit. So the limit does not exist.

Let $g(x) = xf(x) = \begin{cases} x & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$ Then $0 \leqslant g(x) \leqslant x$. Then $\lim_{x\to 0} 0 = 0 = \lim_{x\to 0} x$. Therefore $\lim_{x\to 0} g(x) = 0$ by the function version of the Squeeze Theorem.

3.1.5. SQUEEZE THEOREM FOR FUNCTIONS. *Suppose that f, g, h are functions on $[a-d, a+d]\backslash\{a\}$ such that $f(x) \leqslant g(x) \leqslant h(x)$ and $\lim_{x\to a} f(x) = L = \lim_{x\to a} h(x)$. Then $\lim_{x\to a} g(x) = L$.*

The following analogue of Proposition 2.2.10 is established in the same way as for sequences. We leave the proofs as an exercise.

3.1.6. PROPOSITION. *Suppose that f and g are functions on $(a-d, a+d)$ such that $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$, and let $r \in \mathbb{R}$. Then*

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

(2) $\lim_{x\to a} rf(x) = rL$.

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

(4) *If $M \neq 0$, then there is a $\delta_0 > 0$ so that $g(x) \neq 0$ for $0 < |x-a| < \delta_0$; and*
$\lim_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{L}{M}$.

3.1.7. EXAMPLE. Consider $\lim_{\theta\to 0} \dfrac{\sin\theta}{\theta}$. Remember that we are using radians. Note that if $-\frac{\pi}{2} < \theta < 0$, then $\dfrac{\sin\theta}{\theta} = \dfrac{\sin|\theta|}{|\theta|}$ because this is an even function. So it is enough to work with $0 < \theta < \frac{\pi}{2}$. Draw a circle of radius 1, centre O, and mark the points A and B on rays separated by angle θ. Draw the lines perpendicular to $\overline{OA}$ through A and B, and mark points C and D as in the figure. Note that segments

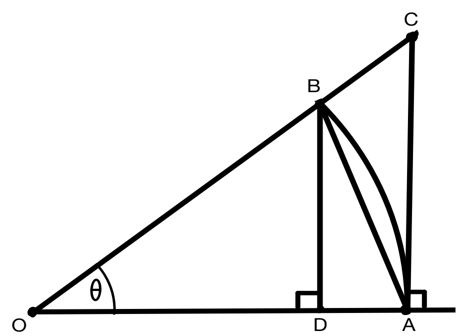


FIGURE 3.1. Limit of $\frac{\sin x}{x}$

$\overline{OA}$ and $\overline{OB}$ have length 1, while $\overline{BD}$ has length $\sin\theta$, and $\overline{AC}$ has length $\tan\theta$. Observe that

$$\triangle OAB \subset \text{sector } OAB \subset \triangle OAD.$$

Thus their areas compare:

$$\frac{1}{2}\sin\theta \leqslant \frac{1}{2}\theta\, 1^2 \leqslant \frac{1}{2}\tan\theta = \frac{\sin\theta}{2\cos\theta}.$$

Rearranging we get

$$\cos\theta \leqslant \frac{\sin\theta}{\theta} \leqslant 1.$$

In particular, $0 < \sin\theta < \theta$ if $0 < \theta < \frac{\pi}{2}$. Therefore if $0 < \theta < 1$,

$$\cos\theta = \sqrt{1 - \sin^2\theta} \geqslant \sqrt{1 - \theta^2} \geqslant 1 - \theta^2.$$

Thus $1 - \theta^2 \leqslant \dfrac{\sin\theta}{\theta} \leqslant 1$. By the Squeeze Theorem, $\displaystyle\lim_{\theta\to 0} \frac{\sin\theta}{\theta} = 1$.

Another important limit that is a consequence of this is

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2} = \lim_{x\to 0} \frac{1 - (1 - 2\sin^2\frac{x}{2})}{x^2} = \lim_{x\to 0} \frac{2\sin^2\frac{x}{2}}{4(\frac{x}{2})^2} = \frac{1}{2}$$

We can use these two limits to compute the derivative of the $\sin x$ and $\cos x$ functions. We will not study the derivative formally until Chapter 5. However most students in this course have seen the derivative in their high school calculus course. So we will make use of that knowledge until we get to the theory later on.

$$\begin{aligned}
\frac{d}{dx}(\sin x)\Big|_{x=a} &= \lim_{h\to 0} \frac{\sin(a+h) - \sin a}{h} \\
&= \lim_{h\to 0} \frac{\sin a \cos h + \cos a \sin h - \sin a}{h} \\
&= \lim_{h\to 0} \sin a \frac{(\cos h - 1)}{h} + \cos a \frac{\sin h}{h} \\
&= \sin a(0) + \cos a(1) = \cos a.
\end{aligned}$$

and

$$\begin{aligned}
\frac{d}{dx}(\cos x)\Big|_{x=a} &= \lim_{h\to 0} \frac{\cos(a+h) - \cos a}{h} \\
&= \lim_{h\to 0} \frac{\cos a \cos h - \sin a \sin h - \cos a}{h} \\
&= \lim_{h\to 0} \cos a \frac{(\cos h - 1)}{h} - \sin a \frac{\sin h}{h} \\
&= \cos a(0) - \sin a(1) = -\sin a.
\end{aligned}$$

Thus

$$\frac{d}{dx}(\sin x) = \cos x \quad \text{and} \quad \frac{d}{dx}(\cos x) = -\sin x.$$

Now I want to graph the function $f(x) = \dfrac{\sin x}{x}$ for $x \neq 0$. Before we draw anything, I will collect some information. First we know that $\sin x$ oscillates between ± 1 with period 2π. Therefore $f(x)$ oscillates between the graphs of $y = \pm\frac{1}{x}$. Since $\sin x$ has zeros at all integer multiples of π, $\pi\mathbb{Z}$, $f(x) = 0$ at $\{n\pi : n \in \mathbb{Z}\backslash\{0\}\}$; but as we appraoch $x = 0$, we have $\displaystyle\lim_{x\to 0} f(x) = 1$. Also $f(x)$ touches the curves $y = \pm\frac{1}{x}$ at alternate odd multiples of $\frac{\pi}{2}$. The function f is *even*, meaning that $f(-x) = \dfrac{\sin(-x)}{-x} = \dfrac{\sin x}{x} = f(x)$.

It won't help much to solve for $f'(x) = 0$ because we know that when the curve touches $y = \pm\frac{1}{x}$, it will be tangent there, but won't have derivative zero. Thinking of $x > 0$, the zeroes of $f'(x)$ will happen a bit before touching the curve $y = \pm\frac{1}{x}$, reaching an extremal point and turning slightly to line up with the other curve. We can't actually solve explicitly for these points, and the information only helps a little bit. For example, the minimum of the function occurs at points $\pm x_0$ where x_0 is a bit smaller than $\frac{3\pi}{2}$. The exception is the behaviour near $x = 0$. Here it helps.

$$f'(x) = \frac{x\cos x - \sin x}{x^2} \quad \text{and} \quad \lim_{x\to 0} f'(x) = \lim_{x\to 0} \frac{\cos x - \frac{\sin x}{x}}{x}.$$

But we just showed that $0 > \cos x - \frac{\sin x}{x} > \cos x - 1$, and $\lim_{x\to 0} \frac{1-\cos x}{x} = 0$. Thus by the Squeeze Theorem, $\lim\limits_{x\to 0} f'(x) = 0$. This means that the curve flattens out as it approached $x = 0$ at $(0, 1)$.

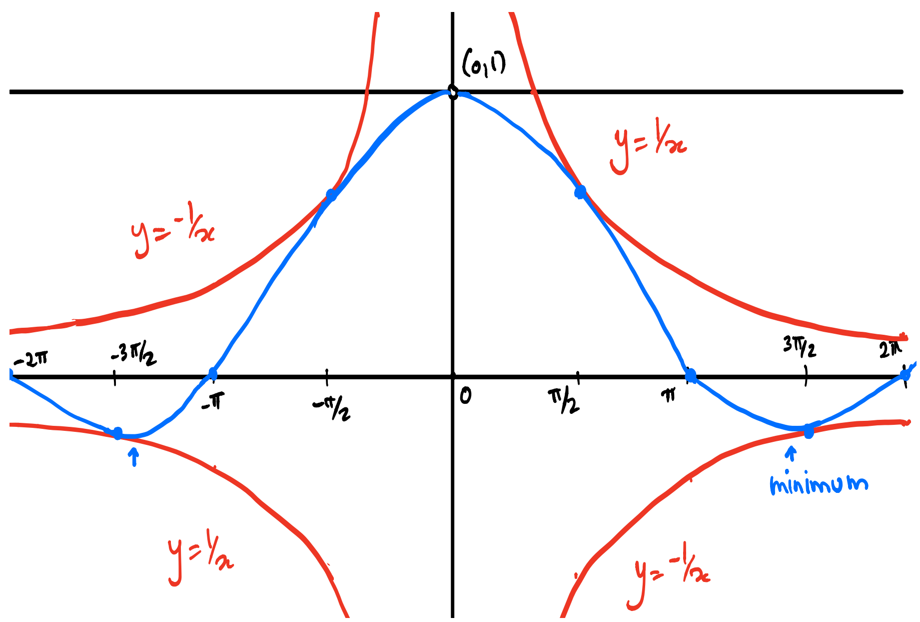


FIGURE 3.2. Graph of $\frac{\sin x}{x}$

3.1.8. EXAMPLE. Graph $g(x) = x\sin\frac{1}{x}$ on $\mathbb{R}\backslash\{0\}$. This is also an even function. As x approaches 0, $\frac{1}{x}$ approaches $\pm\infty$. So $\sin\frac{1}{x}$ oscillates faster and faster between ± 1. So $g(x)$ oscillates between the two functions $y = \pm x$. In particular, $\lim\limits_{x\to 0} g(x) = 0$. The zeroes occur at $\left\{\frac{1}{n\pi} : n \in \mathbb{Z}\backslash\{0\}\right\}$. At the points $\frac{2}{(2n+1)\pi}$ for $n \geqslant 0$, $g(x) = (-1)^n x$. So the largest point where the curve touches the bounding lines is at $x = \frac{2}{\pi}$. Use the even property to reflect this over to the negative real axis.

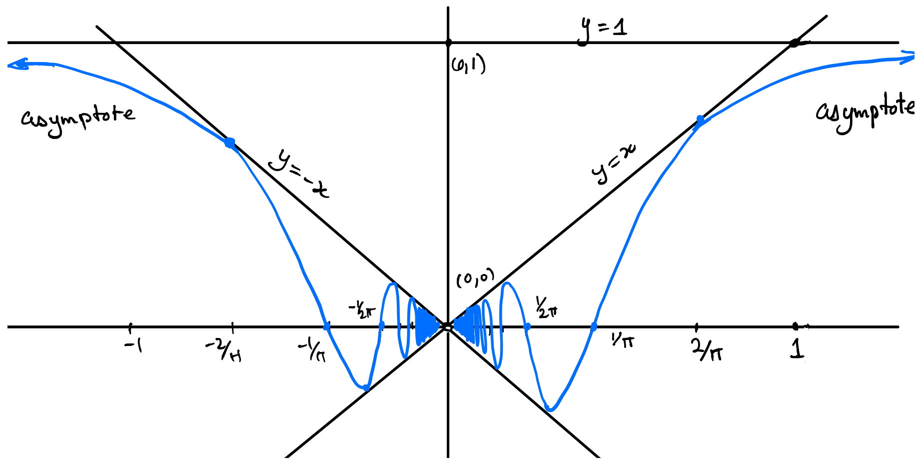


FIGURE 3.3. Graph of $x \sin \frac{1}{x}$

This explains the behaviour as we approach $x = 0$. What happens as $x \to \infty$? In this limit, substitute $u = \frac{1}{x}$. As $x \to \infty$, $u \to 0^+$.

$$\lim_{x\to\infty} x \sin \tfrac{1}{x} = \lim_{u\to 0^+} \frac{\sin u}{u} = 1.$$

Thus this curve has a horizontal asymptote $y = 1$ as $x \to \infty$ and by symmetry, as $x \to -\infty$ as well.

3.2. The natural logarithm

For $0 < a < b$, define $A(a, b)$ to be the area under the graph $y = \frac{1}{x}$ between $x = a$ and $x = b$. Set $A(b, a) = -A(a, b)$ when $b > a$ and $L(a) = A(1, a)$. Observe that $A(a, a) = 0$ and $A(a, b) + A(b, c) = A(a, c)$.

We are not going to do any integral calculus here. Rather I am going to argue geometrically to obtain the properties that we need. For $s > 0$, consider the linear transformation

$$T_s(x, y) = (sx, \tfrac{y}{s}) \quad \text{for} \quad (x, y) \in \mathbb{R}^2.$$

This takes a square S with corners (x, y), $(x + h, y)$, $(x, y + h)$ and $(x + h, y + h)$ to the rectangle $R = T_sS$ with corners $(sx, \frac{y}{s})$, $(sx + sh, \frac{y}{s})$, $(sx, \frac{y}{s} + \frac{h}{s})$, and $(sx + sh, \frac{y}{s} + \frac{h}{s})$. Now S has area h^2 while R has area $(sh)(\frac{h}{s}) = h^2$. Any nice planar region can be approximated by a union of small squares. Since T_s preserves area on each square, it preserves area for any nice region.

Next observe that the points on the curve $y = \frac{1}{x}$ for $x > 0$, say $(x, \frac{1}{x})$, are mapped by T_s to $(sx, \frac{1}{sx})$. So T_s maps the curve onto itself. It similarly maps the x-axis onto itself. The line $x = a$, which consists of points (a, y), is mapped onto

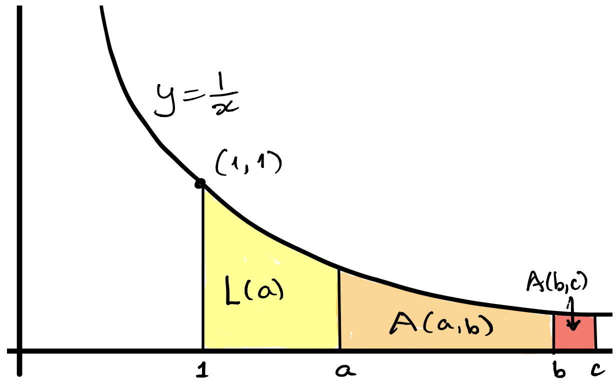


FIGURE 3.4. Area under $y = \frac{1}{x}$

the line $x = sa$. It is now easy to see that the region under the curve $y = \frac{1}{x}$ between $x = a$ and $x = b$ is mapped by T_s onto the region under $y = \frac{1}{x}$ between $x = sa$ and $x = sb$. Therefore, since T_s preserves area,

$$A(a, b) = A(sa, sb) \quad \text{for} \quad 0 < a < b \text{ and } s > 0.$$

3.2.1. PROPOSITION. *For $a, b > 0$, $L(ab) = L(a) + L(b)$.*

PROOF.
$$L(ab) = A(1, ab) = A(1, a) + A(a, ab)$$
$$= A(1, a) + A(1, b) = L(a) + L(b).$$
■

3.2.2. COROLLARY. *For $a > 0$ and $n \in \mathbb{Z}$, $L(a^n) = nL(a)$.*

PROOF. First consider $n \geqslant 0$. Clearly $L(a^0) = L(1) = A(1, 1) = 0$ and $L(a^1) = L(a)$. Proceed by induction. Assume that the formula is true for n. Then

$$L(a^{n+1}) = L(aa^n) = L(a) + L(a^n) = L(a) + nL(a) = (n + 1)L(a).$$

Thus by induction, the formula is valid for all $n \in \mathbb{N}_0$.

Next consider $n = -1$.

$$L(\tfrac{1}{a}) = A(1, \tfrac{1}{a}) = -A(\tfrac{1}{a}, 1) = -A(1, a) = -L(a).$$

Finally, if $n \in \mathbb{N}$, $L(a^{-n}) = -L(a^n) = -nL(a)$. So the formula is valid for $n \in \mathbb{Z}$.
■

3.2.3. COROLLARY. *$L(x)$ is a strictly increasing function on $(0, \infty)$, and*

$$\lim_{x\to\infty} L(x) = +\infty \quad \text{and} \quad \lim_{x\to 0^+} L(x) = -\infty.$$

PROOF. Strictly increasing is clear because $A(a, b) > 0$ for $a < b$. Therefore

$$\lim_{x\to\infty} L(x) = \lim_{n\to\infty} L(2^n) = \lim_{n\to\infty} nL(2) = +\infty$$

because $L(2) > 0$. Similarly,

$$\lim_{x\to 0^+} L(x) = \lim_{n\to -\infty} L(2^n) = \lim_{n\to -\infty} nL(2) = -\infty. \qquad \blacksquare$$

3.2.4. PROPOSITION. *For* $a, b > 0$, $\dfrac{1}{b} < \dfrac{A(a,b)}{b-a} < \dfrac{1}{a}$. *Thus for* $x > 1$, $\dfrac{1}{x} < \dfrac{L(x)}{x-1} < 1$.

PROOF. Notice that the region with area $A(a, b)$ contains the rectangle on $[a, b]$ with height $\frac{1}{b}$ and is contained in the rectangle on $[a, b]$ with height $\frac{1}{a}$. Therefore

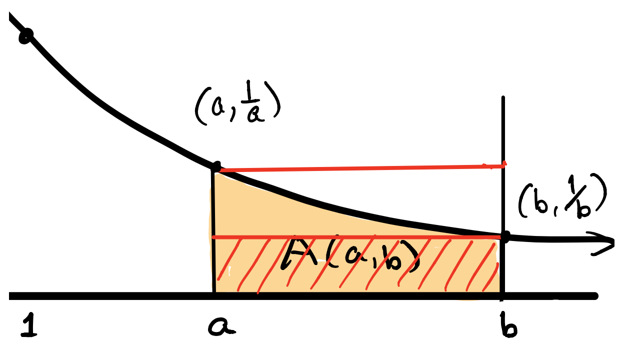


FIGURE 3.5. Bounds for $A(a, b)$

$$\frac{b-a}{b} \leqslant A(a,b) \leqslant \frac{b-a}{a}.$$

Divide by $b - a$. Take $a = 1$ and $b = x$ to get the second formula. $\blacksquare$

3.2.5. COROLLARY. $\dfrac{d}{dx} L(x) = \dfrac{1}{x}$ *for* $x > 0$.

PROOF. For $x > 0$ and $h > 0$, we have

$$\frac{1}{x+h} < \frac{A(x, x+h)}{h} = \frac{L(x+h) - L(x)}{h} < \frac{1}{x}.$$

Let $h \to 0^+$ and use the Squeeze Theorem to get

$$\lim_{h\to 0^+} \frac{L(x+h) - L(x)}{h} = \frac{1}{x}.$$

Similarly if $0 < h < x$,

$$\frac{1}{x} < \frac{A(x-h,x)}{h} = \frac{L(x-h)-L(x)}{-h} < \frac{1}{x-h}.$$

Thus

$$\lim_{h\to 0^-} \frac{L(x+h)-L(x)}{h} = \frac{1}{x}.$$

Therefore $L'(x) = \dfrac{1}{x}$. ■

3.2.6. DEFINITION. The *natural logarithm*, written $\ln x$ or $\log x$, is defined as $\ln x = L(x)$ for $x > 0$. The number e is the unique real number such that $\ln e = 1$.

The main property of any logarithm function is given in Proposition 3.2.1: $\ln(ab) = \ln a + \ln b$. The *base* of the logarithm is the number, in this case e, such that $\ln e = 1$. This specific number is chosen so that the derivative is $\frac{1}{x}$, not some multiple. Corollary 3.2.2 shows that $\ln e^n = n$, but in fact it says much more. Consider a rational number $\frac{m}{n}$. Then

$$m = \ln e^m = \ln((e^{m/n})^n) = n \ln e^{m/n}.$$

Therefore $\ln e^{m/n} = \frac{m}{n}$.

If we have a log function such that a number $a > 1$ had log equal to 1, we call this function $\log_a x$. Because of the product property, it follows that $\log_a a^{\frac{m}{n}} = \frac{m}{n}$ and $\ln a^{\frac{m}{n}} = \frac{m}{n} \ln a$. Therefore $\log_a x = \dfrac{\ln x}{\ln a}$. Therefore $\dfrac{d}{dx} \log_a x = \dfrac{1}{x \ln a}$.

We want to get a rough estimate for the value of e. We will find more precise methods later. First estimate $\ln 2 = A(1,2)$ is bounded above by the trapezoid with vertices $(1,0), (1,1), (2,\frac{1}{2}), (2,0)$ which has base 1 and average height $\frac{1}{2}(1+\frac{1}{2}) = \frac{3}{4}$. So $\ln 2 < 0.75$. Next look at $\ln 2.5$. From 2 to 2.5, we bound the curve by

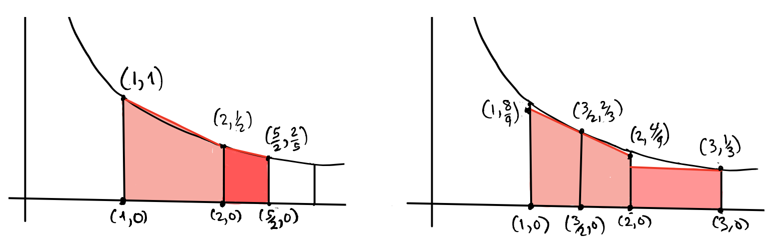


FIGURE 3.6. Estimating e

another trapezoid including vertices $(2.5, 0.4)$ and $(2.5, 0)$. Its area is greater than $A(2, 2.5)$. This trapezoid has base $\frac{1}{2}$ and average height $\frac{1}{2}(\frac{1}{2}+\frac{2}{5}) = \frac{9}{20}$, so it adds an additional $\frac{9}{40}$. Therefore $\ln 2.5 < \frac{3}{4} + \frac{9}{40} = \frac{39}{40} < 1$. Hence $2.5 < e$.

On the other hand, the area $A(1,3)$ under $\frac{1}{x}$ from 1 to 3 is bounded below by a trapezoid on $[1,2]$ with upper edge tangent to $\frac{1}{x}$ at $(\frac{3}{2},\frac{2}{3})$ plus the rectangle on $[2.3]$ of height $\frac{1}{3}$. This provides a lower bound $\ln 3 = A(1,3) > 1(\frac{2}{3}) + \frac{1}{3} = 1$. See the figures. Thus $2.5 < e < 3$.

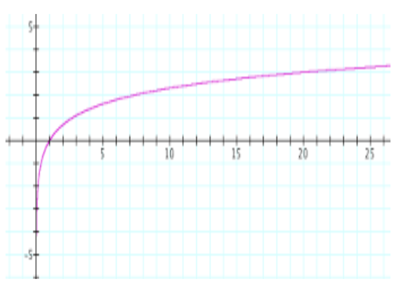
3.3. The exponential function

Since $\ln x$ is strictly increasing on $(0,\infty)$ and maps onto the whole line, $\mathbb{R}$, there is an inverse function $E : \mathbb{R} \to (0,\infty)$ satisfying

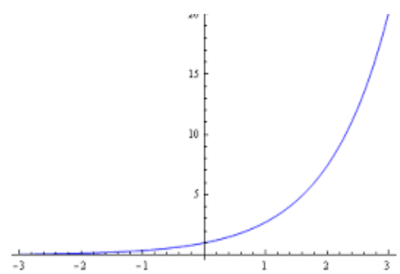
$$E(\ln x) = x \text{ for } x > 0 \quad \text{and} \quad \ln(E(x)) = x \text{ for } x \in \mathbb{R}.$$

In particular, E is strictly increasing, and $E(0) = 1$ and $E(1) = e$. Each point (x,y) on the graph of $\ln x$ converts to the point (y,x) on the graph of E.

Since $\ln(ab) = \ln a + \ln b$, apply E to get $ab = E(\ln a + \ln b)$. Substitute $x = \ln a$ and $y = \ln b$, which are arbitrary real numbers, to get $E(x)E(y) = E(x+y)$. Thus $E(x)^n = E(nx)$ and so $E(\frac{x}{n})^n = E(x)$. Thus $E(\frac{mn}{n}) = E(\frac{x}{n})^m = E(x)^{\frac{m}{n}}$. Now take $x = 1$ to get $E(\frac{m}{n}) = e^{\frac{m}{n}}$. As E is monotone increasing, we obtain that $E(x) = e^x$ for all real numbers. This is known as the *exponential function*.



(A) Graph of $\ln x$



(B) Graph of e^x

3.3.1. EXAMPLES. Let's try to understand the growth rates of these functions more precisely.

(1) $\lim\limits_{x\to\infty} \dfrac{\ln x}{x}$. If $e^n \leqslant x \leqslant e^{n+1}$, we have $n \leqslant \ln x \leqslant n+1$, and hence

$$\frac{n}{e^{n+1}} \leqslant \frac{\ln x}{x} \leqslant \frac{n+1}{e^n}.$$

Claim: $\lim\limits_{n\to\infty} \dfrac{n}{e^n} = 0$. Set $a_n = \dfrac{n}{e^n}$ and compute

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{ne} \leqslant \frac{2}{e} < 0.8.$$

Therefore by repeated application, we get $a_{n+1} \leqslant (.8)^n a_1 \to 0$. Thus by the Squeeze Theorem, $\lim_{x\to\infty} \frac{\ln x}{x} = 0$.

(2) Let $a > 0$ (think of a as small). Then in the next limit, substitute $y = x^a$ and note that as $x \to \infty$, then $y \to \infty$.

$$\lim_{x\to\infty} \frac{\ln x}{x^a} = \lim_{y\to\infty} \frac{\ln y^{1/a}}{y} = \frac{1}{a} \lim_{y\to\infty} \frac{\ln y}{y} = 0.$$

This says that $\ln x$ grows more slowly than any positive power of x, and thus increases very slowly for large x. The curve gets flatter and flatter as x increases.

(3) This converts to a statement about e^x. Here we substitute $y = e^x$, and note that as $x \to \infty$, then $y \to \infty$.

$$\lim_{x\to\infty} \frac{e^x}{x} = \lim_{y\to\infty} \frac{y}{\ln y} = +\infty.$$

Then for any $b > 0$ (think of b as very large), and again substitute $y = e^x$,

$$\lim_{x\to\infty} \frac{e^x}{x^b} = \lim_{y\to\infty} \frac{y}{(\ln y)^b} = \lim_{y\to\infty} \left(\frac{y^{1/b}}{\ln y}\right)^b = +\infty.$$

Hence e^x grows faster than any power of x.

(4) Now consider the behaviour of $\ln x$ as $x \to 0^+$. Again let $a > 0$. So as $x \to 0^+$, we have $\ln x \to -\infty$ and $x^a \to 0$. We will substitute $y = \frac{1}{x}$, so that $y \to \infty$.

$$\lim_{x\to 0^+} x^a \ln x = \lim_{y\to\infty} y^{-a} \ln y^{-1} = \lim_{y\to\infty} \frac{-\ln y}{y^a} = 0.$$

Moreover it takes negative values. So while $\ln x \to -\infty$, it goes more slowly than x^a goes to 0 for any $a > 0$.

(5) If I substitute $y = -x$ in the following:

$$\lim_{x\to-\infty} |x|^b e^x = \lim_{y\to\infty} \frac{y^b}{e^y} = 0.$$

This shows that e^x goes to zero faster than any polynomial x^n goes to ∞ as $x \to -\infty$.

3.3.2. EXAMPLE. Graph $f(x) = x \ln x$ for $x > 0$. Then $f(x) = 0$ only at $x = 1$, where it changes from negative to positive. It increases to infinity a bit faster than a straight line. The interesting behaviour is between 0 and 1. We have $\lim_{x\to 0^+} f(x) = 0$. Now look at the derivative: $f'(x) = \ln x + 1$. The is monotone increasing, so the curve always curves upwards as a faster rate as x increases. Note that $0 = f'(x)$ implies that $\ln x = -1$, so that $x = \frac{1}{e}$. This is a minimum at $(\frac{1}{e}, -\frac{1}{e})$.

Notice also that $\lim\limits_{x\to 0^+} f'(x) = -\infty$. This means that the curve is becoming vertical. There is a vertical tangent at the limit point $(0,0)$. See the figure.

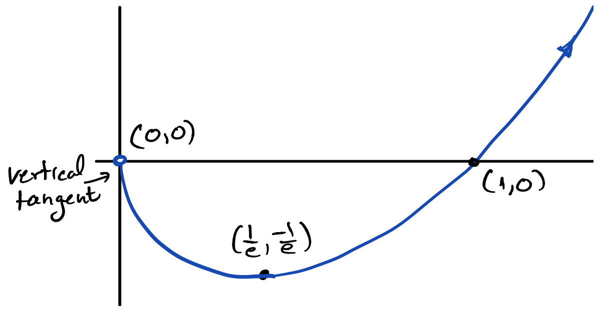


FIGURE 3.8. Graph of $x \ln x$.

3.3.3. PROPOSITION. $\dfrac{d}{dx}(e^x) = e^x$ *for* $x \in \mathbb{R}$.

PROOF. Apply the chain rule to the identity $\ln(e^x) = x$ to get

$$\frac{1}{e^x}\frac{d}{dx}(e^x) = 1 \quad \text{or} \quad \frac{d}{dx}(e^x) = e^x. \qquad \blacksquare$$

3.3.4. PROPOSITION. $\lim\limits_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x$ *for* $x \in \mathbb{R}$.

PROOF. By Proposition 3.2.4 applied to $1 + \frac{x}{n}$, we have

$$\frac{1}{1+\frac{x}{n}} < \frac{\ln(1+\frac{x}{n})}{x/n} < \frac{1}{1}.$$

Therefore

$$\frac{x}{1+\frac{x}{n}} < \ln\left(1 + \tfrac{x}{n}\right)^n < x.$$

Exponentiate to get

$$e^{\frac{x}{1+\frac{x}{n}}} < \left(1 + \tfrac{x}{n}\right)^n < e^x.$$

Now let $n \to \infty$. By the Squeeze Theorem, $\lim\limits_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x$. $\blacksquare$

In the next result, we need Bernouilli's inequality.

3.3.5. LEMMA. $(1+x)^n \geqslant 1 + nx$ *for* $1 + x > 0$ *and* $n \geqslant 1$.

PROOF. Proceed by induction on n. For $n = 1$, $1 + x = 1 + x$. Assume that it is true for $n - 1$. Then

$$(1+x)^n = (1+x)^{n-1}(1+x) \geqslant (1+(n-1)x)(1+x) = 1+nx+(n-1)x^2 \geqslant 1+nx.$$

Hence it is true for n. By induction, this holds for all $n \geqslant 1$. ■

3.3.6. PROPOSITION. *Let $a_n = (1 + \frac{1}{n})^n$ and $b_n = (1 + \frac{1}{n})^{n+1}$. Then a_n is monotone increasing, b_n is monotone decreasing and*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = e$$

Similarly $c_n = (1 - \frac{1}{n})^{n-1}$ and $d_n = (1 - \frac{1}{n})^n$ are monotone decreasing and increasing respectively with limit $\frac{1}{e}$.

PROOF. Set $x = 1$ in the previous Proposition and get $\lim\limits_{n\to\infty} a_n = e^1$. Also

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} a_n(1 + \tfrac{1}{n}) = e(1) = e.$$

Compute, using Bernouilli's inequality at the last step.

$$\begin{aligned} \frac{a_{n+1}}{b_n} &= \Big(\frac{n+2}{n+1}\Big)^{n+1}\Big(\frac{n}{n+1}\Big)^{n+1} = \Big(\frac{(n+2)n}{(n+1)^2}\Big)^{n+1} \\ &= \Big(1 - \frac{1}{(n+1)^2}\Big)^{n+1} \geqslant 1 - \frac{n+1}{(n+1)^2} = \frac{n}{n+1}. \end{aligned}$$

Thus $a_{n+1} \geqslant \frac{n}{n+1} b_n = a_n$. Similarly we can invert this to get

$$\frac{b_n}{a_{n+1}} = \Big(1 + \frac{1}{n^2+2n}\Big)^{n+1} \geqslant 1 + \frac{n+1}{n^2+2n} > \frac{n+2}{n+1}.$$

Thus $b_{n+1} = \frac{n+2}{n+1} a_{n+1} < b_n$.

Note that $c_n = \dfrac{1}{a_{n-1}}$ and $d_n = \dfrac{1}{b_{n-1}}$. The monotonicity and limit follows. ■

3.3.7. EXAMPLE. We will graph a more complicated function, $g(x) = xe^{1/x}$. We first collect as much information as we can about zeros and limits approaching $\pm\infty$ and any points where $g(x)$ is undefined. Then we will look for critical points.

- g is undefined at $x = 0$. Also g has no zeros.
- Check the behaviour as $x \to 0$. Substitute $y = \frac{1}{x}$ in the limits.

$$\lim_{x\to 0^+} xe^{1/x} = \lim_{y\to+\infty} \frac{e^y}{y} = +\infty.$$

This is a vertical *asymptote*. Notice that the limits from the two sides differ!

$$\lim_{x\to 0^-} xe^{1/x} = \lim_{y\to-\infty} \frac{e^y}{y} = 0^-.$$

- Behaviour at $+\infty$. $\lim_{x\to+\infty} xe^{1/x} = +\infty$. However $e^{1/x} \to 1$, so in some sense $xe^{1/x} \approx x$. To see how good an approximation this is, we compute another limit. Substitute $y = \frac{1}{x}$.

$$\lim_{x\to+\infty} xe^{1/x} - x = \lim_{x\to+\infty} \frac{e^{1/x}-1}{1/x}$$
$$= \lim_{y\to0^+} \frac{e^y - 1}{y} = \frac{d}{dy}(e^y)|_{y=0} = e^0 = 1.$$

  This means that $g(x)$ approaches the line $y = x + 1$ as $x \to \infty$. So this line is an *oblique asymptote*.

- Behaviour at $-\infty$. $\lim_{x\to-\infty} xe^{1/x} = -\infty$, and again looks like x.

$$\lim_{x\to-\infty} xe^{1/x} - x = \lim_{y\to0^-} \frac{e^y-1}{y} = \frac{d}{dy}(e^y)|_{y=0} = 1.$$

  Thus the line $y = x + 1$ is an oblique asymptote as $x \to -\infty$ as well.

- $g'(x) = e^{1/x} + xe^{1/x}(\frac{-1}{x^2}) = e^{1/x}(1 - \frac{1}{x})$. So $g'(1) = 0$ is the only critical point. Since $g(x) \to +\infty$ as $x \to 0^+$ and as $x \to +\infty$, there is a local minimum at $(1, e)$.

- We should also check how g approaches 0 as $x \to 0^-$.

$$\lim_{x\to0^-} g'(x) = \lim_{x\to0^-} (1 - \tfrac{1}{x})e^{1/x} = \lim_{y\to-\infty} (1-y)e^y = 0.$$

  So there is a horizontal tangent at the limit point $(0, 0)$.

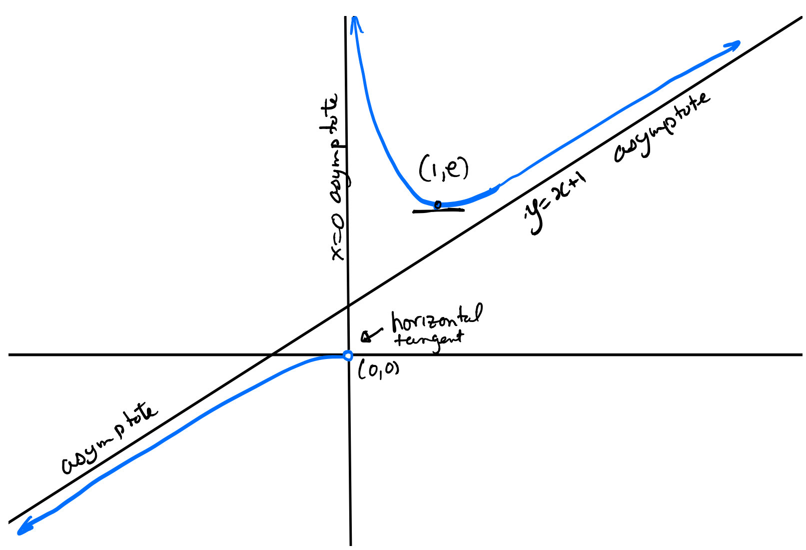


FIGURE 3.9. Graph of $g(x)$.

Exercises for Chapter 3

1. Give a careful ε–δ argument to prove that $\displaystyle\lim_{x\to 8}\frac{3x^{1/3}}{x+4}=\frac{1}{2}$.

2. Compute the following limits:

(a) $\displaystyle\lim_{x\to 1}\frac{3}{x^3-1}-\frac{4}{x^4-1}$

(b) $\displaystyle\lim_{x\to 1}\frac{\sin(x^2-1)}{x-1}$

(c) $\displaystyle\lim_{x\to 0}\frac{\tan x-\sin x}{x^3}$

(d) $\displaystyle\lim_{x\to +\infty} x^{3/2}\big(\sqrt{x+2}+\sqrt{x}-2\sqrt{x+1}\big)$

3. Consider the functions $f(x)=x\sin^2(\frac{1}{x})$ for $x\neq 0$

(a) Where does the graph of this function touch the curves $y=0$, $y=x$ and $y=\frac{1}{x}$?

(b) Graph the function $f(x)$ (*by hand!*). Include graphs of the the auxillary curves $y=x$ and $y=\frac{1}{x}$ on the same graph and show the intersection points from (a). Pay attention to the behaviour of $f(x)$ as x approaches 0 and $\pm\infty$ and identify any asymptotes.
Do not try to compute local maxima and minima or inflection points.

4. Compute the following limits:

(a) $\displaystyle\lim_{x\to +\infty}\frac{\ln x^{1000}}{x^{1/1000}}$

(b) $\displaystyle\lim_{x\to 0} x^{-4}e^{-1/x^2}$

(c) $\displaystyle\lim_{x\to 1}\frac{x^2-1}{\ln x}$

(d) $\displaystyle\lim_{n\to +\infty}\big(1+\tfrac{1}{n}\big)^{n^2}\big(1+\tfrac{1}{n+1}\big)^{-(n+1)^2}$

5. Graph the function $f(x)=\dfrac{1-\cos x}{x^2}$ for $x\neq 0$. Draw the auxillary curve $y=\dfrac{2}{x^2}$ on the same graph and explain where these two curves touch and where f touches the x-axis. Pay attention to the behaviour of $f(x)$ as x approaches 0 and $\pm\infty$ and identify any asymptotes. Do not try to compute local maxima and minima or inflection points.

6. (a) Graph the function $f(x)=\dfrac{\ln x}{x}$ for $x>0$. Pay attention to the behaviour of $f(x)$ as x approaches 0^+ and $+\infty$, asymptotes, maxima and minima, zeros and points of inflection (i.e. points where $f''(x)$ changes sign).

(b) Which is larger, 3^π or π^3?

7. Two of the hyperbolic trig functions are

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \quad \text{and} \quad \cosh(x) = \frac{e^x + e^{-x}}{2}.$$

(a) Sketch the functions $e^x/2$, $e^{-x}/2$, $\sinh(x)$ and $\cosh(x)$ on the same graph. Pay attention to the relationships between the four curves. *I am not expecting a detailed graph here.*
(b) Show that $\cosh^2(x) - \sinh^2(x) = 1$ for all $x \in \mathbb{R}$.
(c) Show that $\sinh(x+y) = \sinh(x)\cosh(y) + \cosh(x)\sinh(y)$ for all $x, y \in \mathbb{R}$.
(d) Solve $\sinh(x) = y$ for x as a function of y.

8. Consider the function $f(x) = xe^{\frac{3x-1}{x^2}}$ for $x \neq 0$. Graph the function $f(x)$ (*by hand!*). Pay attention to the following (show your work):

- asymptotic behaviour at $\pm\infty$ and behaviour at 0.
- compute the derivative and find the critical points, including $\lim_{x\to 0} f'(x)$.
- compute any points of inflection.
- choose a scale that illustrates the key features.

9. For which values of $t > 1$ does the expression $t^{t^{t^{t^{\cdot^{\cdot^{\cdot}}}}}}$ make sense?
HINT: fix $t > 1$ and define $a_0 = 1$ and $a_{n+1} = t^{a_n}$ for $n \geqslant 0$. The question asks when this sequence has a limit. Try $t = \sqrt{2}$ and $t = 2$ on the computer to see what happens.
(a) Show that $a_{n+1} > a_n$ for all $n \geqslant 1$. What does this tell you?
(b) When $L = \lim_{n\to\infty} a_n$ exists, solve for t in terms of L. Use this to find the optimal upper bound for those values of t for which the limit exists. What happens for larger t?
(c) For these values of t, show by induction that a_n is bounded above by e for all $n \geqslant 1$. What does this tell you?

Note: the behaviour when $0 < t < 1$ is very interesting, but much trickier. Compute a few terms using $t = \frac{1}{16}$ and see what occurs. Compare with $t = \frac{1}{4}$.

CHAPTER 4

Continuity

4.1. Continuous functions

We introduce an extremely important property of functions.

4.1.1. DEFINITION. Suppose that $f : [b, c] \to \mathbb{R}$ and $b < a < c$. The function f is *continuous* at a if $\lim_{x\to a} f(x) = f(a)$. That is, for $\varepsilon > 0$, there is a $\delta > 0$ so that if $|x - a| < \delta$, then $|f(x) - f(a)| < \varepsilon$. If it is not continuous at a, then it is *discontinuous* at a.

At endpoints of a closed interval, use one-sided limits. So f is continuous at b if $\lim_{x\to b^+} f(x) = f(b)$; and f is continuous at c if $\lim_{x\to c^-} f(x) = f(c)$. We say that $f(x)$ is a *continuous function* on a set X if it is continuous at each $a \in X$.

4.1.2. REMARK. In the definition of $\lim_{x\to a} f(x)$, we specify that $0 < |x-a| < \delta$. But here we said $|x-a| < \delta$. This is fine because $f(a)$ is defined and is the putative limit. In particular. $|f(a) - f(a)| = 0 < \varepsilon$ is always true.

4.1.3. EXAMPLES.

(1) Let $f(x) = \frac{1}{x}$ for $x \neq 0$. Fix $a \neq 0$. Then

$$|f(x) - f(a)| = \left|\frac{1}{x} - \frac{1}{a}\right| = \frac{|a - x|}{|ax|}.$$

Let $\varepsilon > 0$. If $0 < \delta \leqslant \delta_0 = \frac{|a|}{2}$, then $|x-a| < \delta$ implies $|x| \geqslant |a|-|x-a| \geqslant |a|/2$. Therefore $\dfrac{|a-x|}{|ax|} \leqslant \dfrac{2|a-x|}{a^2}$. Now choose $\delta = \min\{\frac{a^2\varepsilon}{2}, \frac{|a|}{2}\}$. Then $|x - a| < \delta$ implies that $|f(x) - f(a)| < \dfrac{2\delta}{a^2} \leqslant \varepsilon$. Therefore $f(x)$ is continuous on $\mathbb{R}\backslash\{0\}$. Since $\lim_{x\to 0^+} \dfrac{1}{x} = +\infty$, there is no way to define f at 0 to make it continuous there.

(2) Consider $\sin x$. Use the addition formula

$$\sin(a+h)-\sin a = \sin a \cos h + \cos a \sin h - \sin a = \sin a(\cos h - 1) + \cos a \sin h.$$

Using the estimates from Example 3.1.7 for $|h| < \pi/2$,

$$|\sin(a+h) - \sin a| \leqslant 1|\cos h - 1| + 1|\sin h| < h^2 + |h|.$$

Therefore $\lim_{h\to 0} \sin(a+h) = \sin a$. Hence $\sin x$ is continuous.

(3) A function $f : [a,b] \to \mathbb{R}$ is *Lipschitz* with *Lipschitz constant* L if

$$|f(x) - f(y)| \leqslant L|x-y| \quad \text{for all} \quad x, y \in [a,b].$$

Given $\varepsilon > 0$, we can take $\delta = \varepsilon/L$. Then for any x, y in the domain with $|x-y| < \delta$, we have $|f(x) - f(y)| \leqslant L\delta = \varepsilon$. Therefore Lipschitz functions are continuous.

(4) Let $f(x) = \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$. Then $f(x)$ is continuous at $x = 0$ because of Example 3.1.7. It is continuous everywhere else by general facts that we will establish in the next section.

(5) Let $f(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$. Then $f(x)$ is continuous on $\mathbb{R}\backslash\{0\}$, but is discontinuous at 0.

(6) Let $f(x) = \begin{cases} x & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{R}\backslash\mathbb{Q}. \end{cases}$ Then $f(x)$ is continuous at 0, but discontinuous at every $a \neq 0$.

(7) Let $f(x) = \sin \frac{1}{x}$. This has a bad discontinuity at 0. There is no way to define $f(0)$ to make it continuous because every value in $[-1, 1]$ is a limit of $f(x)$ along some subsequence approaching 0.

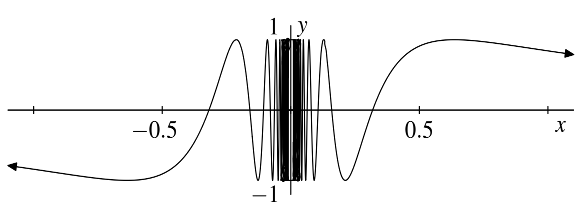


FIGURE 4.1. Graph of $\sin(1/x)$.

(8) *Thomae's function*. Let $f(x) = \begin{cases} \frac{1}{q} & \text{if } x = \frac{p}{q} \in \mathbb{Q},\ \gcd(p,q) = 1,\ q > 0. \\ 0 & \text{if } x \in \mathbb{R}\backslash\mathbb{Q}. \end{cases}$

Claim: $f(x)$ is discontinuous on $\mathbb{Q}$ and continuous on $\mathbb{R}\backslash\mathbb{Q}$.

Every $a \in \mathbb{R}$ is a limit of irrational numbers. So if $a \in \mathbb{Q}$, then since $f(a) \neq 0$, f is discontinuous at a. Now suppose that a is irrational, and $\varepsilon > 0$. Choose N so

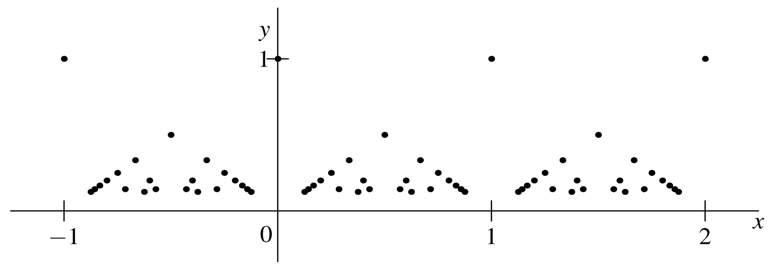


FIGURE 4.2. Thomae's function on $[-1, 2]$

that $\frac{1}{N} \leqslant \varepsilon$. The set $X = \{\frac{p}{q} : 1 \leqslant q \leqslant N,\ p \in \mathbb{Z}\} \cap [a-1, a+1]$ is a finite set of rational numbers. Thus $\operatorname{dist}(a, X) = \min\{|a-x| : x \in X\} = \delta > 0$. If $|x-a| < \delta$, then either $x \notin \mathbb{Q}$ and $f(x) - f(a) = 0$ or $x = \frac{p}{q}$ with $q > N$ and $\gcd(p, q) = 1$. Hence $|f(x) - f(a)| \leqslant \frac{1}{N+1} < \varepsilon$. Therefore f is continuous at a.

4.2. Properties of Continuous Functions

The standard operations of functions that we use preserve continuity.

4.2.1. PROPOSITION. *Suppose that $f, g : [b, c] \to \mathbb{R}$ are both continuous at a. Then $rf + sg$ and fg are continuous at a for any $r, s \in \mathbb{R}$. If $g(a) \neq 0$, then f/g is also continuous at a.*

PROOF. This is an immediate consequence of Proposition 3.1.6. We will use the ε–δ definition to show that fg is continuous at a. We need to control

$$\begin{aligned}|f(x)g(x) - f(a)g(a)| &= |f(x)g(x) - f(a)g(x) + f(a)g(x) - f(a)g(a)| \\ &\leqslant |f(x) - f(a)|\,|g(x)| + |f(a)|, |g(x) - g(a)|.\end{aligned}$$

First bound $|g(x)|$. Take $\varepsilon_0 = 1$ and find $\delta_0 > 0$ so that $|x-a| < \delta_0$ implies that $|g(x) - g(a)| < 1$; and hence $|g(x)| \leqslant |g(a)| + 1$. Now given $\varepsilon > 0$, find $\delta_1 > 0$ so that $|x-a| < \delta_1$ implies that $|f(x) - f(a)| < \frac{\varepsilon}{2(|g(a)|+1)}$. Also choose $\delta_2 > 0$ so that $|x-a| < \delta_2$ implies that $|g(x) - g(a)| < \frac{\varepsilon}{2(|f(a)|+1)}$. Define $\delta = \min\{\delta_0, \delta_1, \delta_2\}$. If $|x-a| < \delta$, then

$$\begin{aligned}|f(x)g(x) - f(a)g(a)| &\leqslant |f(x) - f(a)|\,|g(x)| + |f(a)|, |g(x) - g(a)| \\ &< \frac{\varepsilon}{2(|g(a)|+1)}(|g(a)|+1) + \frac{\varepsilon}{2(|f(a)|+1)}|f(a)| < \varepsilon.\end{aligned}$$

Therefore $f(x)g(x)$ is continuous at a. ∎

Next we show that composition preserves continuity.

4.2.2. THEOREM. *Suppose that $f : [r,s] \to [u,v]$ and $g : [u,v] \to \mathbb{R}$. Suppose that $a \in (r,s)$ and $f(a) = b$. If $\lim_{x\to a} f(x) = f(a)$ and $\lim_{y\to b} g(y) = g(b)$, then $\lim_{x\to a} g\circ f(x) = g\circ f(a)$; i.e., if f is continuous at a and g is continuous at $f(a)$, then $g\circ f$ is continuous at a. If f and g are both continuous, then so is $g\circ f$.*

PROOF. Let $\varepsilon > 0$ be given. Find $\delta_1 > 0$ so that $|y-b| < \delta_1$ implies that $|g(y)-g(b)| < \varepsilon$. Use this δ_1 as an epsilon in the limit of f to obtain $\delta_2 > 0$ so that $|x-a| < \delta_2$ implies that $|f(x)-f(a)| < \delta_1$. And hence $|g(f(x)) - g(f(a))| < \varepsilon$. Therefore $g\circ f$ is continuous at a. ∎

4.2.3. EXAMPLES.

(1) Trig functions. We saw that $\sin x$ is continuous everywhere. Hence $\cos x = \sin(x + \frac{\pi}{2})$ is continuous. The function $\tan x = \dfrac{\sin x}{\cos x}$ is then continuous except at the points where $\cos x = 0$, namely odd multiples of $\frac{\pi}{2}$. However $\tan x$ is not defined at these points. Similarly, $\cot x$, $\sec x$ and $\csc x$ are continuous where they are defined.

(2) If $a > 0$, we define $a^x = e^{x\ln a}$. This is continuous on $\mathbb{R}$.

(3) Rational functions, which are functions of the form $f(x) = \dfrac{p(x)}{q(x)}$ where p and q are polynomials are continuous except at the roots of q.

(4) $f(x) = \cos(x^2 + \frac{1}{x^3})\sin(e^{3\tan x})$ is continuous except at $\{0, \frac{(2n+1)\pi}{2} : n \in \mathbb{Z}\}$.

4.3. Extreme Value Theorem

This fundamental result describing general conditions which guarantee that a function attains its maximum and minimum values depends both on continuity and on the completeness of the real line.

4.3.1. EXTREME VALUE THEOREM. *Let $f(x)$ be a continuous function on a closed, bounded interval $[a,b]$. Then f is bounded and there is a point $c \in [a,b]$ so that $f(c) = \sup_{a\leqslant x\leqslant b} f(x)$, and a point $d \in [a,b]$ so that $f(d) = \inf_{a\leqslant x\leqslant b} f(x)$.*

PROOF. Let $L = \sup_{a\leqslant x\leqslant b} f(x)$. This could possibly be $+\infty$. Pick real numbers $L_1 < L_2 < L_n < L_{n+1} < \dots$ so that $L = \lim_{n\to\infty} L_n$. (E.g., if $L < \infty$, take $L_n = L - \frac{1}{n}$ and if $L = \infty$, take $L_n = n$.) Since $\sup_{a\leqslant x\leqslant b} f(x) > L_n$, there is an $x_n \in [a,b]$ so that $f(x_n) > L_n$. Now $(x_n)_{n\geqslant 1}$ is a bounded sequence. By the Bolzano-Weierstrass Theorem, there is a convergent subsequence (x_{n_i}) with

$\lim_{i\to\infty} x_{n_i} = c$. Since $[a,b]$ is closed, $c \in [a,b]$. By continuity of f,

$$L \geqslant f(c) = \lim_{i\to\infty} f(x_{n_i}) \geqslant \lim_{i\to\infty} L_{n_i} = L.$$

Therefore $L = f(c) < \infty$ and the supremum is attained.

Similarly, there is a point $d \in [a,b]$ so that $f(d) = \inf_{a\leqslant x\leqslant b} f(x)$. ■

4.3.2. EXAMPLES.

(1) Examples where the maximum is attained are familiar. Like $f(x) = 1 - |x|$ on $[-2,2]$. Then max $f(x) = f(0) = 1$.

(2) The real line, $\mathbb{R}$, is closed but not bounded. The function $f(x) = \frac{-1}{1+x^2}$ is continuous and bounded, but does not attain sup $f(x) = 0$.

(3) The continuous function $f(x) = x$ is not even bounded on $\mathbb{R}$.

(4) Let $f(x) = x$ on $(0,1)$. The interval $(0,1)$ is bounded, but not closed. The continuous function f does not attain its supremum or infimum.

(5) Let $f(x) = \frac{1}{x}$ on $(0,1]$. This is continuous, but unbounded. Again the domain $(0,1]$ is not closed.

(6) Define $f(x) = \begin{cases} 0 & \text{if } x = 0 \\ \frac{1}{x} & \text{if } 0 < x \leqslant 1 \end{cases}$. This function is defined on a closed bounded interval, but is unbounded. The problem is that f is not continuous at 0.

(7) A similar example is $f(x) = \begin{cases} \frac{1}{2} & \text{if } x \in \{0,1\} \\ x & \text{if } 0 < x < 1. \end{cases}$ Again $[0,1]$ is closed and bounded, and here f is bounded, but does not attain its supremum or infimum. The problem is that f is not continuous at 0 or 1.

4.4. Intermediate Value Theorem

There is a second important theorem about continuous functions that relies in an essential way on the completeness of $\mathbb{R}$.

4.4.1. INTERMEDIATE VALUE THEOREM. *Let $f : [a,b] \to \mathbb{R}$ be a continuous function such that $f(a) < L < f(b)$. Then there is a point $c \in (a,b)$ so that $f(c) = L$.*

4.4.2. REMARK. Intuitively this says that if the curve $y = f(x)$ starts below the line $y = L$ at a and arrives above the line at b, then it must cross the line somewhere. However if we define $f : \mathbb{Q} \to \mathbb{Q}$ by $f(x) = x^3$, then $f(0) = 0 < 3 < 8 = f(2)$, but there is no rational number x so that $f(x) = 3$. There is a gap in $\mathbb{Q}$ at $\sqrt[3]{3}$. It is the incompleteness of $\mathbb{Q}$ which allows the result to fail.

PROOF. Let $X = \{x \in [a, b] : f(x) < L\}$. Then $a \in X$ and $X \subset [a, b)$, so it is a non-empty bounded set. Hence $c = \sup X$ exists by the Least Upper Bound Principle, and clearly $c \leqslant b$.

Claim: $c < b$. Let $\varepsilon = f(b) - L > 0$. By continuity of f at b, there is a $\delta > 0$ so that if $|x - b| < \delta$, then $|f(x) - f(b)| < \varepsilon$. Hence $f(x) > f(b) - \varepsilon = L$. so on $(b - \delta, b]$, $f(x) > L$. Thus $X \cap (b - \delta, b] = \varnothing$, and so $c \leqslant b - \delta < b$.

Claim: $f(c) = L$. Any $x > c$ is not in X and so $f(x) > L$. Choose a sequence (a_n) in $(c, b]$ and decreasing to c. Then by continuity,

$$f(c) = \lim_{n\to\infty} f(a_n) \geqslant \lim_{n\to\infty} L = L.$$

On the other hand, there is a sequence of points $x_n \in X$ such that $\lim\limits_{n\to\infty} x_n = c$. Thus,

$$f(c) = \lim_{n\to\infty} f(x_n) \leqslant \lim_{n\to\infty} L = L.$$

Therefore, $f(c) = L$. ■

4.4.3. EXAMPLE. Every polynomial of odd degree has a real root. Write the polynomial as $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where n is odd and $a_n \neq 0$. Observe that

$$\lim_{x\to+\infty} p(x) = \lim_{x\to+\infty} a_n x^n \Big(1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\Big) = \begin{cases} +\infty & \text{if } a_n > 0 \\ -\infty & \text{if } a_n < 0. \end{cases}$$

Similarly,

$$\lim_{x\to-\infty} p(x) = \lim_{x\to-\infty} a_n x^n \Big(1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\Big) = \begin{cases} -\infty & \text{if } a_n > 0 \\ +\infty & \text{if } a_n < 0. \end{cases}$$

Therefore $p(x)$ changes sign. By the Intermediate Value Theorem (IVT), there must be a point x_0 so that $p(x_0) = 0$.

4.4.4. EXAMPLE. Every *monic* polynomial of even degree attains its minimum value. Write $p(x) = x^{2n} + a_{2n-1}x^{2n-1} + \cdots + a_1 x + a_0$, Arguing as in the previous example, we see that

$$\lim_{x\to+\infty} p(x) = \lim_{x\to-\infty} p(x) = +\infty.$$

Therefore there is some large number N so that if $|x| > N$, then $p(x) > p(0)+1$. It follows that the infimum of p occurs in $[-N, N]$. By the Extreme Value Theorem, p attains its minimum value.

4.4.5. EXAMPLE. Let $f(x) = x^{179} + \dfrac{163}{1 + x^2 + \sin^2 x}$. The equation $f(x) = 119$ has a solution. Note that f is continuous since the denominator is always non-zero. Now $f(0) = 163 > 119$ and $f(1) = 1 + \dfrac{163}{2 + \sin^2 1} < 1 + 81.5 < 119$. By IVT, there is a point $c \in (0, 1)$ so that $f(c) = 119$.

4.4.6. COROLLARY. *If $f : [a, b] \to \mathbb{R}$ is continuous, then* Ran(f) *is a closed bounded interval.*

PROOF. By the Extreme Value Theorem, the range of f is bounded, and there are points $x_0, x_1 \in [a, b]$ so that

$$f(x_0) = \inf\{f(x) : a \leqslant x \leqslant b\} \quad \text{and} \quad f(x_1) = \sup\{f(x) : a \leqslant x \leqslant b\}.$$

By the Intermediate Value Theorem, every value y with $f(x_0) \leqslant y \leqslant f(x_1)$ is in the range of f. Thus Ran$(f) = [f(x_0), f(x_1)]$ is a closed bounded interval. ■

When f is defined on an interval I which is not closed or not bounded, then by restricting f to an increasing sequence of closed bounded subintervals with union I, you can conclude that the range is an interval, but it may be open, closed or half open independent of I, and it may be unbounded.

4.4.7. COROLLARY. *If I is an interval (open, closed or half closed) and $f : I \to \mathbb{R}$ is continuous and one-to-one, then f is strictly monotone.*

PROOF. We prove this by contradiction. Notice that f is monotone increasing and one-to-one if and only if $a < b < c$ implies $f(a) < f(b) < f(c)$; and likewise f is monotone decreasing if $a < b < c$ implies $f(a) > f(b) > f(c)$. Thus failure to be monotone means that there are points $a < b < c$ so that $f(a) < f(b) > f(c)$ or the inequalities are reversed, since an equality contradicts injectivity. Either way, there is a value L in $(f(a), f(b)) \cap (f(c), f(b))$. By the IVT, there are points $x_1 \in (a, b)$ and $x_2 \in (b, c)$ so that $f(x_1) = L = f(x_2)$, contradicting the fact that f is one to one. Therefore f is monotone. ■

4.5. Monotone Functions

4.5.1. PROPOSITION. *Let $f : [a, b] \to \mathbb{R}$ be a monotone increasing function, and let $a < c < b$. Then $\lim_{x \to c^-} f(x) = L$ and $\lim_{x \to c^+} f(x) = M$ both exist, and $L \leqslant f(c) \leqslant M$.*

PROOF. Let $A = \{f(x) : a \leqslant x < c\}$. This is a non-empty set bounded above by $f(c)$. Hence $L = \sup A$ exists by the LUBP, and $L \leqslant f(c)$. For $\varepsilon > 0$, $L - \varepsilon$ is not an upper bound for A, and hence there is some $x_0 < c$ so that $f(x_0) > L - \varepsilon$. Set $\delta = c - x_0$. If $c - \delta < x < c$, then $L - \varepsilon < f(x_0) \leqslant f(x) \leqslant L$. Therefore $\lim_{x \to c^-} f(x) = L$. Similarly $\lim_{x \to c^+} f(x) = \inf\{f(x) : c < x \leqslant b\} = M \geqslant f(c)$. ∎

4.5.2. DEFINITION. If $\lim_{x \to c^-} f(x) = L$ and $\lim_{x \to c^+} f(x) = M$ both exist, but L, M and $f(c)$ are not all equal, then f is said to have a *jump discontinuity*.

4.5.3. REMARK. At the endpoint a, a similar analysis shows that $\lim_{x \to a^+} f(x) = M$ exists, and $f(a) \leqslant M$. If $f(a) < M$, we call this a jump discontinuity as well. Similarly, $\lim_{x \to b^-} f(x) = L \leqslant f(b)$; and it is a jump discontinuity if $L < f(b)$.

The Proposition shows that the only type of discontinuity that a monotone function can have is a jump discontinuity. This leads to the following useful conclusion.

4.5.4. COROLLARY. *If I is an interval (closed, open or half open) and $f : I \to \mathbb{R}$ is a monotone function, then f is continuous if and only if* $\operatorname{Ran}(f)$ *is an interval.*

PROOF. If f is discontinuous at an interior point $c \in I$, then

$$\lim_{x \to c^-} f(x) = L < M = \lim_{x \to c^+} f(x).$$

Thus $\operatorname{Ran}(f) \subset [f(a), L] \cup \{f(c)\} \cup [M, f(b)]$. The range omits all except possibly one point of (L, M), and hence the range is not an interval. A similar analysis works at an endpoint.

Conversely, if $\operatorname{Ran}(f)$ is an interval, then we see that f has no jump discontinuities. Thus for $a < c < b$, we have $\lim_{x \to c^-} f(x) = f(c) = \lim_{x \to c^+} f(x)$. Therefore f is continuous at c. The argument works similarly at any endpoints. ∎

Recall that when a function f is one-to-one, it is a bijection from its domain onto its range. Thus it has an inverse function f^{-1}. We have seen that when f is continuous and one-to-one on an interval, then it is monotone and the range in an interval. We can now show that the inverse function is also continuous and monotone.

4.5.5. PROPOSITION. *If I is an interval and f is continuous and strictly increasing, then the inverse function f^{-1} is continuous and strictly increasing. Similarly, if f is continuous and strictly decreasing, then the inverse function f^{-1} is continuous and strictly decreasing.*

PROOF. We will deal with the increasing case. Since f is continuous, the range is an interval J. The inverse function $f^{-1} : J \to I$ is a bijection. If $y_1 < y_2 \in J$, then there are unique points $x_1, x_2 \in I$ with $f(x_i) = y_i$. Since f is increasing, $x_1 < x_2$. Thus $f^{-1}(y_1) = x_1 < x_2 = f^{-1}(y_2)$. Thus f^{-1} is strictly increasing. The range is an interval, and thus has no gaps. Therefore f^{-1} is continuous. ■

A *countable* set X is either finite or can be written as a list $X = \{x_n : n \in \mathbb{N}\}$. See the Appendix section A.3.

4.5.6. PROPOSITION. *A monotone function $f : (a, b) \to \mathbb{R}$ is continuous except on a countable set.*

PROOF. We may suppose that f is increasing. Since the only discontinuities are jump discontinuities, we count the jumps based on their size. Since the range of f may be unbounded, we also need to carefully approach the endpoints.

Between $a + \frac{1}{n}$ and $b - \frac{1}{n}$, count the jumps J_n of height at least $\frac{1}{n}$. There can't be more than $n\big(f(b-\frac{1}{n}) - f(a+\frac{1}{n})\big)$, which is finite. Therefore all discontinuities belong to $J = \bigcup_{n \geqslant 1} J_n$, which is countable. ■

4.5.7. EXAMPLE. *Inverse trig functions.* The trig functions are all periodic, either 2π-periodic like $\sin x$ and $\cos x$ or π-periodic like $\tan x$ and $\cot x$. So they are not one-to-one. We get around that by restricting the domain so that it is injective, maps onto the whole range, and the domain is as connected as possible and includes the first quadrant $(0, \frac{\pi}{2})$.

Take $\sin x$. It is monotone increasing on $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and maps onto $[-1, 1]$. Thus $\sin^{-1}(y)$ is the unique $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ with $\sin x = y$. For $\cos x$, we use $[0, \pi]$, on which $\cos x$ is monotone decreasing and maps onto $[-1, 1]$.

The tangent $\tan x$ is defined on $(-\frac{\pi}{2}, \frac{\pi}{2})$, is strictly increasing, and maps onto $\mathbb{R}$. So $\tan^{-1} : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ is a bounded function on the whole line. Similarly $\cot x$ maps $(0, \pi)$ onto $\mathbb{R}$ and is strictly decreasing.

The secant is a problem, because the range of $\sec x$ is $(-\infty, -1] \cup [1, \infty)$, which is the union of two disjoint intervals. Normal practice is to restrict the domain to $[0, \frac{\pi}{2}) \cup (\frac{\pi}{2}, \pi]$. On $[0, \frac{\pi}{2})$, secant is monotone increasing and maps onto $[1, \infty)$; while on $(\frac{\pi}{2}, \pi]$, secant is also increasing with range $(-\infty, -1]$. Hence $\sec^{-1}$ is increasing, and maps $[1, \infty)$ onto $[0, \frac{\pi}{2})$, and maps $(-\infty, -1]$ onto $(\frac{\pi}{2}, \pi]$. Similarly, $\csc^{-1}$ is strictly decreasing, and maps $[-\frac{\pi}{2}, 0)$ onto $(-\infty, -1]$ and $(0, \frac{\pi}{2}]$ onto $[1, \infty)$.

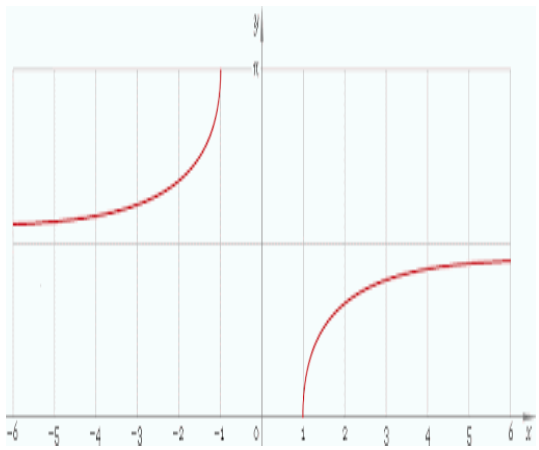


FIGURE 4.3. $\sec^{-1}(x)$

4.5.8. EXAMPLE. Enumerate $\mathbb{Q} \cap (0,1) = \{\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6}, \dots\}$ as $\{r_n : n \geqslant 1\}$. Define $f : [0,1] \to [0,1]$ by

$$f(x) = \sum_{\{n : r_n < x\}} 2^{-n}.$$

Then if $0 \leqslant x < y \leqslant 1$, there is some r_n so that $x < r_n < y$, and therefore $f(x) < f(y)$. So f is strictly monotone increasing. Clearly f has a jump discontinuity at every r_n. However it is continuous at each irrational number as well as 0 and 1. Say c is irrational in $(0,1)$ and $\varepsilon > 0$. Pick N so that $2^{-N} < \varepsilon$. Set $\delta = \operatorname{dist}(x, \{r_n : 1 \leqslant n \leqslant N\}$. If $c - \delta < y < x < z < c + \delta$, then $\{n : y \leqslant r_n < z\} \subset \{n : n > N\}$, and hence

$$f(z) - f(y) = \sum_{\{n : y \leqslant r_n < z\}} 2^{-n} < \sum_{n > N} 2^{-n} = 2^{-N} < \varepsilon.$$

Defining L and M as in Proposition 4.5.1, we see that $M - L < f(z) - f(y) < \varepsilon$. But $\varepsilon > 0$ was arbitrary, and thus $L = M$ and f is continuous at c.

4.5.9. EXAMPLE. The *Cantor function*. Define a function $f : [0,1] \to [0,1]$ as follows: set $f(0) = 0$ and $f(1) = 1$. On the middle third, $[\frac{1}{3}, \frac{2}{3}]$, set $f(x) = \frac{1}{2}$. Then take the middle third from each of the remainder, and set $f(x) = \frac{1}{4}$ on $[\frac{1}{9}, \frac{2}{9}]$ and $f(x) = \frac{3}{4}$ on $[\frac{7}{9}, \frac{8}{9}]$. At the nth stage, there are $2^n - 1$ intervals on which f is defined to take the values $\frac{k}{2^n}$, in order, for $1 \leqslant k \leqslant 2^n - 1$. What remains are 2^n intervals of length 3^{-n} on which f has not yet been defined. In each one, take the middle third and define f to take the average of the values at the two ends of the interval, which will be a number of the form $\frac{2k-1}{2^{n+1}}$ for $1 \leqslant k \leqslant 2^n$. In the end, we have defined a monotone increasing, locally constant function on the union S of all of these intervals, taking all of the values $\frac{k}{2^n}$ for $n \geqslant 1$ and $0 \leqslant k \leqslant 2^n$. These are the diadic rationals numbers in $[0, 1]$. The set S does not include the whole interval, so now we define f on the rest by

$$f(x) = \sup\{f(t) : t \leqslant x,\ t \in S\}.$$

This defines a monotone increasing function on $[0, 1]$ known as the Cantor function.

Notice that the range of f has no gaps because the range includes all diadic rationals,. Therefore f is continuous. On each open interval in S, namely $(\frac{1}{3}, \frac{2}{3})$, $(\frac{1}{9}, \frac{2}{9})$, $(\frac{7}{9}, \frac{8}{9})$, etc., the function f is constant, and thus has derivative 0. The total length of all of these intervals is

$$\frac{1}{3} + 2(\tfrac{1}{9}) + 4(\tfrac{1}{27}) + \cdots = \sum_{k=1}^{\infty} \frac{2^{k-1}}{3^n} = \frac{1/3}{1 - 2/3} = 1.$$

So this is "most" of the interval in some sense. What remains after these open intervals are removed is a closed set C known as the *Cantor set*. It includes the endpoints of the removed intervals, and their limits, which is actually quite a lot more.

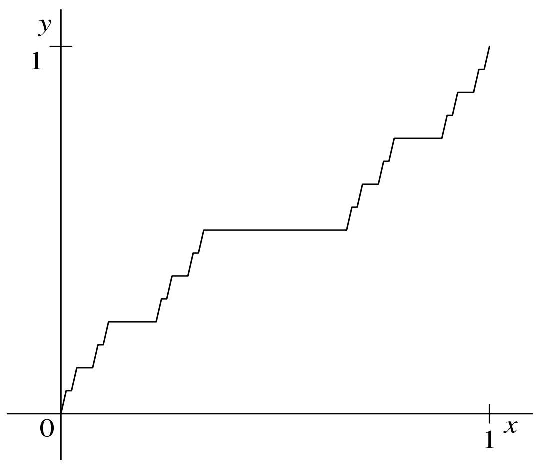


FIGURE 4.4. The Cantor function

To understand this function better, write x in base 3 as $x = (0.x_1x_2x_3\ldots)_{\text{base } 3}$, where $x_i \in \{0, 1, 2\}$. The interval $[\frac{1}{3}, \frac{2}{3}] = \{x : x_1 = 1\}$. The endpoints are $\frac{1}{3} = (0.1000\ldots)_{\text{base } 3}$ and $\frac{2}{3} = (0.1222\ldots)_{\text{base } 3}$. Like in base 10, numbers with a finite expansion in base 3 also have another ending in an infinite string of 2's. Then

$$[\tfrac{1}{9}, \tfrac{2}{9}] = \{x = (0.01x_3x_4\ldots)_{\text{base } 3}\} \quad \text{and} \quad [\tfrac{7}{9}, \tfrac{8}{9}] = \{x = (0.21x_3x_4\ldots)_{\text{base } 3}\}.$$

At the nth stage the intervals are determined by an initial sequence of 0s and/or 2's of length $n - 1$ followed by a 1. The endpoints actually have another expression using only 0s and 2s, $\frac{1}{3} = (0.0222\ldots)_{\text{base } 3}$, $\frac{2}{3} = (0.2000\ldots)_{\text{base } 3}$, $\frac{1}{9} = (0.00222\ldots)_{\text{base } 3}$, $\frac{2}{9} = (0.02000\ldots)_{\text{base } 3}$, $\frac{7}{9} = (0.20222\ldots)_{\text{base } 3}$, etc.

The Cantor set consists of all numbers in $[0, 1]$ which have a ternary expansion with no 1s. Let $(\varepsilon_1, \varepsilon_2, \dots)$ be a sequence of 0s and 1s. Then

$$f((0.(2\varepsilon_1)(2\varepsilon_2)(2\varepsilon_3)\dots)_{\text{base 3}} = (0, \varepsilon_1\varepsilon_2\varepsilon_3\dots)_{\text{base 2}}.$$

The Cantor set is mapped onto $[0, 1]$ by f. This shows that the Cantor set has the same cardinality as $[0, 1]$ and $\mathbb{R}$, (See Appendix A.3.)

4.6. Uniform Continuity

Recall that $f : J \to \mathbb{R}$ is continuous if for each $x \in J$ and $\varepsilon > 0$, there is a $\delta > 0$ so that $|x - y| < \delta$ implies $|f(y) - f(x)| < \varepsilon$. However in checking this in many examples, we find a δ which works for many, sometimes all, values of x simultaneously. This is an important distinction which is captured in this definition. Note that the quantifiers for x and δ come in the reversed order.

4.6.1. Definition. A function $f : J \to \mathbb{R}$ is *uniformly continuous* if for each $\varepsilon > 0$, there is a $\delta > 0$ so that $x, y \in J$ and $|x - y| < \delta$ implies $|f(y) - f(x)| < \varepsilon$.

4.6.2. Examples.

(1) Let $f \in C^1[a, b]$ and set $M = \max_{a \leqslant x \leqslant b} |f'(x)|$, which is finite by the Extreme Value Theorem. Then by the Mean Value Theorem, there is some x_0 so that $\dfrac{|f(y) - f(x)|}{|y - x|} = |f'(x_0)| \leqslant M$. Hence $|f(y) - f(x)| \leqslant M|y - x|$. So f is Lipschitz with constant M. As in Example 4.1.3(3), Lipschitz functions are uniformly continuous using $\delta = \varepsilon/M$. Thus C^1 functions on a closed bounded interval are uniformly continuous.

(2) Let $f(x) = \dfrac{1}{x}$ on $(0, 1]$. Then f is continuous. However if $0 < \varepsilon < 1$ is given and $\delta > 0$, choose $x = \min\{\varepsilon, \delta\}$. Then $|x - \frac{x}{10}| < |x| \leqslant \delta$, but

$$|f(\frac{x}{10}) - f(x)| = \frac{9}{x} > 9 > \varepsilon.$$

So this (arbitrary) choice of δ does not work! Therefore this function is not uniformly continuous. This is a C^1 function, but the derivative blows up as $x \to 0$.

(3) $f(x) = x^2$. On any bounded interval, the derivative is bounded. So by (1), f is uniformly continuous. However on the whole line $\mathbb{R}$, it is not uniformly continuous. For $0 < x < y$,

$$f(y) - f(x) = y^2 - x^2 = (y + x)(y - x).$$

Let $0 < \varepsilon < 1$, and suppose that $\delta > 0$. Take $x = \frac{1}{\delta}$ and $y = x + \delta/2$. Then $|y - x| = \delta/2 < \delta$, but

$$|f(y) - f(x)| = (y + x)(y - x) > 2x\frac{\delta}{2} = 1 > \varepsilon.$$

Hence this (arbitrary) choice of δ does not work! So this function is not uniformly continuous. In this example, the derivative $f'(x) = 2x$ blows up as $x \to \infty$.

(4) $f(x) = \sqrt{x}$ on $[0, \infty)$. This function is continuous, but it fails to be differentiable at $x = 0$. Moreover $f'(x) = \dfrac{1}{2\sqrt{x}}$ is unbounded as $x \to 0^+$. Nevertheless, we will show that f is uniformly continuous.

Let $\varepsilon > 0$. Define $\delta = \varepsilon^2/2$. Suppose that $0 \leqslant x \leqslant y$ and $|y - x| < \delta$. There are two cases. If $x \leqslant \delta$, then $y < 2\delta = \varepsilon^2$. Then $|\sqrt{y} - \sqrt{x}| \leqslant \sqrt{y} < \varepsilon$.

In the second case, $\delta < x$. Then

$$|\sqrt{y} - \sqrt{x}| = \frac{y - x}{\sqrt{y} + \sqrt{x}} < \frac{\delta}{2\sqrt{\delta}} = \frac{\sqrt{\delta}}{2} = \frac{\varepsilon}{2\sqrt{2}} < \varepsilon.$$

There is an important general result which implies uniform continuity.

4.6.3. THEOREM. *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on a closed, bounded interval. Then f is uniformly continuous.*

PROOF. If f is not uniformly continuous, then the definition fails for some $\varepsilon_0 > 0$. That means that no value of δ makes the definition work. So we take $\delta_n = \frac{1}{n}$. Since the definition fails, there are $x_n, y_n \in [a, b]$ so that $|x_n - y_n| < \frac{1}{n}$ and $|f(x_n) - f(y_n)| \geqslant \varepsilon_0$. By the Bolzano-Weierstrass Theorem, the sequence (x_n) has a convergent subsequence $(x_{n_k})_{k \geqslant 1}$, say $c = \lim_{k\to\infty} x_{n_k}$. Therefore

$$\lim_{k\to\infty} y_{n_k} = \lim_{k\to\infty} (y_{n_k} - x_{n_k}) + x_{n_k} = 0 + c = c.$$

By continuity of $f(x)$ at c, we have

$$f(c) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} f(y_{n_k}).$$

Hence

$$0 = \lim_{k\to\infty} |f(x_{n_k}) - f(y_{n_k})| \geqslant \varepsilon_0.$$

This is a contradiction. Thus f must be uniformly continuous. ∎

Exercises for Chapter 4

1. Let $f(x)$ and $g(x)$ be continuous functions on $[a, b]$. Define

$$f \wedge g(x) = \min\{f(x), g(x)\} \quad \text{and} \quad f \vee g(x) = \max\{f(x), g(x)\}.$$

Prove that $f \wedge g$ and $f \vee g$ are continuous on $[a, b]$.

2. (a) Let $f(x) = x^{\frac{1}{1-x}}$ for $x \geqslant 0$, $x \neq 1$. Can f be defined at $x = 1$ in order to make the function continuous there?
(b) Let $g(x) = \dfrac{e^{1/x}}{1 + e^{1/x}}$ for $x \neq 0$. Can g be defined at $x = 0$ in order to make the function continuous there?
(c) Let $f(x) = (\sin^2 x)^{\sec^2 x}$. Where is this function defined? Can f be defined at the missing points in order to make the function continuous there?
HINT: $x^a = e^{a \ln x}$. Look for a derivative in the exponent.

3. Fix a number $d > 0$. A function $f(x)$ on $\mathbb{R}$ is called *d–periodic* if $f(x + d) = f(x)$ for all $x \in \mathbb{R}$. Let f be a continuous d-periodic function on $\mathbb{R}$. Show that f attains its maximum and minimum values.

4. Suppose that $f : \mathbb{R} \to (0, \infty)$ is a continuous function such that $f(0) > 0$ and $\lim_{x \to +\infty} f(x) = 0 = \lim_{x \to -\infty} f(x)$. Prove that f attains its supremum.

5. Let $f : (a, b) \to \mathbb{R}$ be a monotone increasing function.
(a) Show that if f is bounded above, then $\lim_{x \to b^-} f(x)$ exists.
HINT: use MCT for a *sequence* approaching b^-.
(b) Hence show that for all $a < c < b$ that $\lim_{x \to c^-} f(x)$ and $\lim_{x \to c^+} f(x)$ both exist. When is $f(x)$ continuous at c?
(c) Construct a function which is monotone increasing and bounded on $\mathbb{R}$ and is discontinuous at every rational number.
HINT: list all of the rational numbers as $\mathbb{Q} = \{r_1, r_2, r_3, \dots\}$ and introduce a small jump discontinuity at each r_k.

6. Show that $f(x) = |x|^{1/2}$ is continuous but not Lipschitz.

7. Suppose that there is a constant C such that

$$|f(x) - f(y)| \leqslant C|x - y|^2 \quad \text{for all} \quad a \leqslant x, y \leqslant b.$$

Prove that f is constant.

8. Show that $\tan x - 4 \sin x = e^x$ has at least 3 solutions in $(-\frac{\pi}{2}, \frac{\pi}{2})$.

9. Show that there are two *antipodal* points on the equator with exactly the same temperature. HINT: parametrize points on the equator by $[0, 2\pi]$ using the angle θ from the centre of the earth. Assume that the temperature function

$T(\theta)$ is continuous. (Is this reasonable?) You are asked to prove the existence of some θ so that $T(\theta) = T(\theta + \pi)$.

10. Suppose that $f : \mathbb{R} \to (0, \infty)$ is a positive continuous function such that $\lim_{x\to+\infty} f(x) = 0 = \lim_{x\to-\infty} f(x)$. Prove that f attains its supremum.

11. (a) Show that a continuous function on $(-\infty, +\infty)$ cannot take *every* real value *exactly twice*.

(b) Find a continuous function on $(-\infty, +\infty)$ which takes *every* real value *exactly three times*. A sketch of the curve will suffice. An exact formula is not required.

12. Let

$$f(x) = \begin{cases} \dfrac{1}{1 + (\ln x)^2} & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0. \end{cases}$$

Prove that $f(x)$ is uniformly continuous on $[0, \infty)$.
HINT: first prove it separately on $[0, 3]$ and on $[2, \infty)$.

CHAPTER 5

Differentiation

5.1. The derivative

As you have seen in high school calculus, the derivative of a function $f(x)$ at x_0, if it exists, is the slope of the tangent line to the curve $y = f(x)$ at a point $(x_0, f(x_0))$. We compute this by computing the slope of a *secant*, the line segment from $(x_0, f(x_0))$ to $(x_0 + h, f(x_0 + h))$, and taking the limit as $h \to 0$ through both positive and negative values.

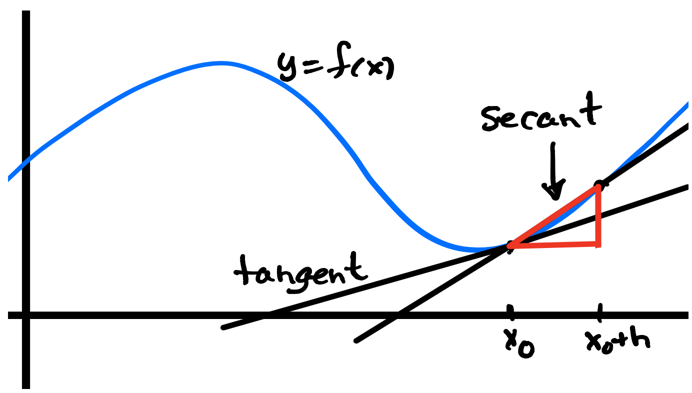


FIGURE 5.1. Tangent line

5.1.1. DEFINITION. Let $f : (a, b) \to \mathbb{R}$ and let $x_0 \in (a, b)$. Then f is *differentiable at* x_0 if

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists and is finite. The limit is called $f'(x_0)$ or $\left(\frac{d}{dx} f\right)(x_0)$.

We say $f(x)$ is *differentiable* on (a, b) if it is differentiable at every point $x_0 \in (a, b)$. When f is defined on a closed interval $[a, b]$, we will say that f is differentiable on $[a, b]$ if it is differentiable on (a, b) and the one-sided derivatives

$$\lim_{h \to 0^+} \frac{f(a + h) - f(a)}{h} \quad \text{and} \quad \lim_{h \to 0^-} \frac{f(b + h) - f(b)}{h}$$

both exists and are finite. The *tangent line* to $f(x)$ at x_0 is

$$T(x) = f(x_0) + f'(x_0)(x - x_0).$$

This is the line through $(x_0, f(x_0))$ with slope $f'(x_0)$.

5.1.2. THEOREM. *If $f(x)$ is differentiable at x_0, then f is continuous at x_0.*

PROOF. This is an easy computation

$$\lim_{x\to x_0} f(x) - f(x_0) = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) = f'(x_0) \cdot 0 = 0. \qquad \blacksquare$$

Now we provide two useful variants which are equivalent to differentiability.

5.1.3. THEOREM. *Let $f : (a, b) \to \mathbb{R}$ and let $x_0 \in (a, b)$. The following are equivalent:*

(1) *f is differentiable at x_0 and $f'(x_0) = m$.*

(2) *there is a linear function $T(x) = f(x_0) + m(x - x_0)$ such that*

$$\lim_{x\to x_0} \frac{f(x) - T(x)}{x - x_0} = 0.$$

(3) *there is a function $\varphi(x)$ which is continuous at x_0 such that $\varphi(x_0) = m$ and $f(x) = f(x_0) + \varphi(x)(x - x_0)$.*

PROOF. (1)⇒(3). Define $\varphi(x) = \dfrac{f(x) - f(x_0)}{x - x_0}$ if $x \neq x_0$ and $\varphi(x_0) = m$. Then $f(x) = f(x_0) + \varphi(x)(x - x_0)$ and

$$\lim_{x\to x_0} \varphi(x) = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) = m = \varphi(x_0).$$

Therefore φ is continuous at x_0, and (3) holds.

(3)⇒(2). Let $T(x) = f(x_0) + m(x - x_0)$. Then

$$\begin{aligned}\lim_{x\to x_0} \frac{f(x) - T(x)}{x - x_0} &= \lim_{x\to x_0} \frac{f(x_0) + \varphi(x)(x - x_0) - f(x_0) - m(x - x_0)}{x - x_0} \\ &= \lim_{x\to x_0} \varphi(x) - m = 0.\end{aligned}$$

(2)⇒(1). If (2) holds, then

$$0 = \lim_{x\to x_0} \frac{f(x) - T(x)}{x - x_0} = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} - m.$$

Therefore $f'(x_0) = \lim_{x\to x_0} \dfrac{f(x) - f(x_0)}{x - x_0} = m$ exists. $\blacksquare$

5.1.4. EXAMPLES.

(1) Let $f(x) = \sin x$. we have shown that $f'(x) = \cos x$. Our proof used the addition formula:

$$\begin{aligned}\sin(x_0 + h) &= \sin x_0 \cos h + \cos x_0 \sin h \\ &= \sin x_0 + \cos x_0\, h + \sin x_0(\cos h - 1) + \cos x_0(h - \sin h).\end{aligned}$$

The tangent line is $T(x_0 + h) = \sin x_0 + \cos x_0\, h$. Define

$$\varphi(x_0 + h) = \cos x_0 + \sin x_0 \frac{\cos h - 1}{h} + \cos x_0 \frac{h - \sin h}{h}.$$

Then $\sin(x_0 + h) = \sin x_0 + \varphi(x_0 + h)h$ and

$$\lim_{h\to 0} \varphi(x_0 + h) = \cos x_0 + \sin x_0 \lim_{h\to 0} \frac{\cos h - 1}{h} + \cos x_0 \lim_{h\to 0} \frac{h - \sin h}{h} = \cos x_0.$$

(2) Let $f(x) = x^n$. Then using the binomial theorem,

$$\begin{aligned}f'(a) &= \lim_{h\to 0} \frac{(a+h)^n - a^n}{h} = \lim_{h\to 0} \frac{\sum_{k=0}^n \binom{n}{k} a^{n-k} h^k - a^n}{h} \\ &= \lim_{h\to 0} \frac{a^n - a^n}{h} + \binom{n}{1} a^{n-1} \frac{h}{h} + \sum_{k=2}^n \binom{n}{k} a^{n-k} h^{k-1} \\ &= na^{n-1} + \lim_{h\to 0} h \sum_{k=2}^n \binom{n}{k} a^{n-k} h^{k-2} = na^{n-1}.\end{aligned}$$

So $f'(x) = nx^{n-1}$.

(3) $f(x) = a^x = e^{x \ln a}$. Then

$$\begin{aligned}f'(x_0) &= \lim_{h\to 0} \frac{a^{x_0+h} - a^{x_0}}{h} = a^{x_0} \lim_{h\to 0} \frac{e^{h \ln a} - 1}{h \ln a} \ln a \\ &= a^{x_0} \ln a \lim_{u\to 0} \frac{e^u - 1}{u} = a^{x_0} \ln a.\end{aligned}$$

5.1.5. EXAMPLE. Let $f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$. For $x \neq 0$, we have

$$f'(x) = 2x \sin \tfrac{1}{x} + x^2 \left(\tfrac{-1}{x^2}\right) \cos \tfrac{1}{x} = 2x \sin \tfrac{1}{x} - \cos \tfrac{1}{x}.$$

However at $x = 0$, we have

$$\lim_{h\to 0} \frac{f(h) - f(0)}{h} = \lim_{h\to 0} h \sin \tfrac{1}{h} = 0$$

by the Squeeze Theorem (since $-|h| \leqslant h \sin \frac{1}{h} \leqslant |h|$). Thus $f'(0) = 0$. Notice however that $f'(x)$ is discontinuous at 0:

$$\lim_{x\to 0} f'(x) = \lim_{x\to 0} 2x \sin \tfrac{1}{x} - \cos \tfrac{1}{x}$$

does not exist!

To graph $f(x)$, we notice that as $x \to 0$, the curve oscillates rapidly between $y = x^2$ and $y = -x^2$. The function is odd. It has zeroes at $\frac{1}{n\pi}$ for $n \in \mathbb{Z}\backslash\{0\}$. It touches the bounding curves in between, and at $\frac{\pm 2}{\pi}$. As $x \to \pm\infty$, we can approximate $x^2 \sin \frac{1}{x} \approx x^2 \frac{1}{x} = x$. Thus we compute (by substituting $u = \frac{1}{x}$)

$$\lim_{x \to \infty} x^2 \sin \frac{1}{x} - x = \lim_{u \to 0^+} \frac{1}{u}\Big(\frac{\sin u}{u} - 1\Big)$$

We showed that for $0 < u < \frac{\pi}{2}$, that $\cos u < \dfrac{\sin u}{u} < 1$, and hence

$$\frac{\cos u - 1}{u} < \frac{1}{u}\Big(\frac{\sin u}{u} - 1\Big) < 0.$$

As $\lim_{u \to 0} \dfrac{\cos u - 1}{u} = 0$, the Squeeze Theorem shows that $\lim\limits_{x \to \infty} x^2 \sin \frac{1}{x} - x = 0$. Thus $y = x$ is an oblique asymptote. A similar analysis shows that as $x \to -\infty$, it is also an asymptote.

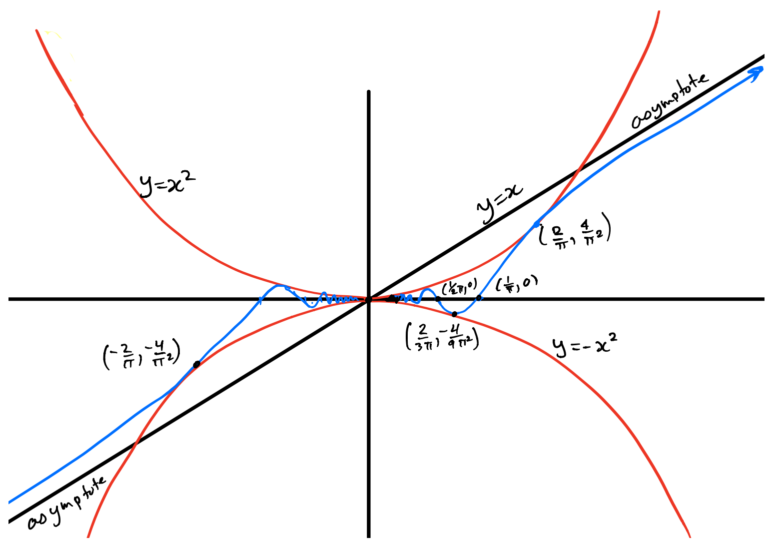


FIGURE 5.2. Graph of $x^2 \sin \frac{1}{x}$

5.2. Differentiation Rules

We first derive the rules for derivatives of sums, products and quotients of functions. The addition rule is routine. The proofs of the product rule and the quotient rule are similar, but the latter is a bit more complicated. We will prove the quotient iule and leave the others as an exercise.

5.2.1. Proposition. *Let f, g be differentiable at x, and let $r, s \in \mathbb{R}$. Then*

(1) $(rf + sg)'(x) = rf'(x) + sg'(x)$.

(2) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.

(3) *If* $g(x) \neq 0$, $\big(f/g\big)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$.

Proof.

$$\lim_{h\to 0} \frac{\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}}{h} = \lim_{h\to 0} \frac{f(x+h)g(x) - f(x)g(x+h)}{hg(x+h)g(x)}$$
$$= \lim_{h\to 0} \frac{1}{g(x+h)g(x)} \Big(\frac{f(x+h)-f(x)}{h} g(x) + f(x) \frac{g(x)-g(x+h)}{h} \Big)$$
$$= \frac{1}{g(x)^2} \big(f'(x)g(x) - f(x)g'(x) \big).$$

We used the fact that differentiability of g implies continuity of g at x. ■

The chain rule for the derivative of a composition will be proven using Theorem 5.1.3.

5.2.2. Chain Rule. *Let $f : (a, b) \to (c, d)$ and $g : (c, d) \to \mathbb{R}$. Suppose that f is differentiable at x_0 and that g is differentiable at $f(x_0)$. Then $g \circ f$ is differentiable at x_0 and*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. Since f is differentiable at x_0, applying Theorem 5.1.3, we can write $f(x) = f(x_0) + \varphi(x)(x - x_0)$ where $\varphi(x_0) = f'(x_0) = \lim\limits_{x\to x_0} \varphi(x)$. Similarly, set $y_0 = f(x_0)$ and write $g(y) = g(y_0) + \psi(y)(y - y_0)$ where $\psi(y_0) = g'(y_0) = \lim\limits_{y\to y_0} \psi(y)$. Set $h(x) = (g \circ f)(x) = g(f(x))$. Then

$$h(x) = h(x_0) + \psi(f(x))(f(x) - f(x_0)) = h(x_0) + \psi(f(x))\varphi(x)(x - x_0).$$

Let $\chi(x) = \psi(f(x))\varphi(x)$. Then $\chi(x_0) = \psi(y_0)\varphi(x_0) = g'(f(x_0))f'(x_0)$. Since f is continuous at x_0, and φ and ψ are continuous at x_0 and $f(x_0)$ respectively,

$$\lim_{x\to x_0} \chi(x) = \lim_{x\to x_0} \psi(f(x))\varphi(x) = \psi(f(x_0))\varphi(x_0) = \chi(x_0).$$

By Theorem 5.1.3, h is differentiable at x_0 and $h'(x_0) = g'(f(x_0))f'(x_0)$. ■

Suppose that $f(x)$ has an inverse function $f^{-1}(x)$. If we knew that f and f^{-1} were differentiable, we could apply the chain rule to $y = f \circ f^{-1}(y)$ to get $1 = f'(f^{-1}(y))(f^{-1})'(y)$. Solving yields

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

For this to make sense, we also need $f'(f^{-1}(y)) \neq 0$. To avoid assuming differentiability, we use Theorem 5.1.3 again.

5.2.3. THEOREM (Inverse functions). *Suppose that f is strictly monotone and maps (a,b) onto (c,d). Let $f^{-1} : (c,d) \to (a,b)$ be the inverse function. Let $y_0 \in (c,d)$ and $x_0 = f^{-1}(y_0)$. If f is differentiable at x_0 and $f'(x_0) \neq 0$, then f^{-1} is differentiable at y_0 and*

$$(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

PROOF. Since f is differentiable at x_0, applying Theorem 5.1.3, we can write $f(x) = f(x_0) + \varphi(x)(x - x_0)$ where $\varphi(x_0) = f'(x_0) = \lim_{x\to x_0} \varphi(x)$. For $y \in (c,d)$, let $x = f^{-1}(y)$. Therefore

$$\begin{aligned} y = f(x) &= f(x_0) + \varphi(x)(x - x_0) \\ &= y_0 + \varphi(f^{-1}(y))(f^{-1}(y) - f^{-1}(y_0)). \end{aligned}$$

Note that if $y \neq y_0$, then $\varphi(f^{-1}(y)) \neq 0$. Solve for $f^{-1}(y)$:

$$f^{-1}(y) = f^{-1}(y_0) + \frac{y - y_0}{\varphi(f^{-1}(y))} = f^{-1}(y_0) + \psi(y)(y - y_0),$$

where $\psi(y) = \dfrac{1}{\varphi(f^{-1}(y))}$ for $y \neq y_0$. Set $\psi(y_0) = \dfrac{1}{f'(x_0)}$. Then

$$\lim_{y\to y_0} \psi(y) = \lim_{y\to y_0} \frac{1}{\varphi(f^{-1}(y))} = \lim_{x\to x_0} \frac{1}{\varphi(x)} = \frac{1}{\varphi(x_0)} = \frac{1}{f'(x_0)} = \psi(y_0).$$

Thus ψ is continuous at y_0, so by Theorem 5.1.3, f^{-1} is differentiable at x_0 and

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$ ■

5.2.4. EXAMPLE. *Trig functions.* We have seen that $\frac{d}{dx}\sin x = \cos x$ and $\frac{d}{dx}\cos x = -\sin x$.

$$\begin{aligned}
\frac{d}{dx}\tan x &= \left(\frac{\sin x}{\cos x}\right)' = \frac{(\cos x)(\cos x) - (\sin x)(-\sin x)}{\cos^2 x} = \sec^2 x. \\
\frac{d}{dx}\cot x &= \left(\frac{\cos x}{\sin x}\right)' = \frac{-\sin^2 x - \cos^2 x}{\sin^2 x} = -\csc^2 x. \\
\frac{d}{dx}\sec x &= \left(\frac{1}{\cos x}\right)' = -\frac{-\sin x}{\cos^2 x} = \frac{\sin x}{\cos x}\frac{1}{\cos x} = \tan x \sec x. \\
\frac{d}{dx}\csc x &= \left(\frac{1}{\sin x}\right)' = -\frac{\cos x}{\sin^2 x} = -\cot x \csc x.
\end{aligned}$$

5.2.5. Example. *Inverse trig functions.* We restrict $\tan x$ to $(-\frac{\pi}{2}, \frac{\pi}{2})$, which is an increasing function with range $\mathbb{R}$. Hence $\tan^{-1} : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$. Moreover, $\lim_{x\to\infty} \tan^{-1}(x) = \frac{\pi}{2}$ shows that there is a horizontal asymptote, and similarly $\lim_{x\to-\infty} \tan^{-1}(x) = -\frac{\pi}{2}$. We compute

$$(\tan^{-1})'(x) = \frac{1}{\tan'(\tan^{-1}(x))} = \frac{1}{\sec^2(\tan^{-1}(x))} = \cos^2(\tan^{-1}(x)).$$

Take a right triangle with angle $\theta = \tan^{-1}(x)$ and compute $\cos\theta = \frac{1}{\sqrt{1+x^2}}$. There-

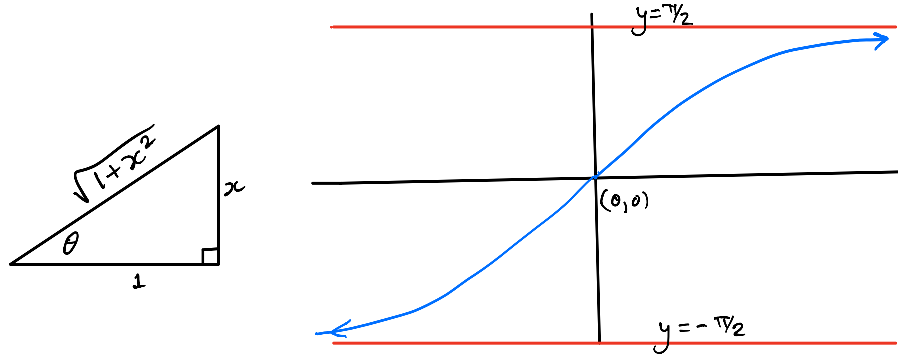


Figure 5.3. Graph of $\tan^{-1}(x)$

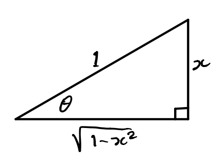
fore $\frac{d}{dx}\tan^{-1}(x) = \dfrac{1}{1+x^2}$.

Next we restrict $\cot x$ to $(0, \pi)$. It has range $\mathbb{R}$, is decreasing, and vertical asymptotes $x = 0$ and $x = \pi$. Thus $\cot^{-1} : \mathbb{R} \to (0, \pi)$ is decreasing, and has horizontal asymptotes $y = 0$ and $y = \pi$. The formula $\cot x = \tan(\frac{\pi}{2} - x)$ implies that $\cot^{-1} x = \frac{\pi}{2} - \tan^{-1} x$. Thus

$$\frac{d}{dx}\cot^{-1}(x) = -\frac{d}{dx}\tan^{-1} x = \frac{-1}{1+x^2}.$$

For $\sin^{-1}$, we restrict $\sin x$ to $[-\frac{\pi}{2}, \frac{\pi}{2}]$ on which $\sin x$ is increasing with range $[-1, 1]$. Thus the inverse $\sin^{-1} : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ is increasing. The derivative is

$$\frac{d}{dx}\sin^{-1}(x) = \frac{1}{\cos(\sin^{-1} x)} = \frac{1}{\sqrt{1-x^2}}.$$



Again we compute the cosine by drawing a right triangle with angle $\theta = \sin^{-1} x$. Notice that the derivative is undefined at $x = \pm 1$. This corresponds to the points $\pm\frac{\pi}{2}$ at which $\sin' x = \cos x = 0$. The function $\sin^{-1} x$ has vertical tangents at $(\pm 1, \pm\frac{\pi}{2})$ corresponding to the horizontal tangents of $\sin x$ at $(\pm\frac{\pi}{2}, \pm 1)$.

Now $\cos x$ maps $[0,\pi]$ onto $[-1,1]$ and is decreasing. We will use the relation $\cos x = \sin(\frac{\pi}{2} - x)$. Therefore $\cos^{-1} x = \frac{\pi}{2} - \sin^{-1} x$. Differentiating shows that $\dfrac{d}{dx}\cos^{-1} x = \dfrac{-1}{\sqrt{1-x^2}}$.

For $\sec^{-1} x$, recall from Example 4.5.7 that $\sec^{-1}$ has two *branches* defined on $(-\infty,-1]$ and $[1,\infty)$, respectively. We can compute the derivative by noting that $\sec^{-1} x = \cos^{-1}\frac{1}{x}$. Therefore by the chain rule,

$$\frac{d}{dx}\sec^{-1} x = \frac{d}{dx}\cos^{-1}(\tfrac{1}{x}) = \frac{-1}{\sqrt{1-\frac{1}{x^2}}}\,\frac{-1}{x^2} = \frac{1}{|x|\sqrt{x^2-1}}.$$

There is a subtlety here. Since $x^2 > 0$, we factor it as $|x|\sqrt{x^2}$ in order to clear the denominator in $\sqrt{1-\frac{1}{x^2}}$. This keeps the derivative positive, instead of changing the sign if we forget the absolute value. Similarly, $\csc^{-1} x = \sin^{-1}\frac{1}{x}$ and $\dfrac{d}{dx}\csc^{-1} x = \dfrac{-1}{|x|\sqrt{x^2-1}}$.

5.2.6. Examples.

(1) Let $a \neq 0$ and let $f(x) = x^a$ for $x > 0$, which is a non-integer power of x. We write $f(x) = e^{a\ln x}$. Thus

$$f'(x) = e^{a\ln x}\frac{a}{x} = \frac{ax^a}{x} = ax^{a-1}.$$

(2) Let $f(x) = \ln|x|$ for $x \neq 0$. When $x > 0$, $f(x) = \ln x$ and so $f'(x) = \frac{1}{x}$. When $x < 0$, $f(x) = \ln(-x)$, and so by the chain rule, $f'(x) = \frac{1}{-x}(-1) = \frac{1}{x}$. Therefore $f'(x) = \frac{1}{x}$.

(3) Implicit differentiation. Consider the curve

$$x^3 + 2x^2y + xy^2 + 3y^3 - 2x^2 - xy - y^2 + x + 7y - 9 = 0.$$

By inspection, $(0,1)$ lies on the curve. However there is no easy way to solve for y as a function of x. We differentiate the whole expression:

$$3x^2 + 2xy + x^2y' + y^2 + 2xyy' + 9y^2y' - 4x - y - xy' - 2yy' + 1 + 7y' = 0.$$

Solve for y':

$$y' = \frac{-(3x^2 + 2xy + y^2 - 4x + 1)}{x^2 + 2xy + 9y^2 - x - 2y + 7}.$$

Plugging in $x = 0$ and $y = 1$ yields $y' = -1/7$ at $(0,1)$.

(4) Logarithmic differentiation. Consider $f(x) = (\tan^2 x)^{\cos x} e^{x^2}(x+1)^4$. Then since $f(x) \geqslant 0$

$$\ln f(x) = 2\cos x \ln|\tan x| + x^2 + 2\ln|x+1|$$

which is valid unless $f(x) = 0$. Differentiate both sides:

$$\frac{f'(x)}{f(x)} = -2\sin x \ln|\tan x| + 2\frac{\cos x}{\tan x}\sec^2 x + 2x + \frac{2}{|x+1|}.$$

Solve for $f'(x)$. At points where $f(x) = 0$, a separate argument is needed.

5.3. Maxima and Minima

Recognizing the maximum and minimum points on a graph, even locally, has many applications.

5.3.1. DEFINITION. If $f : [a,b] \to \mathbb{R}$ is a function, a point x_0 is a *maximum* for f if $f(x_0) = \sup\{f(x) : a \leqslant x \leqslant b\}$. A point x_0 is a *minimum* for f if $f(x_0) = \inf\{f(x) : a \leqslant x \leqslant b\}$.

We say that x_0 is a *local maximum* for f if there is $\delta > 0$ so that it is a maximum on the smaller interval $[x_0 - \delta, x_0 + \delta]$. A point x_0 is a *local minimum* for f if there is $\delta > 0$ so that it is a minimum on $[x_0 - \delta, x_0 + \delta]$.

5.3.2. FERMAT'S THEOREM. *Suppose that $f : [a,b] \to \mathbb{R}$ is a continuous function which attains its maximum or minimum value at x_0. Then either*

(1) $x_0 \in \{a,b\}$ *is an endpoint of* $[a,b]$*,*

(2) $f'(x_0)$ *is undefined,*

or

(3) $f'(x_0) = 0$.

PROOF. Suppose that x_0 is a maximum for f and (1) is false, so $a < x_0 < b$; and that (2) is also false, so that $f'(x_0)$ is defined. Then

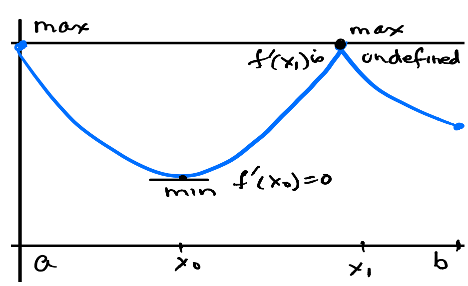


FIGURE 5.4. Fermat's Theorem

$$f'(x_0) = \lim_{h \to 0^+} \frac{f(x_0+h) - f(x_0)}{h} \leqslant 0$$
$$= \lim_{h \to 0^-} \frac{f(x_0+h) - f(x_0)}{h} \geqslant 0.$$

The first inequality is because the numerator is negative and the denominator is positive, while the second inequality is because the numerator is negative and the denominator is negative. Therefore $f'(x_0) = 0$. Minima are treated similarly. ■

We get the following important corollary.

5.3.3. ROLLE'S THEOREM. *Suppose that $f : [a,b] \to \mathbb{R}$ is continuous, is differentiable on (a,b) and $f(a) = f(b)$. Then there is a point $x_0 \in (a,b)$ so that $f'(x_0) = 0$.*

PROOF. If f is constant, any point $x_0 \in (a,b)$ will do. Otherwise there is some x with $f(x) \neq f(a)$. without loss of generality, $f(x) > f(a)$. By the Extreme Value Theorem, f attains its maximum value at some point x_0; and clearly $x_0 \notin \{a,b\}$. By Fermat's Theorem, since f is differentiable at x_0, $f'(x_0) = 0$. ■

5.3.4. EXAMPLE. Snell's Law. A beam of light travels from point A to point B. It starts in one medium in which the speed of light is c_1, but when it crosses the line $\overline{CD}$ into the second medium, the speed is c_2. (Typical media are air, water, glass, vacuum, etc.) The light will travel in a straight line from A to some point X on $\overline{CD}$, but then will change the angle (*refraction*) and follow a straight line to B. The point X is determined by *Fermat's Principle: the light will travel along the path requiring the least time.*

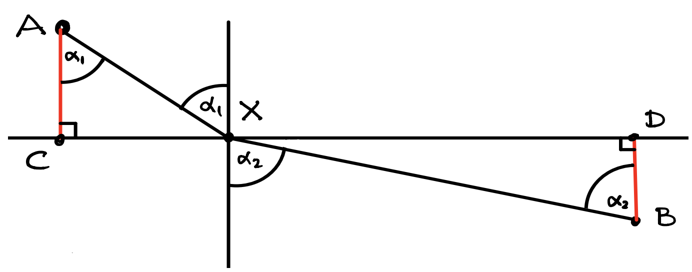


FIGURE 5.5. Snell's law

Drop perpendicular lines from A to a point C on the line, and from B to a point D. Let the distances be $\overline{CD} = L$, $\overline{AC} = h_1$ and $\overline{BD} = h_2$. We will treat X as a variable position along the line, and compute the time it would take the light to travel through X. Then we will minimize the time to find a relationship between the angle of incidence, $\alpha_1 = \angle CAX$, and the angle of refraction, $\alpha_2 = \angle DXB$.

Once we fix the angle α_1, the point X and the angle α_2 are determined as functions of α_1.

The distances travelled in each medium are

$$\overline{AX} = h_1 \sec\alpha_1 \quad \text{and} \quad \overline{XB} = h_2 \sec\alpha_2.$$

Thus the time taken is

$$T(\alpha_1) = \frac{h_1}{c_1}\sec\alpha_1 + \frac{h_2}{c_2}\sec\alpha_2.$$

We also need to work L into this by observing that

$$L = \overline{CX} + \overline{XD} = h_1 \tan\alpha_1 + h_2 \tan\alpha_2.$$

This works for $-\frac{\pi}{2} < \alpha_1 < \frac{\pi}{2}$, so even if the light should travel away from B or beyond B at first (actually impossible), the formula is still valid. We will differentiate it with respect to α_1:

$$0 = \frac{dL}{d\alpha_1} = h_1 \sec^2\alpha_1 + h_2 \sec^2\alpha_2 \frac{d\alpha_2}{d\alpha_1}.$$

Therefore $\dfrac{d\alpha_2}{d\alpha_1} = \dfrac{-h_1\sec^2\alpha_1}{h_2\sec^2\alpha_2}$. Observe that $\lim\limits_{\alpha_1\to\pm\frac{\pi}{2}} T(\alpha_1) = \infty$, and thus T will attain its minimum value in between. Differentiate T:

$$\begin{aligned}\frac{dT}{d\alpha_1} &= \frac{h_1}{c_1}\sec\alpha_1\tan\alpha_1 + \frac{h_2}{c_2}\sec\alpha_2\tan\alpha_2\frac{d\alpha_2}{d\alpha_1}\\ &= \frac{h_1}{c_1}\sec\alpha_1\tan\alpha_1 - \frac{h_2}{c_2}\sec\alpha_2\tan\alpha_2\frac{h_1\sec^2\alpha_1}{h_2\sec^2\alpha_2}\\ &= h_1\sec^2\alpha_1\Big(\frac{\sin\alpha_1}{c_1} - \frac{\sin\alpha_2}{c_2}\Big).\end{aligned}$$

Therefore $T'(\alpha_1) = 0$ if and only if $\dfrac{\sin\alpha_1}{c_1} = \dfrac{\sin\alpha_2}{c_2}$; and this must be the minimum. The relationship between the two angles is known as Snell's law.

5.4. Mean Value Theorem

There is a routine but powerful extension of Rolle's Theorem.

5.4.1. MEAN VALUE THEOREM. *Suppose that $f : [a,b] \to \mathbb{R}$ is continuous, and is differentiable on (a,b). Then there is a point $x_0 \in (a,b)$ so that*

$$f'(x_0) = \frac{f(b)-f(a)}{b-a}.$$

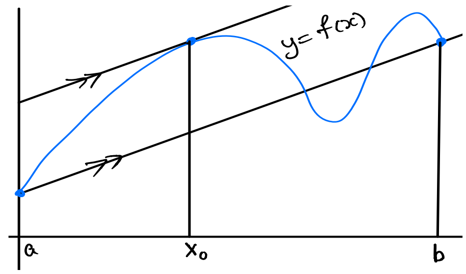


FIGURE 5.6. Mean Value Theorem

PROOF. Let $g(x) = f(x) - \Big(\dfrac{f(b)-f(a)}{b-a}\Big)x$. Then

$$\begin{aligned} g(b) - g(a) &= f(b) - \Big(\frac{f(b)-f(a)}{b-a}\Big)b - f(a) + \Big(\frac{f(b)-f(a)}{b-a}\Big)a \\ &= f(b) - f(a) - \Big(\frac{f(b)-f(a)}{b-a}\Big)(b-a) = 0. \end{aligned}$$

So $g(a) = g(b)$. By Rolle's Theorem, thee is a point $x_0 \in (a,b)$ so that

$$0 = g'(x_0) = f'(x_0) - \frac{f(b)-f(a)}{b-a}.$$

Thus $f'(x_0) - \dfrac{f(b)-f(a)}{b-a}$. ■

The following immediate consequence is used all the time.

5.4.2. COROLLARY. *Suppose that $f : [a,b] \to \mathbb{R}$ is continuous, and is differentiable on (a,b).*

- *If $f'(x) > 0$ on (a,b), then f is strictly increasing.*
- *If $f'(x) \geqslant 0$ on (a,b), then f is increasing.*
- *If $f'(x) < 0$ on (a,b), then f is strictly decreasing.*
- *If $f'(x) \leqslant 0$ on (a,b), then f is decreasing.*

PROOF. We only prove the first statement. Suppose that $a \leqslant x < y \leqslant b$. Then by the Mean Value Theorem applied to $[x,y]$, there is a point $x_0 \in (x,y)$ so that

$$\frac{f(y)-f(x)}{y-x} = f'(x_0) > 0.$$

Hence $f(x) < f(y)$. ■

5.4.3. PROPOSITION. *If $f : [a,b]$ has a continuous derivative $f'(x)$ and $f'(x_0) > 0$, then there is a $\delta > 0$ so that f is strictly increasing on $(x_0 - \delta, x_0 + \delta)$.*

PROOF. Take $\varepsilon = f'(x_0)$. By continuity of f', there is a $\delta > 0$ so that if $x \in (x_0 - \delta, x_0 + \delta)$, then $|f'(x) - f'(x_0)| < \varepsilon$. Hence

$$f(x) \geqslant f(x_0) - |(f(x) - f(x_0)| > \varepsilon - \varepsilon = 0.$$

Thus f is strictly increasing on $(x_0 - \delta, x_0 + \delta)$. ∎

5.4.4. EXAMPLE. This proposition can fail if f' is not continuous. Let

$$f(x) = ax + x^2 \sin \frac{1}{x}$$

for $0 < a < 1$. By Example 5.1.4(4), this function is differentiable everywhere, and $f'(x)$ is continuous except at $x = 0$. Moreover $f'(0) = a > 0$. However

$$f'(x) = a + 2x \sin \frac{1}{x} + x^2 \cos \frac{1}{x} \Big(\frac{-1}{x^2} \Big) = a + 2x \sin \frac{1}{x} - \cos \frac{1}{x}.$$

For $x_n = \frac{1}{2n\pi}$, $f'(x_n) = a - 1 < 0$. By Proposition 5.4.3, f is decreasing on a small interval around x_n. Since $x_n \to 0$, f is not increasing on any interval containing 0.

If $a > 1$, then for any $|x| < (a-1)/2 = \varepsilon$, we have $f'(x) \geqslant a - 2\varepsilon - 1 = 0$. Thus f is strictly increasing on $\big(\frac{1-a}{2}, \frac{a-1}{2}\big)$.

When $a = 1$, we see that $f'(x_n) = 0$. However

$$f''(x) = 2 \sin \tfrac{1}{x} + 2x \cos \tfrac{1}{x} \big(\tfrac{-1}{x^2}\big) + \sin \tfrac{1}{x} \big(\tfrac{-1}{x^2}\big) = \big(2 - \tfrac{1}{x^2}\big) \sin \tfrac{1}{x} - \tfrac{2}{x} \cos \tfrac{1}{x}.$$

This is continuous except at $x = 0$ and $f''(x_n) = \frac{-2}{x_n} < 0$. By Proposition 5.4.3, $f'(x)$ is strictly decreasing on a small interval $(x_n - \delta, x_n + \delta)$ around x_n. Hence $f'(x) < 0$ on $(x_n, x_n + \delta)$. Again f is not increasing on any interval containing 0.

Now we demonstrate another application of the Mean Value Theorem.

5.4.5. EXAMPLE. Let $f(x) = \sin x$. For $0 < x \leqslant \frac{\pi}{2}$, apply MVT on $[0, x]$. There is a point $x_0 \in (0, x)$ so that

$$\frac{\sin x}{x} = \frac{\sin x - \sin 0}{x - 0} = f'(x_0) = \cos x_0 \in (0, 1).$$

Hence $0 < \sin x < x$, an inequality established in Example 3.1.7.

Now let $g(x) = \frac{1}{2}x^2 + \cos x$. Then $g'(x) = x - \sin x > 0$ on $(0, \frac{\pi}{2})$. Thus g is strictly increasing, so that $\frac{1}{2}x^2 + \cos x > g(0) = 1$ on $(0, \frac{\pi}{2}]$. That is,

$$1 - \tfrac{1}{2}x^2 < \cos x < 1 \quad \text{for} \quad 0 < x \leqslant \tfrac{\pi}{2}.$$

Now let $h(x) = \sin x - x + \frac{1}{6}x^3$. This is chosen so that $h(0) = 0$ and $h'(x) = \cos x - 1 + \frac{1}{2}x^2 > 0$. Thus h is strictly increasing on $[0, \frac{\pi}{2}]$. Hence $h(x) > h(0) = 0$. Therefore

$$x - \tfrac{1}{6}x^3 < \sin x < x \quad \text{for} \quad 0 < x \leqslant \tfrac{\pi}{2}.$$

Once more! Let $k(x) = \cos x - (1 - \frac{1}{2}x^2 + \frac{1}{24}x^4)$. Then $k(0) = 0$ and $k'(x) = -\sin x + x - \frac{1}{6}x^3 = -h(x) < 0$. Thus k is strictly decreasing, and so $k(x) < 0$. Therefore

$$1 - \tfrac{1}{2}x^2 < \cos x < 1 - \tfrac{1}{2}x^2 + \tfrac{1}{24}x^4 \quad \text{for} \quad 0 < x \leqslant \tfrac{\pi}{2}.$$

5.4.6. Example. Is $\dfrac{\tan x}{x} > \dfrac{x}{\sin x}$ on $(0, \frac{\pi}{2})$?

This is true if and only if $f(x) = \sin x \tan x - x^2 > 0$. Note that $f(0) = 0$ and

$$f'(x) = \cos x \tan x + \sin x \sec^2 x - 2x = \sin x(1 + \sec^2 x) - 2x.$$

Thus $f'(0) = 0$ and

$$\begin{aligned} f''(x) &= \cos x(1 + \sec^2 x) + \sin x(2\sec x(\sec x \tan x)) - 2 \\ &= (\cos x - 2 + \sec x) + 2\,\frac{\sin^2 x}{\cos^3 x} \\ &= \Big(\sqrt{\cos x} - \frac{1}{\sqrt{\cos x}}\Big)^2 + 2\,\frac{\sin^2 x}{\cos^3 x} > 0. \end{aligned}$$

Thus $f''(x) > 0$ on $(0, \frac{\pi}{2})$, and therefore $f'(x)$ is strictly increasing on $(0, \frac{\pi}{2})$. Since $f'(x) = 0$, we have $f'(x) > 0$, and thus $f(x)$ is strictly increasing. Finally, since $f(0) = 0$, we have $f(x) > 0$ on $(0, \frac{\pi}{2})$. So the answer is yes.

5.4.7. Example. We introduce the *hyperbolic trig functions*.

$$\sinh x = \frac{e^x - e^{-x}}{2} \qquad \cosh x = \frac{e^x + e^{-x}}{2} \qquad \text{and} \qquad \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

Note that $\tanh x$ has $\lim\limits_{x\to\infty} \tanh x = 1$ and $\lim\limits_{x\to-\infty} \tanh x = -1$. Thus $\tanh x$ has horizontal asymptotes $y = \pm 1$. Note that

$$\frac{d}{dx}\sinh x = \frac{e^x + e^{-x}}{2} = \cosh x \quad \text{and} \quad \frac{d}{dx}\cosh x = \frac{e^x - e^{-x}}{2} = \sinh x.$$

Hence

$$\begin{aligned} \frac{d}{dx}\tanh x &= \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} \\ &= \frac{(e^{2x} + 2 + e^{-2x}) - (e^{2x} - 2 + e^{-2x})}{4\cosh^2 x} \\ &= \frac{1}{\cosh^2 x} =: \operatorname{sech}^2 x. \end{aligned}$$

Claim: $\tan^{-1} x < \frac{\pi}{2}\tanh x$. Note that $\frac{\pi}{2}\tanh x$ has the same horizontal asymptotes, $y = \pm\frac{\pi}{2}$, as $\tan^{-1} x$. Also $\tan^{-1} 0 = 0 = \tanh 0$. However

$$\frac{\pi}{2}\tanh'(0) = \frac{\pi}{2} > 1 = \frac{1}{1 + 0^2} = \frac{d}{dx}\tan^{-1}(0).$$

So $\frac{\pi}{2}\tanh x$ increases faster near $x = 0$. However that means that $\tan^{-1} x$ will have to increase faster for larger x to make up the difference. Subtracting them and differentiating will not help us here.

Let $f(x) = \dfrac{\tan^{-1} x}{\tanh x}$ for $x \geqslant 0$. Then $\lim\limits_{x\to\infty} \dfrac{\tan^{-1} x}{\tanh x} = \frac{\pi}{2}$. Differentiate

$$\begin{aligned} f'(x) &= \frac{\frac{1}{1+x^2}\tanh x - \tan^{-1} x \operatorname{sech}^2 x}{\tanh^2 x} \\ &= \frac{\sinh x \cosh x - (1+x^2)\tan^{-1} x}{(1+x^2)\sinh^2 x} \\ &= \frac{\frac{1}{2}\sinh 2x - (1+x^2)\tan^{-1} x}{(1+x^2)\sinh^2 x} \end{aligned}$$

The denominator is positive and $f'(0) = 0$. Let $g(x) = \frac{1}{2}\sinh 2x - (1+x^2)\tan^{-1} x$. Then

$$g'(x) = \cosh 2x - 2x\tan^{-1} x - 1$$

$$g''(x) = 2\sinh 2x - 2\tan^{-1} x - \frac{2x}{1+x^2}$$

and

$$g^{(3)}(x) = 4\cosh 2x - \frac{2}{1+x^2} - \frac{2+2x^2-2x(2x)}{(1+x^2)^2} = 4\big(\cosh 2x - \frac{1}{(1+x^2)^2}\big) > 0$$

because $\cosh 2x > 1 > \frac{1}{(1+x^2)^2}$. Now $0 = g(0) = g'(0) = g''(0)$. Since $g^{(3)} > 0$, g'' is strictly increasing and hence positive. Thus g' is strictly increasing, and hence positive; and thus g is strictly increasing. Finally, this means that $g > 0$, and thus $f' > 0$. Thus $f(x)$ is strictly increasing. In particular, $f(x) < \frac{\pi}{2}$ for all $x > 0$.

5.5. Convexity and the second derivative

5.5.1. DEFINITION. Higher order derivatives. If $f'(x)$ is differentiable, then we write $f''(x) = \frac{d}{dx}f'(x)$ for the *second derivative* of f. Similarly, for $k \geqslant 3$, we write $f^{(k)}(x) = \frac{d}{dx}f^{(k-1)}(x) =: \frac{d^k}{dx^k}f(x)$ for the *kth derivative*. If f has k derivatives and $f^{(k)}$ is continuous, we call f a C^k function.

5.5.2. DEFINITION. A function $f : (a,b) \to \mathbb{R}$ is *convex* if

$$f(tu + (1-t)v) \leqslant tf(u) + (1-t)f(v) \quad \text{for} \quad a < u < v < b \text{ and } 0 < t < 1.$$

And f is *strictly convex* if this is a strict inequality. A function g is *concave* if $f(x) = -g(x)$ is convex.

5.5.3. REMARK. The straight line through $(u, f(u))$ and $(v, f(v))$ is

$$L(tu + (1-t)v) = tf(u) + (1-t)f(v)$$
$$= f(u) + \frac{f(v) - f(u)}{v - u}(x - u) \quad \text{for} \quad t \in \mathbb{R}.$$

The values $0 < t < 1$ yield all points $x = tu + (1-t)v$ between u and v. So convexity means that the function always lies below the chord connecting two points on the graph of f.

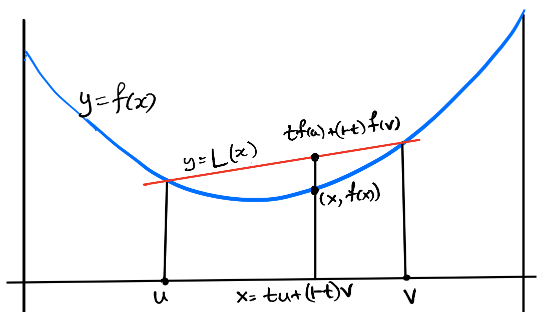


FIGURE 5.7. A convex function

5.5.4. PROPOSITION. *$f''(x) \geqslant 0$ on (a, b) implies that $f'(x)$ is increasing on (a, b), which implies that $f(x)$ is convex on (a, b). Similarly, $f''(x) \leqslant 0$ on (a, b) implies that $f'(x)$ is decreasing on (a, b), which implies that $f(x)$ is concave on (a, b). Strict inequalities yield strict convexity/concavity.*

PROOF. The first step follows from the Mean Value Theorem. Now suppose that $f'(x)$ is increasing (even if $f''(x)$ is not defined). Define

$$g(x) = f(x) - L(x) = f(x) - f(u) - \frac{f(v) - f(u)}{v - u}(x - u).$$

Then $g(u) = g(v) = 0$ and $g'(x)$ is increasing. By Rolle's Theorem, there is a point $x_0 \in (u, v)$ so that $g'(x_0) = 0$. Then $g'(x) \leqslant 0$ on $[u, x_0]$, so that $g(x)$ is decreasing and thus $g(x) \leqslant g(u) = 0$. Likewise $g'(x) \geqslant 0$ on $[x_0, v]$, so that $g(x)$ is increasing, and thus $g(x) \leqslant g(v) = 0$. Thus $g(x) \leqslant 0$ on $[u, v]$; whence $f(x) \leqslant L(x)$.

The other cases are similar. ∎

This next corollary is known as the *second derivative test* for extreme points.

5.5.5. Corollary. *Suppose that f is C^2 on (a,b). If $x_0 \in (a,b)$, $f'(x_0) = 0$ and $f''(x_0) < 0$, then x_0 is a local maximum. If $x_0 \in (a,b)$, $f'(x_0) = 0$ and $f''(x_0) > 0$, then x_0 is a local minimum.*

Proof. Since $f''(x_0) < 0$ and f'' is continuous, there is a $\delta > 0$ so that $f''(x) < 0$ on $(x_0-\delta, x_0+\delta)$. Thus $f'(x)$ is strictly decreasing on $(x_0-\delta, x_0+\delta)$. Since $f'(x_0) = 0$, $f'(x) > 0$ on $(x_0-\delta, x_0)$; and thus $f(x) \leqslant f(x_0)$ there. Similarly, $f'(x) < 0$ on $(x_0, x_0+\delta)$; and thus $f(x) \leqslant f(x_0)$ there too. Hence x_0 is a local maximum, ■

The second derivative $f''(x)$ measures the curvature of the curve $y = f(x)$. When $f''(x) > 0$, the slope $f'(x)$ is increasing, and thus the graph is curving upwards. Similarly if $f''(x) < 0$, the slope is decreasing, and the graph is curving downwards. When $f''(x)$ changes sign, the curve switches from curving down to curving up or vice versa. The transition point is called a *point of inflection*. This can happen when $f''(x) = 0$, but also when there is a vertical tangent

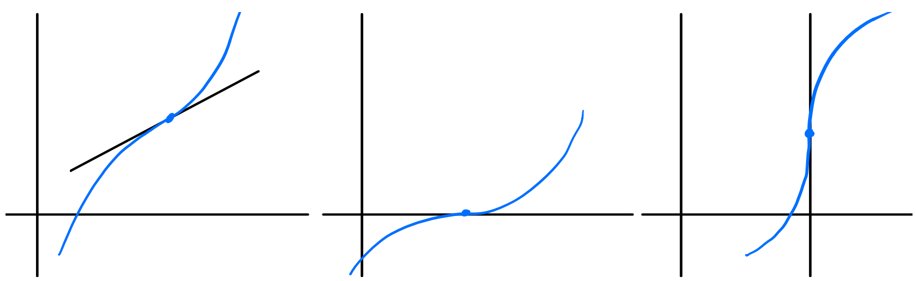


Figure 5.8. Inflection points

What does the third derivative, $f^{(3)}(x)$ represent physically? For a moving vehicle, it measures the change in acceleration. For example, when a subway car starts up, there is a jerk as it takes off. For this reason, the third derivative is sometimes called the jerk, especially in physics.

5.5.6. Example. Graph $f(x) = \dfrac{x^{1/3}}{x-1}$.

- $f(x) = 0$ only at $x = 0$.
- f is undefined at $x = 1$. $\lim\limits_{x\to 1^-} f(x) = -\infty$ and $\lim\limits_{x\to 1^+} f(x) = +\infty$. So $x = 1$ is a vertical asymptote.
- $\lim\limits_{x\to\pm\infty} f(x) = 0^+$. So $y = 0$ is a horizontal asymptote as $x \to \pm\infty$.
- $f'(x) = \dfrac{\frac{1}{3}x^{-2/3}(x-1) - x^{1/3}}{(x-1)^2} = \dfrac{(x-1)-3x}{3x^{2/3}(x-1)^2} = \dfrac{-1-2x}{3x^{2/3}(x-1)^2}$.

So $f'(-\frac{1}{2}) = 0$. Thus $(-\frac{1}{2}, \frac{\sqrt[3]{4}}{3})$ is a critical point. The denominator is positive, and the numerator changes sign from positive to negative at $x = -\frac{1}{2}$. Hence this is

a local maximum.
Also $f'(x)$ is undefined at $x = 0$ and 1. At the point $(0, 0)$, $\lim_{x\to 0} f'(x) = -\infty$. This is a vertical tangent. It is also an inflection point.
We already know what is happening near $x = 1$.
• For $f''(x)$, use the product rule rather than the quotient rule.

$$\begin{aligned} f''(x) &= \frac{-2}{3x^{2/3}(x-1)^2} + \frac{(-1-2x)(-\frac{2}{3})}{3x^{5/3}(x-1)^2} + \frac{(-1-2x)(-2)}{3x^{2/3}(x-1)^3} \\ &= \frac{-2x(x-1) + (1+2x)(\frac{2}{3})(x-1) + (1+2x)(2x)}{3x^{5/3}(x-1)^3} \\ &= \frac{2(5x^2+5x-1)}{9x^{5/3}(x-1)^3}. \end{aligned}$$

Now $f''(x) = 0$ at the roots of $5x^2 + 5x - 1 = 0$, namely $x = \dfrac{-5 \pm 3\sqrt{5}}{10}$ which are about 0.17 and -1.17. These are inflection points.
Also f'' is undefined at $0, 1$, but this doesn't add new information.

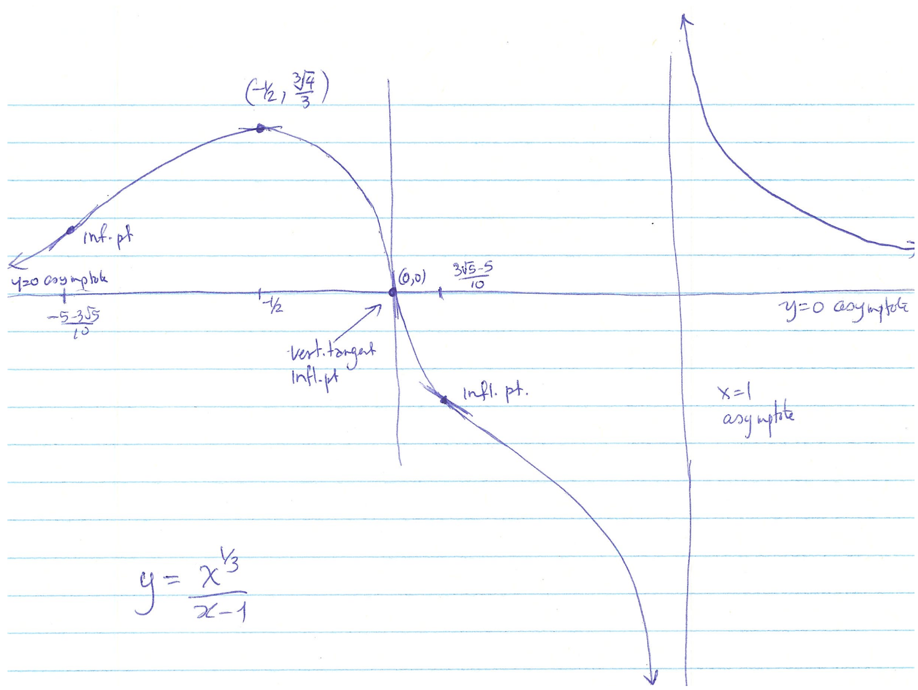


FIGURE 5.9. Graph of $f(x)$

5.6. Convexity and Jensen's Inequality

We first look at how close convex functions are to being differentiable.

5.6.1. SECANT LEMMA. *Let $f : (a, b) \to \mathbb{R}$ be convex and let $a < x < y < z < b$. Then*

$$\frac{f(y) - f(x)}{y - x} \leqslant \frac{f(z) - f(x)}{z - x} \leqslant \frac{f(z) - f(yx)}{y - y}.$$

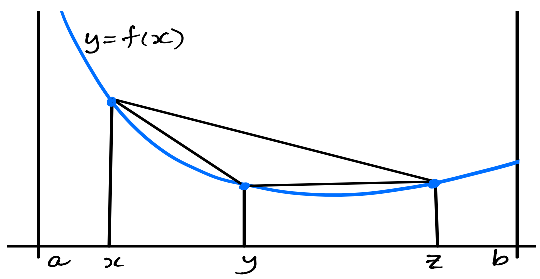


FIGURE 5.10. Secant Lemma

PROOF. Let $t = \frac{z-y}{z-x}$ and observe that $y = tx + (1 - t)z$. By convexity, $f(y) \leqslant tf(x) + (1 - t)f(z)$. Therefore

$$f(y) - f(x) \leqslant (1 - t)(f(z) - f(x)).$$

Since $y - x = (1 - t)(z - x)$, we can divide and obtain

$$\frac{f(y) - f(x)}{y - x} \leqslant \frac{f(z) - f(x)}{z - x}.$$

The second inequality is proven in the same manner. ■

5.6.2. COROLLARY. *If $f : (a, b) \to \mathbb{R}$ is convex, then it is continuous.*

PROOF. It is enough to show that f is continuous on $[c, d]$ if $a < c < d < b$. Pick c' and d' so that $a < c' < c < d < d' < b$. If $c \leqslant x < y \leqslant d$, then by the Secant Lemma,

$$C := \frac{f(c) - f(c')}{c - c'} \leqslant \frac{f(y) - f(x)}{y - x} \leqslant \frac{f(d') - f(d)}{d' - d} =: D.$$

Hence $|f(y) - f(x)| \leqslant \max\{|C|, |D|\}|y - x|$. Therefore f is Lipschitz on $[c, d]$, and so is continuous. ■

5.6.3. Remark. A convex function can fail to be continuous at an endpoint. For example, $f(x) = \begin{cases} 0 & \text{if } 0 < x < 1 \\ 1 & \text{if } x \in \{0,1\} \end{cases}$.

Also a convex function does not need to be differentiable everywhere. The function $f(x) = |x|$ on $(-1,1)$ is typical.

5.6.4. Definition. A function $f : (a,b) \to \mathbb{R}$ has a *right derivative* at x if $D_+f(x) = \lim_{h\to 0^+} \dfrac{f(x+h)-f(x)}{h}$ exists. Similarly there is a *left derivative* at x if $D_-f(x) = \lim_{h\to 0^-} \dfrac{f(x+h)-f(x)}{h}$ exists.

5.6.5. Theorem. *Let $f : (a,b) \to \mathbb{R}$ be convex. Then f has left and right derivatives at every point. If $a < x < y < b$, then*

$$D_-f(x) \leqslant D_+f(x) \leqslant D_-f(y) \leqslant D_+f(y).$$

Proof. Let $0 < h < k < \min\{x-a, \frac{1}{2}(y-x), b-y\} =: \delta$. By the Secant Lemma,

$$\frac{f(x)-f(x-k)}{k} \leqslant \frac{f(x)-f(x-h)}{h} \leqslant \frac{f(x+h)-f(x)}{h} \leqslant \frac{f(x+k)-f(x)}{k}.$$

Therefore $g(h) = \dfrac{f(x+h)-f(x)}{h}$ is an increasing function on $(-\delta,\delta)\backslash\{0\}$. By Proposition 4.5.1,

$$\lim_{h\to 0^-} \frac{f(x+h)-f(x)}{h} = \sup_{h<0} \frac{f(x+h)-f(x)}{h} = D_-f(x)$$

exists, and similarly

$$\lim_{h\to 0^+} \frac{f(x+h)-f(x)}{h} = \sup_{h>0} \frac{f(x+h)-f(x)}{h} = D_+(x)$$

exists and $D_-f(x) \leqslant D_+f(x)$.

The Secant Lemma also shows that $\dfrac{f(x+h)-f(x)}{h} \leqslant \dfrac{f(x)-f(x-k)}{k}$, and therefore $D_+f(x) \leqslant D_-f(y)$. ∎

5.6.6. Corollary. *A convex function is differentiable except on a countable set.*

Proof. By Proposition 4.5.6, the monotone function $D_-f(x)$ is continuous except on a countable set. However if $D_-f(x)$ is continuous at x, then for $\varepsilon > 0$, there is a $\delta > 0$ so that $|y-x| < \delta$ implies that $D_-f(y) - D_-f(x)| < \varepsilon$. Now if

$y > x$, choose $x < y < y' < x + \delta$ and note that

$$D_-f(x) \leqslant D_+f(x) \leqslant D_+f(y) \leqslant D_-f(y') < D_-f(x) + \varepsilon.$$

Since ε is arbitrarily small, we have that $D_+f(x) = D_-f(x)$, and it is also continuous at x. Hence f is differentiable at each x except for the countable set of discontinuities of D_-f. ■

In a certain sense, the following is just a repeated application of the defintion of convex function. However it has some surprising applications.

5.6.7. JENSEN'S INEQUALITY. *Let $f : (a, b) \to \mathbb{R}$ be convex. Suppose that $x_, \dots, x_n \in (a, b)$, $t_1, \dots, t_n \geqslant 0$ and $\sum\limits_{i=1}^{n} t_i = 1$. Then*

$$f\Big(\sum_{i=1}^{n} t_i x_i\Big) \leqslant \sum_{i=1}^{n} t_i f(x_i).$$

If f is strictly convex, and $t_i > 0$ for all i, then equality only holds if $x_1 = \dots = x_n$.

PROOF. Proceed by induction on $n \geqslant 2$. The case $n = 2$ is the definition of convexity. Now assume that the result is true for n, and let $x_, \dots, x_{n+1} \in (a, b)$, $t_1, \dots, t_{n+1} \geqslant 0$ and $\sum\limits_{i=1}^{n+1} t_i = 1$. Define

$$t = \sum_{i=1}^{n} t_i = 1 - t_{n+1} \quad \text{and} \quad y = \sum_{i=1}^{n} \frac{t_i}{t} x_i.$$

Then $y \in (a, b)$ and $\sum_{i=1}^{n} t_i = 1$, so the n case yields

$$f(y) \leqslant \sum_{i=1}^{n} \frac{t_i}{t} f(x_i).$$

Observe that $0 < t < 1$ and

$$ty + t_{n+1}x_{n+1} = \sum_{i=1}^{n+1} t_i x_i.$$

By convexity of f,

$$\begin{aligned} f\Big(\sum_{i=1}^{n+1} t_i x_i\Big) &= f(ty + t_{n+1}x_{n+1}) \leqslant tf(y) + t_{n+1}f(x_{n+1}) \\ &\leqslant t\sum_{i=1}^{n} \frac{t_i}{t} f(x_i) + t_{n+1}f(x_{n+1}) = \sum_{i=1}^{n+1} t_i f(x_i). \end{aligned}$$

Now if f is strictly convex, then the $n = 2$ case is a strict inequality if $x_1 \neq x_2$ and $0 < t_1 < 1$. So if every $t_i > 0$, then in the argument above, equality in the first line forces $y = x_{n+1}$; and equality in the inequality for $f(y)$, by the n case, forces $x_1 = \cdots = x_n = y$. ■

5.6.8. EXAMPLE. Let $f(x) = e^x$. Since $f''(x) = e^x > 0$, this is a strictly convex function on $\mathbb{R}$. Suppose that $a_i > 0$ and $t_i > 0$ for $1 \leqslant i \leqslant n$ and $\sum\limits_{i=1}^{n} t_i = 1$. Let $x_i = \ln a_i$ and apply Jensen's inequality. It says

$$a_1^{t_1} a_2^{t_2} \dots a_n^{t_n} = e^{\sum_{i=1}^n t_i \ln a_i} \leqslant \sum_{i=1}^{n} t_i e^{\ln a_i} = \sum_{i=1}^{n} t_i a_i.$$

This is known as the *generalized geometric mean–arithmetic mean inequality*.

The usual GM–AM inequality takes $t_1 = t_2 = \cdots = t_n = \frac{1}{n}$. It says that

$$\sqrt[n]{a_1 a_2 \dots a_n} \leqslant \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

5.6.9. EXAMPLE. Let $0 < s < t$ and $a_i > 0$ for $1 \leqslant i \leqslant n$. We will show that

$$\Big(\tfrac{1}{n} \sum_{i=1}^{n} a_i^s\Big)^{1/s} \leqslant \Big(\tfrac{1}{n} \sum_{i=1}^{n} a_i^t\Big)^{1/t}.$$

Let $f(x) = x^{t/s}$ for $x \geqslant 0$. Then $f''(x) = \frac{t}{s}(\frac{t}{s} - 1)x^{\frac{t}{s}-2} > 0$ if $x > 0$. Hence f is strictly convex. Let $x_i = a_i^s$ and $t_i = \frac{1}{n}$. Then by Jensen's inequality,

$$f\Big(\frac{1}{n} \sum_{i=1}^{n} x_i\Big) \leqslant \frac{1}{n} \sum_{i=1}^{n} f(x_i);$$

which becomes

$$\Big(\tfrac{1}{n} \sum_{i=1}^{n} a_i^s\Big)^{t/s} \leqslant \tfrac{1}{n} \sum_{i=1}^{n} a_i^t.$$

Take the tth root to obtain the inequality.

5.6.10. EXAMPLE. For $n \geqslant 3$, find the n-gon inscribed in a circle which has the greatest area.

Choose points $A_1, \dots, A_n$ in order around the circumference of a circle of radius r. Then the chord $\overline{A_i A_{i+1}}$ subtends an angle $\alpha_i \in (0, \pi]$ and $\sum_{i=1}^n \alpha_i = 2\pi$. (Here we mean A_1 when we write A_{n+1} and O is the centre of the circle.) The triangle OA_iA_{i+1} is isosceles with base $2r \sin \frac{\alpha_i}{2}$ and height $r \cos \frac{\alpha_i}{2}$. Thus it has area $r^2 \sin \frac{\alpha_i}{2} \cos \frac{\alpha_i}{2} = \frac{r^2}{2} \sin \alpha_i$. Therefore the area of the polygon is

$$A = A(\alpha_1, \dots, \alpha_n) = \frac{r^2}{2} \sum_{i=1}^{n} \sin \alpha_i.$$

Consider $f(x) = \sin x$ on $[0, \pi]$. Since $f''(x) = -\sin x < 0$ on $(0, \pi)$, this is a *strictly concave* function. Equivalently, $-f(x)$ is strictly convex. This reverses

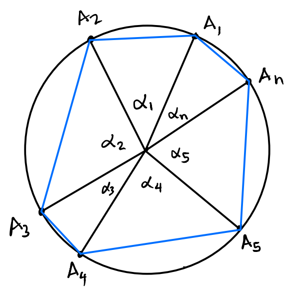


FIGURE 5.11. Maximize area of the inscribed polygon

the inequality in Jensen's formula. This becomes

$$A = \frac{r^2 n}{2}\,\frac{1}{n}\sum_{i=1}^{n} f(\alpha_i) \leqslant \frac{r^2 n}{2}\, f\Big(\frac{1}{n}\sum_{i=1}^{n}\alpha_i\Big) = \frac{r^2 n}{2}\sin\frac{2\pi}{n}.$$

Moreover the inequality is strict unless $\alpha_1 = \cdots = \alpha_n = \frac{2\pi}{n}$. The maximum area occurs only at the regular n-gon which has area $\frac{r^2}{2} n \sin\frac{2\pi}{n}$.

5.7. L'Hôpital's Rule

We begin with an intermediate value theorem for derivatives, which need not be continuous.

5.7.1. DARBOUX'S THEOREM. *Suppose that f, g are differentiable on $[a, b]$ and $f'(a) < L < f'(b)$. Then there is an $x_0 \in (a, b)$ so that $f'(x_0) = L$.*

PROOF. Let $g(x) = f(x) - Lx$. This is differentiable on $[a, b]$ with $g'(x) = f'(x) - L$; so $g'(a) < 0 < g'(b)$. Since g is differentiable, it is continuous. By the Extreme Value Theorem, g attains its minimum value at some point $x_0 \in [a, b]$. Near $x = a$,

$$0 > g'(a) = \lim_{h\to 0^+} \frac{g(a+h) - g(a)}{h}.$$

Therefore there is a $\delta > 0$ so that for $0 < h < \delta$, $g(a + h) < g(a)$, and thus the minimum does not occur at a. Similarly the minimum cannot occur at b. By Fermat's Theorem, since g is differentiable, $g'(x_0) = 0$. Therefore $f'(x_0) = L$. ■

5.7.2. CAUCHY'S MVT. *Suppose that f, g are continuous on $[a, b]$ and differentiable on (a, b). If $g'(x) \neq 0$ on (a, b), then there is a $x_0 \in (a, b)$ so that*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)}.$$

PROOF. Let $h(x) = \big(f(b) - f(a)\big)g(x) - f(x)\big(g(b) - g(a)\big)$. Then

$$h(a) = f(b)g(a) - f(a)g(b) = h(b).$$

By Rolle's Theorem, there is an $x_0 \in (a, b)$ so that

$$0 = h'(x_0) = \big(f(b) - f(a)\big)g'(x_0) - f'(x_0)\big(g(b) - g(a)\big).$$

Since $g'(x) \neq 0$, sign$(g'(x))$ is contant by Darboux's Theorem. Thus $g(x)$ is strictly monotone, and thus $g(b) - g(a) \neq 0$. Now divide by $g'(x_0))\big(g(b) - g(a)\big)$ to get the result. ■

The main result of this section is very popular with student's, but in practice, there are usually superior methods. The author warns the reader to never ever apply L'Hôpital's rule to $\lim_{x\to 0} \frac{\sin x}{x} = 1$. This limit must be known before one can differentiate $\sin x$, and thus the argument is circular! The hypotheses of this result are crucial, and need to be verified in any application.

5.7.3. L'HÔPITAL'S RULE. *Suppose that f, g are differentiable on an open interval J with c as an endpoint ($\pm\infty$ are allowed). Suppose that*

(1) $g(x) \neq 0$ *and* $g'(x) \neq 0$ *for* $x \in J$.

(2) $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$ *or* $\lim_{x\to c} |f(x)| = \lim_{x\to c} |g(x)| = +\infty$.

(3) $\lim_{x\to c} \frac{f'(x)}{g'(x)} = L$ *exists.*

Then $\lim_{x\to c} \frac{f(x)}{g(x)} = L$.

PROOF. Case 1. Suppose that $c \in \mathbb{R}$ and $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$. (This may be a one-sided limit, in which case $x \to c^+$ or $x \to c_-$ as appropriate.) Since $\lim_{x\to c} \frac{f'(x)}{g'(x)} = L$, given $\varepsilon > 0$, there is a $\delta > 0$ so that $0 < |x - c| < \delta$ implies that

$$L - \varepsilon < \frac{f'(x)}{g'(x)} < L + \varepsilon.$$

If we define $f(c) = g(c) = 0$, then both f and g are continuous at c. Then we can apply the Cauchy Mean Value Theorem on the interval $[c, x]$ (or $[x, c]$) with $|x - c| < \delta$ to get

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(c)}{g(x) - g(c)} = \frac{f'(x_0)}{g'(x_0)} \in (L - \varepsilon, L + \varepsilon).$$

Since $\varepsilon > 0$ was arbitrary, $\lim_{x\to c} \frac{f(x)}{g(x)} = L$.

Case 2. Suppose that $c \in \mathbb{R}$ and $\lim_{x\to c} |f(x)| = \lim_{x\to c} |g(x)| = +\infty$. For $\varepsilon > 0$, find $\delta > 0$ for $\lim_{x\to c} \frac{f'(x)}{g'(x)} = L$ as before. Consider points $c < x < y < c + \delta$ (or $c - \delta < y < x < c$) and apply the Cauchy Mean Value Theorem to $[x, y]$ to get a point $x_0 \in (x, y)$ so that

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(x) - f(y)}{g(x) - g(y)} = \frac{\frac{f(x)}{g(x)} - \frac{f(y)}{g(x)}}{1 - \frac{g(y)}{g(x)}} \in (L - \varepsilon, L + \varepsilon).$$

Let $x \to c$ while holding y fixed. Then $\frac{f(y)}{g(x)} \to 0$ and $\frac{g(y)}{g(x)} \to 0$. Therefore for x sufficiently close to c, this quantity will be within ε of $\frac{f(x)}{g(x)}$, and hence $\left|\frac{f(x)}{g(x)} - L\right| < 2\varepsilon$. This means that $\lim_{x\to c} \frac{f(x)}{g(x)} = L$.

Case 3. When $c = \pm\infty$, make the substitution $u = \frac{1}{x}$. Make J smaller to exclude 0 if necessary. Set $F(u) = f(\frac{1}{u})$ and $G(u) = g(\frac{1}{u})$ and $I = \{\frac{1}{x} : x \in J\}$. Then

(1) $G(u) = g(\frac{1}{u}) \neq 0$ and $G'(u) = g'(\frac{1}{u})\big(\frac{-1}{u^2}\big) \neq 0$ for $u \in I$.

(2) $\lim_{u\to 0} F(u) = \lim_{x\to c} |f(x)| \in \{0, \infty\}$ and similarly $\lim_{u\to 0} G(u) = \lim_{x\to c} |g(x)| \in \{0, \infty\}$.

(3) $\displaystyle \lim_{u\to 0} \frac{F'(u)}{G'(u)} = \lim_{u\to 0} \frac{f'(\frac{1}{u})\big(\frac{-1}{u^2}\big)}{g'(\frac{1}{u})\big(\frac{-1}{u^2}\big)} = \lim_{x\to c} \frac{f'(x)}{g'(x)} = L.$

So the hypotheses for either Case 1 or 2 is satisfied. Therefore

$$L = \lim_{u\to 0} \frac{F(u)}{G(u)} = \lim_{x\to c} \frac{f(x)}{g(x)}. \qquad \blacksquare$$

5.7.4. EXAMPLE. Let $a > 0$. Find $\displaystyle \lim_{x\to a^+} \frac{\sqrt{x} + \sqrt{x-a} - \sqrt{a}}{\sqrt{x^2 - a^2}}$.

Both numerator $f(x)$ and denominator $g(x)$ tend to 0 as $x \to a^+$, and $g(x) = \sqrt{x^2 - a^2} \neq 0$. Compute $g'(x) = \dfrac{2x}{2\sqrt{x^2 - a^2}} \neq 0$ as well. So we compute

$$\begin{aligned} \lim_{x\to a^+} \frac{f'(x)}{g'(x)} &= \lim_{x\to a^+} \frac{\frac{1}{2\sqrt{x}} + \frac{1}{2\sqrt{x-a}}}{\frac{2x}{2\sqrt{x^2-a^2}}} = \lim_{x\to a^+} \frac{(\sqrt{x-a} + \sqrt{x})}{2\sqrt{x}\sqrt{x-a}} \frac{\sqrt{x^2 - a^2}}{x} \\ &= \lim_{x\to a^+} \frac{\sqrt{x^2 - a^2} + \sqrt{x(x+a)}}{2x\sqrt{x}} = \frac{\sqrt{2}a}{2a\sqrt{a}} = \frac{1}{\sqrt{2a}}. \end{aligned}$$

Thus $\lim_{x\to a^+} \dfrac{\sqrt{x}+\sqrt{x-a}-\sqrt{a}}{\sqrt{x^2-a^2}} = \frac{1}{\sqrt{2a}}$.

5.7.5. EXAMPLE. Compute $\lim_{x\to 0}\left(\dfrac{\tan x}{x}\right)^{1/x^2}$. This function is even, so consider $x > 0$. Then take the logarithm $\dfrac{\ln\tan x - \ln x}{x^2} =: \dfrac{f(x)}{g(x)}$. We now have a "0/0" situation because $\lim_{x\to 0} \ln\dfrac{\tan x}{x} = \ln 1 = 0$; and the denominator x^2 and its derivative $2x$ never vanish on $(0,1)$. Compute

$$\begin{aligned}\lim_{x\to 0^+}\frac{f'(x)}{g'(x)} &= \lim_{x\to 0^+}\frac{\frac{\sec^2 x}{\tan x}-\frac{1}{x}}{2x} = \lim_{x\to 0^+}\frac{x-\sin x\cos x}{2x^2\sin x\cos x}\\ &= \lim_{x\to 0^+}\frac{2x-\sin 2x}{2x^2\sin 2x} = \lim_{x\to 0^+}\frac{2x-\sin 2x}{4x^3}\,\frac{2x}{\sin 2x}\end{aligned}$$

Now $\lim_{x\to 0^+}\dfrac{2x}{\sin 2x} = 1$ is a known limit. Consider $\lim_{x\to 0^+}\frac{2x-\sin 2x}{4x^3}$. This is another "0/0" situation, and we have (setting $y = 2x$)

$$\lim_{x\to 0^+}\frac{2-2\cos 2x}{12x^2} = \lim_{y\to 0^+}\frac{2-2\cos y}{3y^2} = \frac{1}{3}.$$

by Example 3.1.7. Therefore L'Hôpital's rule applies, and $\lim_{x\to 0^+}\dfrac{2x-\sin 2x}{4x^3} = \frac{1}{3}$. Thus $\lim_{x\to 0^+}\dfrac{f'(x)}{g'(x)} = \frac{1}{3}$. So by L'Hôpital's rule again, $\lim_{x\to 0^+}\dfrac{f(x)}{g(x)} = \frac{1}{3}$. Finally we have

$$\lim_{x\to 0}\left(\frac{\tan x}{x}\right)^{1/x^2} = e^{1/3}.$$

5.7.6. EXAMPLE. Consider $\lim_{x\to\infty}\dfrac{x-\sin x}{x+\sin x}$. Here we have an "$\infty/\infty$" situation, and the denominator is never 0 for $x > 1$. However, $\lim_{x\to\infty}\dfrac{1-\cos x}{1+\cos x}$ does not exist, in part because the denominator vanishes periodically. So L'Hôpital's rule does not apply. In fact,

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

5.8. Newton's Method

This is an algorithm for approximately solving as equation $f(x) = 0$ using calculus. The idea is to repeatedly solve for where the tangent line crosses the axis. The power of the method lies in its rapid convergence.

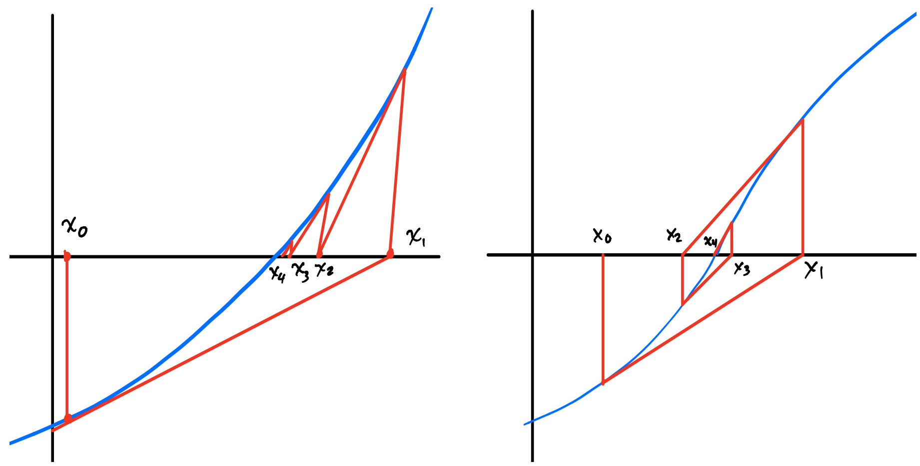


FIGURE 5.12. Newton's method

Given an approximation x_n, find the tangent line

$$Tx = f(x_n) + f'(x_n)(x - x_n).$$

Solve for the root of T, $x = x_n - \dfrac{f(x_n)}{f'(x_n)}$.

Hypotheses

- $f(x)$ has a zero x_*, i.e., $f(x_*) = 0$.
- f is C^2 near x_*.
- $f'(x_*) \neq 0$.

Algorithm.

- Start with an initial guess x_0 "sufficiently close" to x_* with $f(x_0)$ 'small'.
- Given x_n, set $\quad x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$.

Error estimates.
Fix a small interval $[a, b]$ containing x_* and x_0 with x_* near the middle on which $f'(x) \neq 0$. This is necessary to ensure that the iterates stay in the interval. Let

$$m = \min_{a \leqslant x \leqslant b} |f'(x)| \quad \text{and} \quad C = \max_{a \leqslant x \leqslant b} |f''(x)|.$$

• $|x_n - x_*| \leqslant \dfrac{|f(x_n)|}{m}$. This follows from the MVT

$$\frac{f(x_n)}{x_n - x_*} = \frac{f(x_n) - f(x_*)}{x_n - x_*} = f'(c)$$

for some point c between x_n and x_*, so $c \in [a, b]$. Thus

$$|x_n - x_*| = \frac{|f(x_n)|}{|f'(c)|} \leqslant \frac{|f(x_n)|}{m}.$$

• Consider the map $Sx = x - \dfrac{f(x)}{f'(x)}$. Then

$$S'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2}.$$

Then $S(x_*) = x_*$ and $S'(x_*) = 0$. So there is an interval around x_* on which $|S'(x)| < \frac{1}{2}$. Then the MVT shows that

$$\frac{|x_{n+1} - x_*|}{|x_n - x_*|} = \frac{|S(x_n) - S(x_*)|}{|x_n - x_*|} = |S'(c)| < \frac{1}{2}.$$

Therefore, $|x_{n+1} - x_*| < \frac{1}{2}|x_n - x_*|$. In particular, this shows that $\lim_{n\to\infty} x_n = x_*$. This is decent convergence, but in fact it improves as we get closer because the estimate on S' improves.

• $|x_{n+1} - x_*| < \dfrac{C}{m}|x_n - x_*|^2$. In the first estimate, we found a point c so that $x_n - x_* = \dfrac{f(x_n)}{f'(c)}$ and the algorithm says $x_{n+1} - x_n = -\dfrac{f(x_n)}{f'(x_n)}$. Thus

$$\begin{aligned} x_{n+1} - x_* &= (x_n - x_*) + (x_{n+1} - x_n) \\ &= \frac{f(x_n)}{f'(c)} - \frac{f(x_n)}{f'(x_n)} \\ &= \frac{f(x_n)(f'(x_n) - f'(c))}{f'(c)f'(x_n)} \end{aligned}$$

We apply MVT to $f'(x_n) - f'(c)$ to find a point d between x_n and c, and hence between x_n and x_*, so that $f'(x_n) - f'(c) = f''(d)(x_n - c)$. Plugging this back in yields

$$\begin{aligned} x_{n+1} - x_* &= \frac{f(x_n)(f'(x_n) - f'(c))}{f'(c)f'(x_n)} \\ &= \frac{f(x_n)}{f'(c)} \frac{f''(d)(x_n - c)}{f'(x_n)} \\ &= \frac{f''(d)}{f'(x_n)}(x_n - x_*)(x_n - c) \end{aligned}$$

Therefore, since $|x_n - c| < |x_n - x_*|$,

$$|x_{n+1} - x_*| \leqslant \frac{C}{m}|x_n - x_*|^2.$$

This is known as quadratic convergence, and says that the number of decimals of accuracy roughly doubles with each iteration. Very quickly on a computer, the issue becomes the ability to do high precision calculation.

5.8.1. Example. Compute $\sqrt{149}$.
Let $f(x) = x^2 - 149$ and restrict the domain to $[12, 13]$. Then $f'(x) = 2x$ and $f''(x) = 2$. Thus $m = \min_{x\in[12,13]} 2x = 24$ and $C = \max_{x\in[12,13]} 2 = 2$. Hence $\frac{C}{m} = \frac{1}{12}$.

Start with $x_0 = 12$ and $x_* = \sqrt{149}$. The first estimate shows that

$$|x_0 - x_*| \leqslant \frac{|f(x_0)|}{m} = \frac{5}{24} < 0.21.$$

The algorithm is

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 149}{2x_n} = \frac{x_n^2 + 149}{2x_n} = \frac{1}{2}\Big(x_n + \frac{149}{x_n}\Big).$$

The map S has derivative $S'(x) = \frac{|f(x)f''(x)|}{f'(x)^2} \leqslant \frac{(169-149)2}{24^2} < \frac{1}{12}$. So

$$|x_{n+1} - x_*| \leqslant \frac{1}{12}|x_n - x_*|,$$

but the quadratic estimate actually is better immediately.

$$x_1 = 12.2083 \qquad |x_1 - x_*| \leqslant \frac{1}{12}(0.21)^2 < 3.7\,10^{-3}$$

$$x_2 = 12.2065557 \qquad |x_2 - x_*| \leqslant \frac{1}{12}(3.7 \cdot 10^{-3})^2 < 1.16\,10^{-6}$$

$$x_3 = 12.2065556157337036 \qquad |x_3 - x_*| \leqslant \frac{1}{12}(1.16 \cdot 10^{-6})^2 < 3.6\,10^{-13}$$

$$x_4 = 12.20655561573370295189\ldots \qquad |x_4 - x_*| \leqslant \frac{1}{12}(3.6 \cdot 10^{-13})^2 < 1.2\,10^{-26}$$

So in four steps, I have surpassed the accuracy of my computer's significant digits.

5.8.2. Example. Solve $2^x = 4x$.
Let $f(x) = 2^x - 4x$. By inspection, $x = 4$ is a solution, but the graph shows a second smaller solution. Compute $f'(x) = 2^x \ln 2 - 4$ and $f''(x) = 2^x(\ln 2)^2 > 0$. Since $f'' > 0$, the curve is convex, and hence there are only two solutions. Now $f(0) = 1 > 0 > -2 = f(1)$, so the second solution is in $(0, 1)$ by IVT. We will work in $[0, 1]$. Then

$$m = \min_{x\in[0,1]} |f'(x)| = 4 - 2\ln 2 \approx 2.6 \quad \text{and} \quad C = \max_{x\in[0,1]} |f''(x)| = 2(\ln 2)^2 < 1.$$

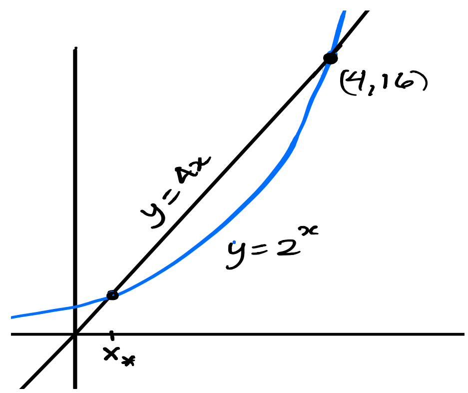


FIGURE 5.13. Graph of 2^x and $4x$

Thus $\frac{C}{m} < 0.37$. Start with $x_0 = 0.25$. Then

$$|x_* - x_0| < \frac{|f'(x_0)|}{m} = \frac{0.1892}{2.6} < 0.073.$$

Applying the algorithm we get

$x_1 = 0.309579$ $\quad |x_1 - x_*| \leqslant 0.37(.073)^2 < .002$

$x_2 = 0.30990692$ $\quad |x_2 - x_*| \leqslant 0.37(.002)^2 < 1.5\,10^{-6}$

$x_3 = 0.3099069323806$ $\quad |x_3 - x_*| \leqslant 0.37(1.5 \cdot 10^{-6})^2 < 8.3\,10^{-13}$

$x_4 = 0.30990693238069056545 46$ $\quad |x_4 - x_*| \leqslant 0.37(8.3 \cdot 10^{-13})^2 < 2.5\,10^{-25}$

Exercises for Chapter 5

1. Let $f(x) = e^{-1/x^2}$ for $x \neq 0$ and $f(0) = 0$.
(a) Compute the derivative of f at $x = 0$ from the definition.
(b) Compute the derivative of f for $x \neq 0$. Is $f'(x)$ continuous?

2. (a) Simplify the expression $f(x) = \sec(\tan^{-1}(\sin(\tan^{-1} x)))$ for $x \in \mathbb{R}$.
(b) Graph $f(x)$ including identification of inflection points.
HINT: work with the simplified formula.

3. Two corridors meet at right angles. They have widths a meters and b meters, respectively. Find the length of the longest ladder that can be moved around the corner (while keeping the ladder horizontal). HINT: use an angle as your variable.

4. (a) A movie screen 10m high is mounted on the wall 5m above the floor. At what distance from the screen does a point on the floor subtend the greatest angle and find the angle.

HINT: consider $A = (0,5)$, $B = (0,15)$ and $X = (x,0)$ for $x \geqslant 0$. Find an expression for the angle $\angle AXB$ as a function of x.

(b) **Bonus.** Show that the solution to (a) may be obtained by finding the circle through A and B which is tangent to the x-axis. Explain why this is the solution using Euclidean geometry.

5. (a) Find the trapezoid that fits inside a semicircle of radius r with the two sides parallel to the diameter which has the largest area.

(b) **Bonus.** Find the quadrilateral of largest area that fits inside a semicircle of radius r.

6. (a) Sketch $\csc(x)$ for $-\pi/2 \leqslant x \leqslant \pi/2$, $x \neq 0$ and its inverse function $\csc^{-1}(x)$.

(b) Compute the derivative of $\csc^{-1}(x)$. **Warning**: be careful with signs.

7. (a) Show that $\tan x > x + \dfrac{x^3}{3} + \dfrac{2x^5}{15}$ for $0 < x < \dfrac{\pi}{2}$.

(b) Show that $\tan x < x + \dfrac{x^3}{3} + \dfrac{2x^5}{5}$ for $0 < x < 1$. HINT: $\cos x > 1 - \dfrac{x^2}{2}$ on $(0, \frac{\pi}{2})$.

8. (a) Let $f(x) = x^2 \sin(1/x)$ for $x \neq 0$ and $f(0) = 0$. Show that 0 is a critical point of f that is not a local maximum nor a local minimum nor an inflection point.

(b) Let $f : (a,b) \to \mathbb{R}$ be a continuous function which is differentiable except possibly at x_0, and $\lim\limits_{x \to x_0} f'(x) = L$ exists. Prove that $f'(x_0) = L$. HINT: use MVT on $[x_0, x_0 + h]$.

(c) Let $g(x) = 2x^2 + f(x)$. Show that g does have a global minimum at 0, but $g'(x)$ changes sign infinitely often on both $(0,\varepsilon)$ and $(-\varepsilon,0)$ for any $\varepsilon > 0$.

9. Let $f : (a,b) \to \mathbb{R}$ be a continuous function which is differentiable except possibly at x_0, and $\lim\limits_{x \to x_0} f'(x) = L$ exists. Prove that $f'(x_0) = L$. HINT: use MVT on $[x_0, x_0 + h]$.

10. Let f be a differentiable function on $[a,b]$, but the derivative may be discontinuous.

(a) Suppose that $f'(a) < 0 < f'(b)$. Show that the minimum of f does not occur at an endpoint. What can you conclude?

(b) Suppose that $f'(a) < L < f'(b)$. Prove that there is an $c \in (a,b)$ such that $f'(c) = L$.

(c) Prove that if $f'(x)$ is monotone, then $f'(x)$ is continuous.

11. Graph $f(x) = \exp\left(\frac{x^2-3}{x^2-x}\right)$. You can use a computer program to find the approximate roots of the degree 5 polynomial that occurs in the second derivative.

12. Let $n \geqslant 1$, and let f and g have nth order derivatives on (a, b).
Show that $(fg)^{(n)}(x) = \sum\limits_{k=0}^{n} \binom{n}{k} f^{(n-k)}(x) g^{(k)}(x)$.

13. (a) Show that $\ln(1 + e^x)$ is convex. Sketch it.
(b) Show that for any $a_i > 0$, $\left(1 + \sqrt[n]{a_1 a_2 \ldots a_n}\right)^n \leqslant \prod_{i=1}^{n}(1 + a_i)$.
(c) Suppose that $0 < a = x_0 < x_1 < \cdots < x_n = b$. Show that the maximum of $\dfrac{x_0 x_1 \ldots x_n}{(x_0 + x_1)(x_1 + x_2) \ldots (x_{n-1} + x_n)}$ occurs when $\dfrac{x_{i+1}}{x_i}$ are all equal for $0 \leqslant i < n$.

14. (a) Suppose that $f''(a)$ exists. Show that

$$\lim_{h \to 0} \frac{f(a+h) + f(a-h) - 2f(a)}{h^2} = f''(a).$$

(b) Show by example that this limit may exist even when $f''(a)$ does not.

15. Compute the following limits.
(a) $\lim\limits_{x \to 0^+} \dfrac{(1+x)^{1/x} - e}{x}$
(b) $\lim\limits_{x \to 0} \cot^2 x - \dfrac{1}{x^2}$

16. Let $f(x) = e^{-2x}(\cos x + 2 \sin x)$ and $g(x) = e^{-x}(\cos x + \sin x)$.
Find all of the errors (if any) in the following L'Hôpital's rule argument:

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)} = \lim_{x \to \infty} \frac{5}{2} e^{-x} = 0.$$

17. Suppose that f is differentiable on $[a, b]$, $f(a) = 0$ and $|f'(x)| \leqslant A|f(x)|$ for all $x \in [a, b]$, where A is a positive real number. Prove that f is constant.

18. Show that $p(x) = x^3 + x + 1$ has exactly one real root. Find this root of $p(x)$ to 6 decimal places using Newton's method. Provide the algorithm and show your error estimates.

You can use a computer to do the calculations, but an analysis of the error should be done by hand. Try using Maple: (punctuation is critical!)

```
with(Student[Calculus1]):
NewtonsMethod(x^5+x+1,x=-0.5,
    view=[-1.5 .. 0.0, DEFAULT],output=sequence);
```

This only yield 5 terms, so substitute the last result back in and repeat.

CHAPTER 6

The Riemann Integral

6.1. Archimedes

Archimedes of Syracuse (287 BC to 212 BC) developed many ideas in mathematics, physics and astronomy which have greatly influenced modern thinking. He is famous for showing that the area of a circle of radius r is πr^2 and computing π to some accuracy. For example, he showed that $\frac{223}{71} < \pi < \frac{22}{7}$. He also computed the area of a section of a parabola. He always used geometric ideas from Euclidean geometry. We will derive his formula is a more analytic/algebraic manner.

Consider the following figure: the parabola is given by $y = ax^2 + bx + c$ and $x_1 < x_2$ are given. Here $a > 0$. Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ be the corresponding points on the curve. We wish to compute the area of the sector of the parabola determined by the curve and the line between P_1 and P_2.

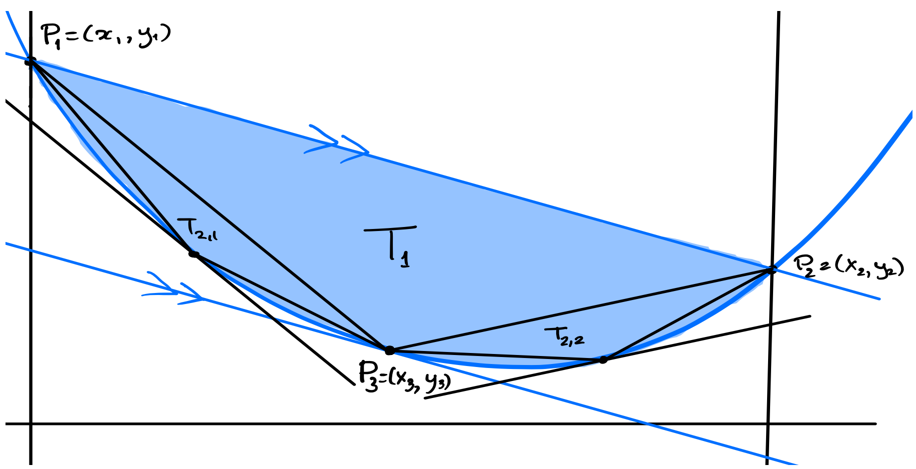


FIGURE 6.1. Area of a sector of a parabola

The line $\overline{P_1P_2}$ from P_1 to P_2 has slope

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{a(x_2^2 - x_1^2) + b(x_2 - x_1) + c - c}{x_2 - x_1} = a(x_2 + x_1) + b.$$

There is a line parallel to $\overline{P_1P_2}$ which is tangent to the parabola. Archimedes showed (without calculus) that it intersects the curve at the midpoint, $P_3 = (x_3, y_3)$

where $x_3 = \frac{x_1+x_2}{2}$. We can see this by differentiating $f(x) = ax^2 + bx + c$ to get $f'(x) = 2ax + b$ and observing that $f'(x_3) = a(x_1 + x_2) + b$, which is the same slope. Since $f'(x)$ is a non-constant line, there is only one point with this slope.

The triangle $\triangle P_1P_2P_3$ has base $x_2 - x_1$ and average height $\frac{y_1+y_2}{2} - y_3$. Thus its area is

$$\begin{aligned} T_1 &= \frac{x_2 - x_1}{4}(y_1 + y_2 - 2y_3) \\ &= \frac{x_2 - x_1}{4}\Big(a\big(x_1^2 + x_2^2 - 2\tfrac{(x_1+x_2)^2}{4}\big) + b\big(x_1 + x_2 - 2\tfrac{x_1+x_2}{2}\big) + (c + c - 2c)\Big) \\ &= \frac{x_2 - x_1}{4}\frac{a(x_1 - x_2)^2}{2} = \frac{a}{8}(x_2 - x_1)^3. \end{aligned}$$

Archimedes method is to repeat the process with the two parabolic sectors remaining; and then with four sectors, etc., which will tile the sector with an infinite family of triangles of the same nature. Summing all of the areas of these triangles will yield a formula for the area. That is, in the sector of the parabola from P_1 to P_3, we obtain a triangle $P_1P_3P_4$, where $P_4 = (x_4, y_4)$ and $x_4 = \frac{x_1+x_3}{2}$, has area

$$T_{2,1} = \frac{a}{8}(x_3 - x_1)^2 = \frac{a}{32}(x_2 - x_1)^2.$$

Similarly setting $P_5 = (x_5, y_5)$ with $x_5 = \frac{x_3+x_2}{2}$ yields a triangle $P_2P_3P_5$ with area

$$T_{2,2} = \frac{a}{8}(x_2 - x_3)^2 = \frac{a}{32}(x_2 - x_1)^2 = T_{2,1} =: T_2 = \frac{T_1}{4}.$$

This produces two triangles of $\frac{1}{4}$ the original size. At the next stage, there are four triangles with a base half again as big, so they will have area $T_3 = \frac{T_2}{4} = 4^{-2}T_1$. Thus at the nth stage, we obtain 2^{n-1} triangles of area $T_n = 4^{1-n}T_1$. Summing, we get

$$\text{area of sector} = \sum_{n\geqslant 1} 2^{n-1}4^{1-n}T_1 = T_1 \sum_{n\geqslant 1} 2^{1-n} = 2T_1 = \frac{a}{4}(x_2 - x_1)^3.$$

This is a rather complicated and ingenious procedure. Such an approach requires a new idea for each geometric region. Moreover the applicability is rather limited. Before the Fundamental Theorem of Calculus, providing an excellent general method for computing areas, mathematicians developed a variety of more sophisticated methods for computing areas. The main take-away from this should be that the methods of calculus are powerful.

6.2. The Riemann Integral

Riemann's method of integration is to approximate the area under a curve from both above and below by a collection of rectangles. This yields an upper and lower bound for the area under the curve. See figure. If these two estimates converge to the same value as the width of the rectangles decreases to 0, we say that the function is Riemann integrable, and assign this limit to be the value of the integral. We will make this precise in the following definitions.

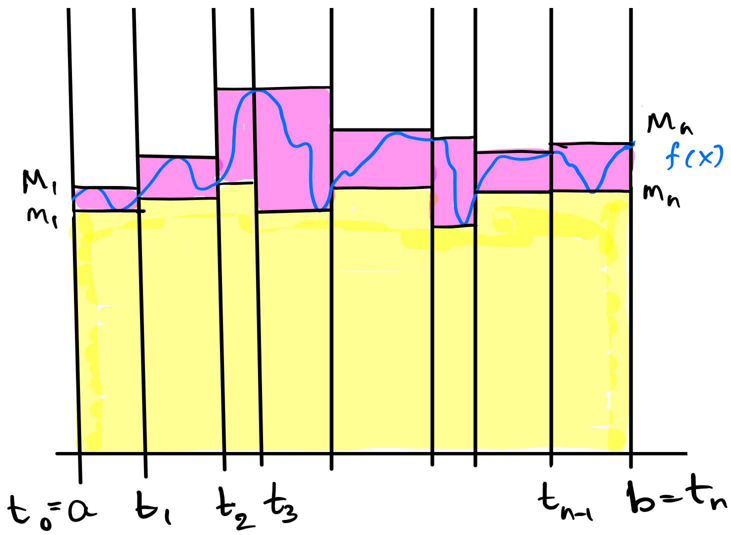


FIGURE 6.2. A Riemann sum

6.2.1. DEFINITION. A *partition* of an interval $[a, b]$ is a finite sequence $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_n = b\}$. A partition $\mathcal{Q}$ *refines* $\mathcal{P}$ if $\mathcal{P} \subset \mathcal{Q}$. Given two partitions $\mathcal{P}_1$ and $\mathcal{P}_2$, the *common refinement* $\mathcal{P}_1 \vee \mathcal{P}_2$ is the ordered list using the points of $\mathcal{P}_1 \cup \mathcal{P}_2$ is order without repetition. The *mesh* of $\mathcal{P}$ is

$$\operatorname{mesh} \mathcal{P} = \max_{1 \leqslant i \leqslant n} t_i - t_{i-1}.$$

Set $\Delta t_i = t_i - t_{i-1}$.

6.2.2. DEFINITION. Let $f(x)$ be a *bounded* function defined on $[a, b]$ and let $\mathcal{P}$ be a partition of $[a, b]$. Set

$$m_i = m_i(f, \mathcal{P}) = \inf\{f(x) : t_{i-1} \leqslant x \leqslant t_i\}$$
$$M_i = M_i(f, \mathcal{P}) = \sup\{f(x) : t_{i-1} \leqslant x \leqslant t_i\}.$$

The *lower sum* $L(f, \mathcal{P})$ and *upper sum* $U(f, \mathcal{P})$ are given by

$$L(f, \mathcal{P}) = \sum_{i=1}^{n} m_i(t_i - t_{i-1}) = \sum_{i=1}^{n} m_i \Delta t_i$$
$$U(f, \mathcal{P}) = \sum_{i=1}^{n} M_i(t_i - t_{i-1}) = \sum_{i=1}^{n} M_i \Delta t_i.$$

Given a collection of points $\{x_i\}$ with $t_{i-1} \leqslant x_i \leqslant t_i$ for $1 \leqslant i \leqslant n$, the *Riemann sum* is

$$R(f, \mathcal{P}, \{x_i\}) = \sum_{i=1}^{n} f(x_i)(t_i - t_{i-1}) = \sum_{i=1}^{n} f(x_i)\Delta t_i.$$

We make the following routine observations which will be useful.

(1) $L(f, \mathcal{P}) \leqslant R(f, \mathcal{P}, \{x_i\}) \leqslant U(f, \mathcal{P})$.

(2) If $\mathcal{P} \subset \mathcal{Q}$, then

$$L(f, \mathcal{P}) \leqslant L(f, \mathcal{Q}) \leqslant U(f, \mathcal{Q}) \leqslant U(f, \mathcal{P}).$$

To see this, consider an interval $[t_{i-1}, t_i]$ of the partition $\mathcal{P}$. In the refinement $\mathcal{Q}$, this interval is divided into smaller intervals. Thus the lower bounds on these smaller intervals will be at least as large as m_i, but possible larger; and the upper bounds will be no greater than M_i, and possibly smaller.

(3) Given $\mathcal{P}_1$ and $\mathcal{P}_2$, since $\mathcal{P}_1 \vee \mathcal{P}_2$ refines both, we obtain

$$L(f, \mathcal{P}_1) \leqslant L(f, \mathcal{P}_1 \vee \mathcal{P}_2) \leqslant U(f, \mathcal{P}_1 \vee \mathcal{P}_2) \leqslant U(f, \mathcal{P}_2).$$

(4) It follows that $\{L(f, \mathcal{P}) : \mathcal{P} \text{ a partition}\}$ is bounded above by any element of $\{U(f, \mathcal{P}) : \mathcal{P} \text{ a partition}\}$; and likewise this latter set is bounded below.

(5) This procedure doesn't make sense if f is unbounded.

Based on these observations, we are led to the following definition.

6.2.3. DEFINITION. If f is a bounded function on $[a, b]$, define

$$L(f) = \sup \{L(f, \mathcal{P}) : \mathcal{P} \text{ a partition}\}$$
$$U(f) = \inf \{U(f, \mathcal{P}) : \mathcal{P} \text{ a partition}\}.$$

Say that $f(x)$ is Riemann integrable if $L(f) = U(f)$. Denote this common value by $\displaystyle\int_a^b f(x)\,dx$.

Note that $L(f) \leqslant U(f)$. Moreover if we can find a sequence of partitions $\mathcal{P}_n$ so that

$$\lim_{n\to\infty} L(f, \mathcal{P}_n) = L = \lim_{n\to\infty} U(f, \mathcal{P}_n),$$

then necessarily f is integrable with integral L. We will spell this out in more detail after a couple of examples.

6.2.4. EXAMPLE. Let $c > 1$ and set $f(x) = c^x$; and fix $a < b$. Consider the partition $\mathcal{P}_n$ with evenly spaced points $t_i = a + i\frac{b-a}{n}$ for $0 \leqslant i \leqslant n$. Then $\Delta t_i = \frac{b-a}{n}$. Since f is monotone increasing, we have $m_i = c^{t_{i-1}}$ and $M_i = c^{t_i}$ for

$1 \leqslant i \leqslant n$. Compute

$$\begin{aligned} L_n = L(f, \mathcal{P}_n) &= \sum_{i=1}^{n} c^{t_{i-1}} \Delta t_i \\ &= \sum_{i=1}^{n} c^a c^{\frac{b-a}{n}(i-1)} \frac{b-a}{n} = \frac{b-a}{n} c^a \sum_{i=1}^{n} \left(c^{\frac{b-a}{n}}\right)^{i-1} \\ &= \frac{b-a}{n} c^a \frac{c^{b-a} - 1}{c^{\frac{b-a}{n}} - 1} = \left(c^b - c^a\right) \frac{h_n}{c^{h_n} - 1} \end{aligned}$$

In the last line, we sum a geometric series and set $h_n = \frac{b-a}{n}$. Observe that

$$\lim_{n\to\infty} \frac{c^{h_n} - 1}{h_n} = f'(0) = c^x \ln c\Big|_{x=0} = \ln c.$$

Therefore

$$\lim_{n\to\infty} L_n = \frac{c^b - c^a}{\ln c}.$$

Similarly we compute

$$U_n = U(f, \mathcal{P}_n) = \sum_{i=1}^{n} c^{t_i} \Delta t_i = c^{\frac{b-a}{n}} L_n.$$

Hence

$$\lim_{n\to\infty} U_n = \lim_{n\to\infty} c^{\frac{b-a}{n}} \lim_{n\to\infty} L_n = \frac{c^b - c^a}{\ln c}.$$

Hence f is integrable, and $\displaystyle\int_a^b c^x \, dx = \frac{c^b - c^a}{\ln c}$.

6.2.5. Example. Let $f(x) = x^p$ for $p \geqslant 0$ and let $0 \leqslant a < b$. In this example, we will use a different partition. Let $h_n = \left(\frac{b}{a}\right)^{1/n}$ and set $t_i = a h_n^i$ for $1 \leqslant i \leqslant n$. Then $\Delta t_i = t_{i-1}(h_n - 1)$. Since f is monotone increasing, we have $m_i = t_{i-1}^p$ and $M_i = t_i^p = m_i h_n^p$. Calculate

$$\begin{aligned} L_n = L(f, \mathcal{P}_n) &= \sum_{i=1}^{n} m_i \Delta t_i = \sum_{i=1}^{n} t_{i-1}^p t_{i-1}(h_n - 1) \\ &= (h_n - 1) \sum_{i=1}^{n} \left(a h_n^{i-1}\right)^{p+1} = a^{p+1}(h_n - 1) \sum_{i=1}^{n} h_n^{(i-1)(p+1)} \\ &= a^{p+1}(h_n - 1) \frac{h_n^{n(p+1)} - 1}{h_n^{p+1} - 1} = a^{p+1}\left(\left(\tfrac{b}{a}\right)^{p+1} - 1\right) \frac{h_n - 1}{h_n^{p+1} - 1} \\ &= (b^{p+1} - a^{p+1}) \frac{h_n - 1}{h_n^{p+1} - 1}. \end{aligned}$$

Now we compute the limit

$$\lim_{n\to\infty} \frac{h_n^{p+1} - 1}{h_n - 1} = \frac{d}{dx}(x^{p+1})\Big|_{x=1} = p + 1.$$

Therefore

$$\lim_{n\to\infty} L_n = \frac{b^{p+1} - a^{p+1}}{p+1}.$$

Similarly,

$$U_n = U(f, \mathcal{P}_n) = \sum_{i=1}^{n} M_i \Delta t_i = \sum_{i=1}^{n} m_i h_n^p \Delta t_i = L_n h_n^p.$$

Since $\lim_{n\to\infty} h_n = 1$, we have

$$\lim_{n\to\infty} U_n = \lim_{n\to\infty} L_n = \frac{b^{p+1} - a^{p+1}}{p+1}.$$

Hence f is integrable, and $\displaystyle\int_a^b x^p\, dx = \frac{b^{p+1} - a^{p+1}}{p+1}$.

6.2.6. RIEMANN'S CONDITION. *Let $f(x)$ be a bounded function on $[a, b]$. Then f is Riemann integrable if and only if Riemann's condition holds:*

$(\ast)$ *for all $\varepsilon > 0$, there is a partition $\mathcal{P}$ so that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon$.*

PROOF. If f is Riemann integrable, we have $L(f) = U(f)$. We can choose partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ so that

$$L(f, \mathcal{P}_1) > L(f) - \frac{\varepsilon}{2} \quad \text{and} \quad U(f, \mathcal{P}_2) < U(f) + \frac{\varepsilon}{2}.$$

Set $\mathcal{P} = \mathcal{P}_1 \vee \mathcal{P}_2$. Then

$$L(f) - \frac{\varepsilon}{2} < L(f, \mathcal{P}_1) \leqslant L(f, \mathcal{P}) \leqslant U(f, \mathcal{P}) \leqslant U(f, \mathcal{P}_2) < U(f) + \frac{\varepsilon}{2} = L(f) + \frac{\varepsilon}{2}.$$

Hence $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon$. So $(\ast)$ holds.

Conversely if $(\ast)$ holds, take $\varepsilon > 0$ and find the appropriate partition $\mathcal{P}$. Then

$$U(f) - L(f) \leqslant U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon.$$

This is valid for all $\varepsilon > 0$; hence $U(f) = L(f)$ and f is Riemann integrable. ■

This allows us to provide a number of conditions equivalent to integrability.

6.2.7. THEOREM. *Let $f(x)$ be a bounded function on $[a, b]$. Then the following are equivalent:*

(1) *f is Riemann integrable.*

(2) *For all $\varepsilon > 0$, there is a partition $\mathcal{P}$ so that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon$.*

(3) *For all* $\varepsilon > 0$*, there is a* $\delta > 0$ *so that* mesh $\mathcal{P} < \delta$ *implies*
$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon.$$

(4) *For all* $\varepsilon > 0$*, there is a* $\delta > 0$ *so that if* mesh $\mathcal{P} < \delta$ *and* $x_i \in [t_{i-1}, t_i]$*, then* $R(f,\mathcal{P},\{x_i\}) - L(f) < \varepsilon$*.*

(5) *There are partitions* $\mathcal{P}_n$ *so that* $\lim\limits_{n\to\infty} L(f,\mathcal{P}_n) = \lim\limits_{n\to\infty} U(f,\mathcal{P}_n)$*.*

PROOF. We just established the equivalence of (1) and (2). Clearly (3) implies (2). Conversely, suppose that $U(f,\mathcal{P}) - L(f,\mathcal{P}) < \frac{\varepsilon}{2}$. Let n be the size of $\mathcal{P}$ and let $M = \sup\{|f(x)| : a \leqslant x \leqslant b\}$. Define $\delta = \dfrac{\varepsilon}{4n(M+1)}$. Suppose that mesh $\mathcal{Q} < \delta$. Set $\mathcal{R} = \mathcal{P} \vee \mathcal{Q}$. Then

$$L(f,\mathcal{R}) - L(f,\mathcal{R}) < U(f,\mathcal{P}) - L(f,\mathcal{P}) < \frac{\varepsilon}{2}.$$

The idea is to compare the upper and lower sums for $\mathcal{R}$ and $\mathcal{Q}$. Since $\mathcal{P}$ has $n-1$ points t_i in (a,b), at most $n-1$ intervals of $\mathcal{Q}$ are subdivided by some t_i in $\mathcal{R}$. Say $\mathcal{Q} = \{a = s_0 < s_1 < \dots < s_m = b\}$ and for some (or all) $1 \leqslant i \leqslant n-1$, there are j_i so that $s_{j_i - 1} < t_i < s_{j_i}$. On $[s_{j_i-1}, s_{j_i}]$, we have $-M \leqslant m_{j_i} \leqslant M_{j_i} \leqslant M$. In the calculation of $U(f,\mathcal{Q}) - L(f,\mathcal{Q})$, this interval contributes

$$(M_{j_i} - m_{j_i})\Delta s_{j_i} \leqslant 2M\delta < \frac{\varepsilon}{2n}.$$

For the same interval, the contribution to $U(f,\mathcal{R}) - L(f,\mathcal{R})$ is positive. On the remaining intervals, the contribution to $U(f,\mathcal{Q}) - L(f,\mathcal{Q})$ and $U(f,\mathcal{R}) - L(f,\mathcal{R})$ are equal. Thus,

$$U(f,\mathcal{Q}) - L(f,\mathcal{Q}) < U(f,\mathcal{R}) - L(f,\mathcal{R}) + (n-1)\frac{\varepsilon}{2n} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Assuming (3), we know that (1) holds so that $L(f) = U(f)$. Given $\varepsilon > 0$, use (3) to get δ. Then

$$R(f,\mathcal{P},\{x_i\}) - L(f) < U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon.$$

Conversely, if (4) holds and $\varepsilon > 0$, take the δ corresponding to $\varepsilon/2$. Given $\mathcal{P}$ with mesh $\mathcal{P} < \delta$ and size n, choose $x_i \in [t_{i-1}, t_i]$ with $f(x_i) > M_i - \frac{\varepsilon}{2(b-a)}$. Then

$$\begin{aligned} U(f,\mathcal{P}) - L(f,\mathcal{P}) &= \sum_{i=1}^{n} (M_i - m_i)\Delta t_i < \sum_{i=1}^{n} \big(f(x_i) + \frac{\varepsilon}{2(b-a)} - m_i\big)\Delta t_i \\ &= R(f,\mathcal{P},\{x_i\}) - L(f,\mathcal{P}) + \frac{\varepsilon}{2(b-a)} \sum_{i=1}^{n} \Delta t_i < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

So (3) and (4) are equivalent.

If (2) holds, we can choose $\mathcal{P}_n$ so that $U(f,\mathcal{P}_n) - L(f,\mathcal{P}_n) < \frac{1}{n}$. Then

$$\lim_{n\to\infty} L(f,\mathcal{P}_n) \leqslant \lim_{n\to\infty} U(f,\mathcal{P}_n) \leqslant \lim_{n\to\infty} L(f,\mathcal{P}_n) + \tfrac{1}{n} = \lim_{n\to\infty} L(f,\mathcal{P}_n).$$

So (5) holds. Conversely, if (5) holds, then

$$\lim_{n\to\infty} U(f,\mathcal{P}_n) - L(f,\mathcal{P}_n) = 0.$$

Hence for any $\varepsilon > 0$, there is an n so that $U(f,\mathcal{P}_n) - L(f,\mathcal{P}_n) < \varepsilon$. Therefore (2) and (5) are equivalent. ■

6.3. Basic Properties of the Integral

In this section, we verify some elementary properties which will simplify our calculations.

6.3.1. PROPOSITION. *Let $f(x), g(x)$ be Riemann integrable functions on $[a,b]$.*

(1) *If $c \in \mathbb{R}$, then $cf(x)$ is Riemann integrable and $\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$.*

(2) *$f+g$ is Riemann integrable, and*

$$\int_a^b f(x)+g(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

(3) *If $f(x) \leqslant g(x)$ on $[a,b]$, then $\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx$.*

(4) *If $a < c < b$, then f is Riemann integrable on $[a,c]$ and $[c,b]$, and*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

PROOF. (1) If $c \geqslant 0$, we have $m_i(cf) = cm_i(f)$ and $M_i(cf) = cM_i(f)$. So the result is straightforward. If $c < 0$, we have $m_i(cf) = cM_i(f)$ and $M_i(cf) = cm_i(f)$; and the argument is similar.

(2) If $\mathcal{P}$ is any partition of $[a,b]$,

$$\begin{aligned} m_i(f+g) &= \inf_{t_{i-1}\leqslant x\leqslant t_i} f(x)+g(x) \\ &\geqslant \inf_{t_{i-1}\leqslant x\leqslant t_i} f(y) + \inf_{t_{i-1}\leqslant y\leqslant t_i} g(y) = m_i(f)+m_i(g). \\ M_i(f+g) &= \sup_{t_{i-1}\leqslant x\leqslant t_i} f(x)+g(x) \\ &\leqslant \sup_{t_{i-1}\leqslant x\leqslant t_i} f(y) + \sup_{t_{i-1}\leqslant y\leqslant t_i} g(y) = M_i(f)+M_i(g). \end{aligned}$$

Hence

$$L(f,\mathcal{P})+L(g,\mathcal{P}) = \sum_{i=1}^{n} \Big(m_i(f)+m_i(g)\Big)\Delta t_i \leqslant \sum_{i=1}^{n} m_i(f+g)\Delta t_i = L(f+g,\mathcal{P}).$$

$$U(f+g,\mathcal{P}) = \sum_{i=1}^{n} M_i(f+g)\Delta t_i \leqslant \sum_{i=1}^{n} \big(M_i(f)+M_i(g)\big)\Delta t_i = U(f,\mathcal{P})+U(g,\mathcal{P}).$$

Thus if $\mathcal{P}$ and $\mathcal{Q}$ are partitions,

$$L(f,\mathcal{P}) + L(g,\mathcal{Q}) \leqslant L(f,\mathcal{P}\vee\mathcal{Q}) + L(g,\mathcal{P}\vee\mathcal{Q}) \leqslant L(f+g,\mathcal{P}\vee\mathcal{Q})$$

and consequently, $L(f) + L(g) \leqslant L(f+g)$. Similarly, $U(f+g) \leqslant U(f) + U(g)$. Therefore if f and g are integrable, then

$$\begin{aligned}\int_a^b f(x)\,dx + \int_a^b g(x)\,dx &= L(f) + L(g) \leqslant L(f+g)\\ &\leqslant U(f+g) \leqslant U(f) + U(g) = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.\end{aligned}$$

Hence $L(f+g) = U(f+g) = \displaystyle\int_a^b f(x)\,dx + \int_a^b g(x)\,dx$. It follows that $f+g$ is Riemann integrable and

$$\int_a^b f(x) + g(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

(3) is left as an exercise.

(4) follows by using a partition containing the point c. In this case, we see that

$$\begin{aligned}L(f,\mathcal{P}) &= L(f|_{[a,c]},\mathcal{P}|_{[a,c]}) + L(f|_{[c,b]},\mathcal{P}|_{[c,b]})\\ U(f,\mathcal{P}) &= U(f|_{[a,c]},\mathcal{P}|_{[a,c]}) + U(f|_{[c,b]},\mathcal{P}|_{[c,b]}).\end{aligned}$$

Details are left to the reader. ■

The following is an immediate consequence of (3).

6.3.2. COROLLARY. *Let $f(x)$ be a Riemann integrable function on $[a,b]$. If $m \leqslant f(x) \leqslant M$, then*

$$m(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M(b-a).$$

In particular, if $|f(x)| \leqslant M$, then $\left|\displaystyle\int_a^b f(x)\,dx\right| \leqslant M(b-a)$.

6.4. Riemann integrable functions

In this section, we prove that two large classes of functions are Riemann integrable.

6.4.1. THEOREM. *Every monotone function f on $[a, b]$ is Riemann integrable.*

PROOF. We may suppose that $f(x)$ is monotone increasing. Consider the partition $\mathcal{P}_n$ with evenly spaced points $t_i = a + i\frac{b-a}{n}$ for $0 \leqslant i \leqslant n$. Then $\Delta t_i = \frac{b-a}{n}$. Since f is monotone increasing, we have $m_i = f(t_{i-1})$ and $M_i = f(t_i)$ for $1 \leqslant i \leqslant n$. Compute

$$L(f, \mathcal{P}_n) = \sum_{i=1}^{n} m_i \Delta t_i = \sum_{i=1}^{n} f(t_{i-1}) \frac{b-a}{n}$$

$$U(f, \mathcal{P}_n) = \sum_{i=1}^{n} M_i \Delta t_i = \sum_{i=1}^{n} f(t_i) \frac{b-a}{n}.$$

Subtracting, we obtain

$$U(f, \mathcal{P}_n) - L(f, \mathcal{P}_n) = \big(f(t_n) - f(t_0)\big)\frac{b-a}{n} = \frac{\big(f(b) - f(a)\big)(b-a)}{n}.$$

Hence for $\varepsilon > 0$, pick n so large that $U(f, \mathcal{P}_n) - L(f, \mathcal{P}_n) < \varepsilon$. Therefore f satisfies Riemann's condition, and hence is Riemann integrable. ■

We can extend this to cover most functions in our everyday experience.

6.4.2. DEFINITION. A function $f(x)$ is *piecewise monotone* if there is a partition $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_n = b\}$ of $[a, b]$ so that f is monotone on $[t_{i-1}, t_i]$ for $1 \leqslant i \leqslant n$.

6.4.3. COROLLARY. *Every piecewise monotone function is Riemann integrable.*

PROOF. Begin with a partition $\mathcal{P}$ so that f is monotone on each segment. Then apply Theorem 6.4.1 on each segment. ■

6.4.4. EXAMPLE. Here is an example of a non-integrable function. Let

$$f(x) = \begin{cases} 0 & \text{if } x \notin \mathbb{Q} \\ 1 & \text{if } x \in \mathbb{Q} \end{cases} \quad \text{for} \quad x \in [a, b].$$

Then for any partition $\mathcal{P}$, we have $m_i = 0$ and $M_i = 1$. Therefore $L(f, \mathcal{P}) = 0$ and $U(f, \mathcal{P}) = b - a$. So $L(f) = 0 \neq b - a = U(f)$; and thus f is not Riemann integrable.

6.4.5. EXAMPLE. Here is a discontinuous example of an integrable function. Let

$$f(x) = \begin{cases} 0 & \text{if } x \notin \mathbb{Q} \\ \frac{1}{q} & \text{if } x = \frac{p}{q} \in \mathbb{Q}, q \in \mathbb{N}, \gcd(p,q) = 1 \end{cases} \quad \text{for} \quad x \in [0,1].$$

This is known as Thomae's function. It is easy to see that f is discontinuous at every rational point.

Let $\varepsilon > 0$. Choose an integer $N \geqslant 4$ so that $\frac{1}{N} < \varepsilon/2$. Let $\mathcal{P}$ be a partition with mesh $\mathcal{P} < \frac{\varepsilon}{2N^2}$ so that $t_i \notin \mathbb{Q}$ if $1 \leqslant i \leqslant n-1$. Since every interval contains irrational points, $m_i = 0$ for all i; so $L(f,\mathcal{P}) = 0$. The function f takes values $\geqslant \frac{1}{N}$ only on $0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \dots, \frac{1}{N}, \dots, \frac{N-1}{N}$. This consists at most $2 + 1 + 2 + \cdots + (N-1) = 2 + \frac{N(N-1)}{2} < N^2$ points. So on $n - N^2$ intervals, $M_i < \frac{1}{N^2}$ and on the remaining N^2 intervals, $M_i \leqslant 1$. Therefore

$$U(f,\mathcal{P}) < \frac{1}{N^2}(1) + 1(N^2)\frac{\varepsilon}{2N^2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since ε is arbitrary, f satisfies Riemann's condition. Hence $\displaystyle\int_0^1 f(x)\,dx = 0$.

We have seen that there are continuous functions which oscillate rapidly and are not piecewise monotone. However they are still integrable.

6.4.6. THEOREM. *Every continuous function f on $[a,b]$ is Riemann integrable.*

PROOF. Recall Theorem 4.6.3 that shows that every continuous function on a closed bounded interval $[a,b]$ is uniformly continuous. This means that given $\varepsilon > 0$, there is a $\delta > 0$ so that $|x-y| < \delta$ implies that $|f(x) - f(y)| < \varepsilon$. Given $\varepsilon > 0$, take a $\delta > 0$ which works for $\frac{\varepsilon}{b-a}$.

Let $\mathcal{P}$ be any partition with mesh $\mathcal{P} < \delta$. Then

$$M_i - m_i = \sup_{t_{i-1} \leqslant x \leqslant t_i} f(x) - \inf_{t_{i-1} \leqslant y \leqslant t_i} f(y) = \sup_{t_{i-1} \leqslant x,y \leqslant t_i} f(x) - f(y) \leqslant \frac{\varepsilon}{b-a}.$$

Therefore

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) = \sum_{i=1}^{n} (M_i - m_i)\Delta t_i \leqslant \frac{\varepsilon}{b-a} \sum_{i=1}^{n} \Delta t_i = \varepsilon.$$

This verifies Riemann's condition, and thus $f(x)$ is Riemann integrable. ■

For continuous functions, we have an integral version of the Mean Value Theorem.

6.4.7. THEOREM. *Suppose that $f(x)$ is continuous on $[a,b]$. Then there is a point $c \in [a,b]$ so that*

$$\frac{1}{b-a}\int_a^b f(x)\,dx = f(c).$$

PROOF. Let $m = \inf_{a \leqslant x \leqslant b} f(x)$ and $M = \sup_{a \leqslant x \leqslant b} f(x)$. By the Extreme Value Theorem, the minimum and maximum values are attained, say $f(c_1) = m$ and $f(c_2) = M$. Corollary 6.3.2 shows that

$$m \leqslant L := \frac{1}{b-a}\int_a^b f(x)\,dx \leqslant M.$$

By the Intermediate Value Theorem, there is a point c between c_1 and c_2 so that $f(c) = L$. ∎

6.5. More integrable functions

Here are some constructions that preserve integrability.

6.5.1. PROPOSITION. *Let $f(x), g(x)$ be Riemann integrable functions on $[a,b]$.*

(1) $(f \vee g)(x) = \max\{f(x), g(x)\}$ *and* $(f \wedge g)(x) = \min\{f(x), g(x)\}$ *are Riemann integrable. In particular,* $|f|(x) = |f(x)|$ *is Riemann integrable.*

(2) $(fg)(x) = f(x)g(x)$ *is Riemann integrable.*

(3) *if* $\inf_{a \leqslant x \leqslant b} |g(x)| = \delta > 0$, *then* f/g *is Riemann integrable.*

PROOF. (1) First consider $f \vee 0$. Let $\varepsilon > 0$ and let $\mathcal{P}$ be a partition so that $U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon$. Let m_i and M_i have the usual definition. Then

$$n_i = \inf\{(f \vee 0)(x) : t_{i-1} \leqslant x \leqslant t_i\} = m_i \vee 0$$

and

$$N_i = \sup\{(f \vee 0)(x) : t_{i-1} \leqslant x \leqslant t_i\} = M_i \vee 0.$$

Therefore, $N_i - n_i \leqslant M_i - m_i$ with equality only if $m_i \geqslant 0$ or $m_i = M_i$. Hence

$$\begin{aligned} U(f \vee 0, \mathcal{P}) - L(f \vee 0, \mathcal{P}) &= \sum_{i=1}^{n} (N_i - n_i)\Delta t_i \\ &\leqslant \sum_{i=1}^{n} (M_i - m_i)\Delta t_i \\ &= U(f,\mathcal{P}) - L(f\mathcal{P}) < \varepsilon. \end{aligned}$$

Therefore $f \vee 0$ is Riemann integrable by Riemann's condition. Since

$$f \vee g = f + ((g - f) \vee 0) \quad \text{and} \quad f \wedge g = f - ((f - g) \vee 0),$$

they are also Riemann integrable. In particular, $|f| = f \vee -f$ is Riemann integrable.

(2) Let $F = 1 + \sup\{|f(x)| : a \leqslant x \leqslant b\}$ and $G = 1 + \sup\{|g(x)| : a \leqslant x \leqslant b\}$. Given $\varepsilon > 0$, choose a partition $\mathcal{P}$ so that

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \frac{\varepsilon}{2G} \quad \text{and} \quad U(g, \mathcal{P}) - L(g, \mathcal{P}) < \frac{\varepsilon}{2F}.$$

Let m_i, M_i come from f, and let n_i and N_i come from g. Similarly let

$$l_i = \inf\{f(x)g(x) : t_{i-1} \leqslant x \leqslant t_i\} \quad \text{and} \quad L_i = \sup\{f(x)g(x) : t_{i-1} \leqslant x \leqslant t_i\}.$$

Now

$$\begin{aligned} f(x)g(x) - f(y)g(y) &= \big(f(x) - f(y)\big)g(y) + f(x)\big(g(x) - g(y)\big) \\ &\leqslant G(M_i - m_i) + F(N_i - n_i). \end{aligned}$$

Thus

$$L_i - l_i = \sup_{t_{i-1} \leqslant x, y \leqslant t_i} f(x)g(x) - f(y)g(y) \leqslant G(M_i - m_i) + F(N_i - n_i).$$

Therefore

$$\begin{aligned} U(fg, \mathcal{P}) - L(fg, \mathcal{P}) &= \sum_{i=1}^{n} (L_i - l_i)\Delta t_i \\ &\leqslant G \sum_{i=1}^{n} (M_i - m_i)\Delta t_i + F \sum_{i=1}^{n} (N_i - n_i)\Delta t_i \\ &= G\big(U(f, \mathcal{P}) - L(f, \mathcal{P})\big) + F\big(U(g, \mathcal{P}) - L(g, \mathcal{P})\big) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Thus fg satisfies Riemann's condition, and therefore is integrable.

(3) First consider $1/g$. Note that

$$\Big|\frac{1}{g(x)} - \frac{1}{g(y)}\Big| = \Big|\frac{g(x) - g(y)}{g(x)g(y)}\Big| \leqslant \delta^{-2}|g(x) - g(y)|.$$

For $\varepsilon > 0$, choose a partition $\mathcal{P}$ so that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \delta^2\varepsilon$. Let

$$k_i = \inf\big\{\tfrac{1}{g(x)} : t_{i-1} \leqslant x \leqslant t_i\big\} \quad \text{and} \quad K_i = \sup\big\{\tfrac{1}{g(x)} : t_{i-1} \leqslant x \leqslant t_i\big\}.$$

Then $K_i - k_i \leqslant \delta^{-2}(M_i - m_i)$. Therefore

$$\begin{aligned} U(\tfrac{1}{g}, \mathcal{P}) - L(\tfrac{1}{g}, \mathcal{P}) &= \sum_{i=1}^{n} (K_i - k_i)\Delta t_i \leqslant \delta^{-2} \sum_{i=1}^{n} (M_i - m_i)\Delta t_i \\ &= \delta^{-2}\big(U(g, \mathcal{P}) - L(g, \mathcal{P})\big) < \delta^{-2}\delta^2\varepsilon = \varepsilon. \end{aligned}$$

Thus $\frac{1}{g}$ satisfies Riemann's condition, and therefore is integrable. Finally the quotient $\frac{f}{g} = f\left(\frac{1}{g}\right)$ is integrable by (2). ∎

6.6. Fundamental Theorem of Calculus

The Fundamental Theorem of Calculus is called fundamental because it provides a deep link between integration and differentiation. This link allows for a variety of useful computational methods for computing integrals.

Recall that a function $g(x)$ on $[a, b]$ is Lipschitz if there is a constant C so that $|g(x) - g(y)| \leqslant C|x - y|$ for $x, y \in [a, b]$. In Example 4.1.3(3), we showed that Lipschitz functions are uniformly continuous. And Example 4.6.2(1) showed that if g is differentiable and g' is bounded by M, then g is Lipschitz with constant M.

6.6.1. Fundamental Theorem of Calculus. *Let $f(x)$ be a Riemann integrable function on $[a, b]$, and define $F(x) = \int_a^x f(t)\,dt$. Then F is Lipschitz, and hence continuous. If f is continuous at a point $x_0 \in [a, b]$, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.*

Proof. Riemann integrable functions are bounded. Let

$$M = \|f\|_\infty := \sup\{|f(x)| : a \leqslant x \leqslant b\}.$$

Then if $a \leqslant x < y \leqslant b$, we have

$$|F(y) - F(x)| = \Big|\int_x^y f(t)\,dt\Big| \leqslant \int_x^y M\,dt = M|y - x|.$$

Hence F is Lipschitz with constant M, and thus is continuous.

Suppose that f is continuous at x_0. Given $\varepsilon > 0$, pick $\delta > 0$ so that $|y - x_0| < \delta$ implies that $|f(y) - f(x_0)| < \varepsilon$. Then for $0 < h < \delta$, we have

$$\begin{aligned}\left|\frac{F(x_0 + h) - F(x_0)}{h} - f(x_0)\right| &= \left|\frac{1}{h}\int_{x_0}^{x_0+h} f(t)\,dt - \frac{1}{h}\int_{x_0}^{x_0+h} f(x_0)\,dt\right| \\ &\leqslant \frac{1}{h}\int_{x_0}^{x_0+h} |f(t) - f(x_0)|\,dt < \frac{1}{h}\int_{x_0}^{x_0+h} \varepsilon\,dt = \varepsilon.\end{aligned}$$

Similarly one shows that

$$\left|\frac{F(x_0) - F(x_0 - h)}{h} - f(x_0)\right| < \varepsilon.$$

Therefore $F'(x_0) = \lim_{h\to 0} \frac{F(x_0 + h) - F(x_0)}{h} = f(x_0)$. ∎

The most important consequence of this theorem is the following method for computing integrals.

6.6.2. COROLLARY. *Suppose that $f(x)$ is continuous on $[a,b]$ and G is a differentiable function such that $G'(x) = f(x)$ for $x \in [a,b]$. Then*

$$\int_a^b f(x)\,dx = G(b) - G(a).$$

PROOF. Let $F(x) = \int_a^x f(t)\,dt$. By the Fundamental Theorem of Calculus, $F'(x) = f(x) = G'(x)$ for $x \in [a,b]$. Therefore $(G-F)' = G' - F' = 0$. By the Mean Value Theorem, $G - F$ is constant, say c. Therefore $G(x) = c + \int_a^x f(t)\,dt$. Hence $G(b) - G(a) = \int_a^b f(t)\,dt$. ■

6.6.3. DEFINITION. If $f(x)$ is a continuous function on $[a,b]$, an *antiderivative* of f is a continuous function $F(x)$ on $[a,b]$ which is differentiable on (a,b) with $F'(x) = f(x)$.

Note that if F and G are two antiderivatives of f, then $(F-G)' = 0$ on (a,b) and thus $F - G$ is constant. Thus the set of all antiderivatives of f have the form $F(x) + c$ for some constant c. We will indicate the antiderivative of f by

$$\int f(x)\,dx = F(x) + c.$$

6.6.4. EXAMPLE. Let

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

Then one can see geometrically that

$$F(x) = \int_{-1}^x f(t)\,dt = \begin{cases} -1 - x & \text{if } -1 \leqslant x < 0 \\ \ \ x - 1 & \text{if } \quad 0 \leqslant x \leqslant 1 \end{cases} = |x| - 1.$$

Then F is differentiable on $[-1,0) \cup (0,1]$, but fails to be differentiable at $x = 0$, which is a point of discontinuity for f.

The Fundamental Theorem of Calculus can be pushed a bit further, dropping continuity of f provided that f is a derivative and also Riemann integrable.

6.6.5. Fundamental Theorem of Calculus II. *Suppose that $f(x)$ is a Riemann integrable function on $[a, b]$ and that $F(x)$ is a differentiable function such that $F'(x) = f(x)$ for $x \in [a, b]$. Then*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

PROOF. Let $\varepsilon > 0$. Choose a partition $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_n = b\}$ of $[a, b]$ so that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon$. By the Mean Value Theorem, there is a point $x_i \in (t_{i-1}, t_i)$ so that

$$\frac{F(t_i) - F(t_{i-1})}{t_i - t_{i-1}} = F'(x_i) = f(x_i).$$

Therefore

$$R(f, \mathcal{P}, \{x_i\}) = \sum_{i=1}^{n} f(x_i)(t_i - t_{i-1}) = \sum_{i=1}^{n} F(t_i) - F(t_{i-1}) = F(b) - F(a).$$

Now both $R(f, \mathcal{P}, \{x_i\})$ and $\displaystyle\int_a^b f(x)\,dx$ lie in the interval $[L(f, \mathcal{P}), U(f, \mathcal{P})]$. Therefore

$$\left| \int_a^b f(x)\,dx - \big(F(b) - F(a)\big) \right| = \left| \int_a^b f(x)\,dx - R(f, \mathcal{P}, \{x_i\}) \right| < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we have $\displaystyle\int_a^b f(x)\,dx = F(b) - F(a)$. ■

The hypothesis that f is Riemann integrable does not follow just because f is the derivative of a function $F(x)$, even if it is bounded. That is the point of the following rather difficult example.

6.6.6. Example. We start with a variant of the Cantor set constructed as follows. Start with $[0, 1]$. Remove the middle open interval of length $\frac{1}{4}$, i.e. $(\frac{3}{8}, \frac{5}{8})$. Then from the middle of the remaining intervals, remove an interval of length 4^{-2}, namely $(\frac{5}{32}, \frac{7}{32})$ and $(\frac{25}{32}, \frac{27}{32})$. At the nth stage, remove an open interval of length 4^{-n} from the middle of each of the remaining 2^{n-1} closed intervals. Then the total length of the pieces removed is

$$\sum_{n=1}^{\infty} \frac{2^{n-1}}{4^n} = \frac{1}{2} \sum_{n=1}^{\infty} 2^{-n} = \frac{1}{2}.$$

Let $\{I_k = (a_k, b_k) : k \geqslant 1\}$ be a list of the removed intervals. Define $U = \bigcup_{k \geqslant 1} I_k$ and $C = [0, 1] \backslash U$. Inside each interval I_k, let $J_k = (c_k, d_k)$ be an interval of length $|I_k|^2$ with the same centre; and let $V = \bigcup_{k \geqslant 1} J_k$. Notice that after the nth stage, the 2^n intervals that remain all have the same length, and so it tends to 0. That means that C does not contain any interval. Thus if $x \in C$, then $(x - \varepsilon, x + \varepsilon)$

intersects U. This means that there is a sequence of points in U converging to x. While x could be an endpoint of some I_k, it can be approached from the other side as well. So we can assume that the points belong to distinct intervals I_k, so that the lengths of the I_k will tend to 0. Now we can alter such a sequence by replacing a point of I_k with any point we choose in J_k, and keep the same limit.

Define functions

$$f_k(x) = \begin{cases} 0 & \text{if } x \notin J_k \\ \sin\left(\frac{2\pi(x-c_k)}{d_k-c_k}\right) & \text{if } x \in J_k \end{cases} \quad \text{and} \quad f(x) = \sum_{k\geqslant 1} f_k(x).$$

Note that the functions $f(x)$ are continuous and have disjoint supports. Thus $f(x)$ is bounded by 1, and in particular $f(e_k) = 1$ where $e_k = \frac{3c_k+d}{4}$. However it is not continuous because each $x \in C$ is a limit of points e_k in J_k, and f takes the value 1 at these points and is 0 on C. We will show that $f(x)$ is not Riemann integrable, but is a derivative.

Let $\mathcal{P}$ be any partition of $[0, 1]$. The collection

$$A = \{i : (t_{i-1}, t_i) \subset I_k \text{ for some } k\}$$

has total length

$$\sum_{i\in A} t_i - t_{i-1} \leqslant \sum_{k\geqslant 1} |I_k| = \frac{1}{2}.$$

Let $B = \{1, \dots, n\}\backslash A$. Then

$$\sum_{i\in B} t_i - t_{i-1} = 1 - \sum_{i\in A} t_i - t_{i-1} \geqslant \frac{1}{2}.$$

If $i \in B$, then there is a point $x_i \in (t_{i-1}, t_i) \cap C$. As observed above, there is a sequence of points e_k in certain I_k's converging to x. Thus there is a point $y_i \in (t_{i-1}, t_i)$ which is one of these e_k. Hence $f(y_k) = 1$ and $f(x) = 0$. That is $m_i = 0$ and $M_i = 1$ if $i \in B$. Therefore

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) \geqslant \sum_{i\in B}(M_i - m_i)\Delta t_i = \sum_{i\in B} t_i - t_{i-1} \geqslant \frac{1}{2}.$$

Hence the Riemann condition is not satisfied, so f is not Riemann integrable.

Next we define

$$F_k(x) = \int_0^x f_k(t)\,dt \quad \text{for } k \geqslant 1 \quad \text{and} \quad F(x) = \sum_{k\geqslant 1} F_k(x).$$

Since f_k is continuous, F_k is differentiable with $F_k'(x) = f_k(x)$. Also each $F_k = 0$ for $x \notin J_k$, and $\|F_k\|_\infty = F_k\left(\frac{c_k+d_k}{2}\right) = \frac{1}{\pi}|J_k| = \frac{1}{\pi}|I_k|^2$. The sum of these values converges, and hence the series $G_n = \sum_{k=1}^{n} F_k(x)$ converges uniformly to F. Indeed, for any x, there is at most one f so that $F_k(x) \neq 0$.

Finally we show that F is differentiable. On each interval $I_n = (a_k, b_k)$, we have $F(x) = F_k(x)$. This is differentiable and

$$F'(x) = F_k'(x) = f_k(x) = f(x) \quad \text{for} \quad x \in I_k.$$

Now consider $x \in C$. Then $F(x) = 0$. For $y \neq x$, there are two cases. If $y \notin V$, then $F(y) = 0$ and so $\dfrac{F(y) - F(x)}{y - x} = 0$. Otherwise $y \in J_k$ for some k. Since $x \notin I_k$, the interval (x, y) intersects I_k in either (a_k, y) or (y, d_k). In either case, this has length at least

$$c_k - a_k = \frac{1}{2}(|I_k| - |J_k|) = \frac{1}{2}|I_k|(1 - |I_k|) \geqslant \frac{3}{8}|I_k|.$$

Therefore

$$\left|\frac{F(y) - F(x)}{y - x}\right| \leqslant \frac{|J_k|/\pi}{3|I_k|/8} = \frac{8}{3\pi}|I_k|.$$

As y tends to x, the length of I_k must tend to 0 (since $\frac{3}{8}|I_k| \leqslant |y - x|$). Therefore,

$$F'(x) = \lim_{y \to x} \frac{F(y) - F(x)}{y - x} = 0 = f(x).$$

Hence F is differentiable with bounded derivative f, but f is not Riemann integrable.

Exercises for Chapter 6

1. Let $f(x) = \frac{1}{x}$ on $[a, b]$ where $0 < a < b$. Let

$$\mathcal{P}_n = \{t_i = a(b/a)^{i/n} : 0 \leqslant i \leqslant n\} \quad \text{for} \quad n \geqslant 1.$$

(a) Find $L(f, \mathcal{P}_n)$ and $U(f, \mathcal{P}_n)$.
(b) Show that f is Riemann integrable, and evaluate $\lim_{n\to\infty} U(f, \mathcal{P}_n)$.

2. Evaluate the area of a sector of a parabola using Riemann sums.

3. Let $f(x) = \begin{cases} 1 & \text{for} \quad 0 \leqslant x \leqslant 1 \\ 2 & \text{for} \quad 1 < x \leqslant e \\ 3 & \text{for} \quad e < x \leqslant \pi \\ 4 & \text{for} \quad \pi < x \leqslant 4. \end{cases}$

(a) Verify Riemann's condition.
(b) Given $\varepsilon > 0$, find an *explicit value* for $\delta > 0$ so that every partition $\mathcal{P}$ with $\text{mesh}(\mathcal{P}) < \delta$ satisfies Riemann's condition for ε.

4. Prove Proposition 6.3.1(3): if $f(x) \leqslant g(x)$ are integrable on $[a, b]$, then

$$\int_a^b f(x)\,dx \leqslant \int_x^b g(x)\,dx.$$

5. Prove Proposition 6.3.1(43): if $f(x)$ is Riemann integrable on $[a, b]$ and $a < c < b$, then f is Riemann integrable on $[a, c]$.

6. Let $f(x) = \sin(\frac{1}{x})$ for $x \neq 0$ and $f(0) = 0$. Prove that f is Riemann integrable on $[0, 1]$. HINT: verify Riemann's condition by using a partition with a small t_1 and using the continuity of f on $[t_1, 1]$.

7. Suppose that $f(x)$ and $g(x)$ are continuous, monotone increasing functions on $[a, b]$. Prove that

$$\frac{1}{b-a}\int_a^b f(x)\,dx\,\frac{1}{b-a}\int_a^b g(x)\,dx \leqslant \frac{1}{b-a}\int_a^b f(x)g(x)\,dx.$$

HINT: integrate $\big(f(x) - f(c)\big)\big(g(x) - g(c)\big)$ for a certain value of c.

8. Evaluate $\lim\limits_{n\to\infty} \sum_{k=1}^{n} \dfrac{n^5 + 4kn^4 + 9k^2n^3 + 16k^3n^2 + 25k^4n + 36k^5}{n^6}$.
HINT: why is this in this chapter?

9. Let $f(x)$ be a continuous strictly increasing function from $[a, b]$ onto $[c, d]$.
(a) Let $\mathcal{P} = \{a = t_0 < \cdots < t_n = b\}$ be a partition of $[a, b]$, and define a partition $\mathcal{P}'$ of $[c, d]$ by $t_i' = f(t_i)$ for $0 \leqslant i \leqslant n$. Show that

$$L(f, \mathcal{P}) + U(f^{-1}, \mathcal{P}') = bd - ac.$$

HINT: see figure.
(b) Hence show that for any continuous strictly monotone function $f(x)$

$$\int_c^d f^{-1}(x)\,dx = df^{-1}(d) - cf^{-1}(c) - \int_{f^{-1}(c)}^{f^{-1}(d)} f(x)\,dx.$$

(c) For any positive real numbers p, q, show that

$$\int_0^1 (1 - x^p)^{1/q}\,dx = \int_0^1 (1 - x^q)^{1/p}\,dx.$$

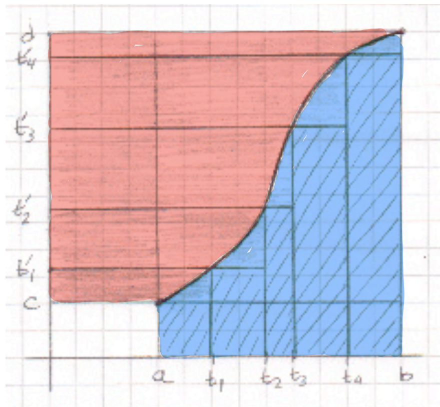


FIGURE 6.3. Exercise (9)

CHAPTER 7

Techniques of Integration

7.1. Simple observations

Here are a few simple techniques to get us started.

7.1.1 Recognizing a derivative. Sometimes you can see an antiderivative by recognizing the integrand as a derivative. For example

$$\int_0^1 e^{x^2} x\,dx = \frac{1}{2}e^{x^2}\Big|_0^1 = \frac{e-1}{2}.$$

$$\int_{-1}^1 \frac{dx}{1+x^2} = \tan^{-1}(x)\Big|_{-1}^1 = \frac{\pi}{4}+\frac{\pi}{4} = \frac{\pi}{2}.$$

$$\int_0^{\pi/4} \sec x \tan x\,dx = \sec x\Big|_0^{\pi/4} = \sqrt{2}-1.$$

7.1.2 Recognizing symmetry. If the integrand is even or odd or periodic, there may be a simplification which is helpful. For example,

$$\int_{-1}^1 e^{x^4}\sin x\,dx = 0 \quad \text{because the integrand is an odd function.}$$

$$\int_{-1}^1 \frac{dx}{1+x^2} = 2\int_0^1 \frac{dx}{1+x^2} \quad \text{because the integrand is an even function.}$$

The following function is 2π periodic, which explains the first two equalities, and it is odd, which explains the last.

$$\int_0^{10\pi} \sin^3 x\,dx = 5\int_0^{2\pi} \sin^3 x\,dx = 5\int_{-\pi}^{\pi} \sin^3 x\,dx = 0$$

7.1.3 The chain rule and FTC. To differentiate $f(x) = \int_{x^2}^{x^3} e^{t^2}\,dt$, let us define $F(x) = \int_0^x e^{t^2}\,dt$. By FTC, $F'(x) = e^{x^2}$. Notice that $f(x) = F(x^3) - F(x^2)$. Therefore by the Chain Rule,

$$f'(x) = F'(x^3)3x^2 - F'(x^2)2x = 3x^2 e^{x^6} - 2xe^{x^4}.$$

7.2. Integration by Parts

We start with the product rule for derivatives:

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

Therefore

$$\begin{aligned}\int_a^b f'(x)g(x)\,dx &= \int_a^b (fg)'(x)\,dx - \int_a^b f(x)g'(x)\,dx \\ &= f(x)g(x)\Big|_a^b - \int_a^b f(x)g'(x)\,dx.\end{aligned}$$

Using this method involves recognizing a derivative as a factor of the integrand.

7.2.1. Example. In the following, $f'(x) = x$ and $g(x) = \tan^{-1}(x)$. Then $f(x) = \frac{x^2}{2}$ and $g'(x) = \frac{1}{1+x^2}$.

$$\begin{aligned}\int_0^1 x\tan^{-1}(x)\,dx &= \frac{x^2}{2}\tan^{-1}(x)\Big|_0^1 - \int_0^1 \frac{x^2}{2}\frac{1}{1+x^2}\,dx \\ &= \frac{x^2}{2}\tan^{-1}(x)\Big|_0^1 - \frac{1}{2}\int_0^1 1 - \frac{1}{1+x^2}\,dx \\ &= \frac{x^2}{2}\tan^{-1}(x)\Big|_0^1 - \frac{x - \tan^{-1}(x)}{2}\Big|_0^1 \\ &= \frac{x^2+1}{2}\tan^{-1}(x) - \frac{x}{2}\Big|_0^1 = \frac{\pi}{4} - \frac{1}{2}.\end{aligned}$$

7.2.2. Example. This works for indefinite integrals as well.

$$\begin{aligned}\int xe^{2x}\,dx &= x\left(\tfrac{1}{2}e^{2x}\right) - \int 1\left(\tfrac{1}{2}e^{2x}\right)dx \\ &= \tfrac{1}{2}xe^{2x} - \tfrac{1}{4}e^{2x} + c = \frac{2x-1}{4}e^{2x} + c.\end{aligned}$$

7.2.3. EXAMPLE. If $a \neq -1$, we integrate x^a and differentiate $\ln x$.

$$\int x^a \ln x\,dx = \frac{x^{a+1}}{a+1}\ln x - \int \frac{x^{a+1}}{a+1}\frac{1}{x}\,dx = \frac{x^{a+1}}{a+1}\ln x - \frac{1}{a+1}\int x^a\,dx$$
$$= \frac{x^{a+1}}{a+1}\ln x - \frac{x^{a+1}}{(a+1)^2} + c.$$

When $a = -1$, we observe that $\frac{1}{x}$ is the derivative of $\ln x$. Therefore we observe that $\dfrac{d}{dx}(\ln x)^2 = \dfrac{2}{x}\ln x$. Hence

$$\int x^{-1}\ln x\,dx = \tfrac{1}{2}(\ln x)^2 + c.$$

This is best handled by the substitution method.

7.2.4. EXAMPLE. What is wrong with the following argument?

$$\int \frac{1}{x}\,dx = \int 1\cdot\frac{1}{x}\,dx = x\,\frac{1}{x} - \int x\Big(\frac{-1}{x^2}\Big)\,dx = 1 + \int \frac{1}{x}\,dx.$$

Therefore $0 = 1$.

7.3. Integration by Substitution

This is a fundamental method which will get a great deal of use. It can be a bit tricky to use because the limits of integration change. In a sense, this is the integral version of the chain rule.

7.3.1. PROPOSITION. *Let $f(x)$ be continuous on $[a,b]$. Suppose $u : [p,q] \to [a,b]$ is a C^1 function. Then*

$$\int_p^q f(u(x))u'(x)\,dx = \int_{u(p)}^{u(q)} f(t)\,dt.$$

PROOF. Let $F(y) = \displaystyle\int_a^y f(t)\,dt$. Since u maps $[p,q]$ into $[a,b]$, we can define $G(x) = F(u(x))$ for $x \in [p,q]$. By the chain rule and FTC,

$$G'(x) = F'(u(x))u'(x) = f(u(x))u'(x).$$

Therefore

$$\int_p^q f(u(x))u'(x)\,dx = \int_p^q G'(x)\,dx = G(x)\Big|_p^q$$
$$= F(u(q)) - F(u(p))$$

$$= F(y)\Big|_{u(p)}^{u(q)} = \int_{u(p)}^{u(q)} f(t)\,dt. \qquad \blacksquare$$

7.3.2. REMARKS.

(1) If u is monotone, which is frequently the case in practice, then it suffices to check that $a \leqslant u(p), u(q) \leqslant b$ to ensure that $u([p,q]) \subset [a,b]$.

(2) If we set $y = u(x)$, then $y' = u'(x)$. So *formally* we get

$$dy = \frac{dy}{dx}\,dx = u'(x)\,dx.$$

Thus we *substitute* y for $u(x)$ and replace $u'(x)\,dx$ by dy to get

$$\int_p^q f(u(x))u'(x)\,dx = \int_p^q f(y)y'(x)\,dx = \int_{u(p)}^{u(q)} f(y)\,dy.$$

The limits of integration change becasue we are now integrating with respect to y. As x runs from p to q, $y = u(x)$ runs from $u(p)$ to $u(q)$.

(3) If u is monotone decreasing, then $u(p) > u(q)$, so that

$$\int_p^q f(u(x))u'(x)\,dx = \int_{u(p)}^{u(q)} f(y)\,dy = -\int_{u(q)}^{u(p)} f(y)\,dy.$$

(4) Often substitution works in the other direction. That is, we are considering $\int_a^b f(t)\,dt$ and we substitute $t = u(x)$. To do this, we need u to be monotone, C^1, and contain $[a,b]$ in its range. We set $t = u(x)$ and hence $dt = u'(x)\,dx$. The new limits will be $u^{-1}(a) =: p$ and $u^{-1}(b) =: q$. That is,

$$\int_a^b f(t)\,dt = \int_{u^{-1}(a)}^{u^{-1}(b)} f(u(x))u'(x)\,dx.$$

7.3.3. EXAMPLE. Consider $\displaystyle\int_0^{\pi/4} \tan x\,dx = \int_0^{\pi/4} \frac{\sin x}{\cos x}\,dx.$

We recognize $-\sin x$ as the derivative of $\cos x$. So we substitute $u = \cos x$ and $du = u'(x)\,dx = -\sin x\,dx$. Thus the integral becomes

$$\begin{aligned}\int_0^{\pi/4} \frac{\sin x}{\cos x}\,dx &= \int_0^{\pi/4} \frac{-1}{\cos x}(-\sin x)\,dx \\ &= \int_{\cos(0)}^{\cos(\pi/4)} \frac{-du}{u} = -\ln|u|\Big|_1^{1/\sqrt{2}} \\ &= -\ln\left|\frac{1}{\sqrt{2}}\right| + \ln|1| = \frac{\ln 2}{2}.\end{aligned}$$

7.3.4. EXAMPLE. Consider $\displaystyle\int_{r/2}^{\sqrt{3}r/2} \frac{1}{\sqrt{r^2-t^2}}\,dt.$

Substitute $t = r\sin x$. This will greatly simplify the quantity under the square root. Then $dt = r\cos x\,dx$. Then $u(x) = r\sin x$, so that $p = u^{-1}(r/2) = \frac{\pi}{6}$ and $q = u^{-1}(\sqrt{3}r/2) = \frac{\pi}{3}$. Thus

$$\begin{aligned}\int_{r/2}^{\sqrt{3}r/2} \frac{1}{\sqrt{r^2-t^2}}\,dt &= \int_{\pi/6}^{\pi/3} \frac{1}{\sqrt{r^2-r^2\sin^2 x}} r\cos x\,dx \\ &= \int_{\pi/6}^{\pi/3} \frac{1}{r\cos x} r\cos x\,dx \\ &= \int_{\pi/6}^{\pi/3} dx = x\Big|_{\pi/6}^{\pi/3} = \frac{\pi}{6}.\end{aligned}$$

For an indefinite integral, one must convert back to the original variable

$$\int \frac{1}{\sqrt{r^2-t^2}}\,dt = x = \sin^{-1}\left(\frac{t}{r}\right) + c.$$

This can be verified by differentiation.

7.3.5. EXAMPLE. Consider $\displaystyle\int \frac{dx}{a^2\sin^2 x + b^2\cos^2 x}.$

First we manipulate this to put it into a more useful form. Multiply by $\dfrac{\sec^2 x}{\sec^2 x}$ to get

$$\int \frac{\sec^2 x}{a^2\tan^2 x + b^2}\,dx = \frac{1}{b^2}\int \frac{\sec^2 x}{\left(\frac{a}{b}\tan x\right)^2 + 1}\,dx$$

Substitute $u = \frac{a}{b}\tan x$. Then $du = \frac{a}{b}\sec^2 x\,dx$. We obtain

$$\begin{aligned}\frac{1}{b^2}\frac{b}{a}\int \frac{\frac{a}{b}\sec^2 x\,dx}{\left(\frac{a}{b}\tan x\right)^2+1} &= \frac{1}{ab}\int \frac{du}{u^2+1} \\ &= \frac{1}{ab}\tan^{-1}(u) + c = \frac{1}{ab}\tan^{-1}\left(\frac{a}{b}\tan x\right) + c.\end{aligned}$$

7.3.6. EXAMPLE. Consider $\displaystyle\int \sec x\,dx.$

This requires a clever trick using the fact that $\frac{d}{dx}\tan x = \sec^2 x$ and $\frac{d}{dx}\sec x = \sec x\tan x$.

$$\int \sec x\,dx = \int \sec x \frac{\sec x + \tan x}{\sec x + \tan x}\,dx = \int \frac{\sec^2 x + \sec x\tan x}{\sec x + \tan x}\,dx.$$

So the numerator is the derivative of $u = \sec x + \tan x$. Making this substitution, we get

$$\begin{aligned}\int \sec x\,dx &= \int \frac{\sec^2 x + \sec x \tan x}{\sec x + \tan x}\,dx \\ &= \int \frac{u'(x)}{u(x)}\,dx = \int \frac{du}{u} \\ &= \ln|u| + c = \ln\big|\sec x + \tan x\big| + c.\end{aligned}$$

Similarly, $\int \csc x\,dx = \ln\big|\csc x - \cot x\big| + c$. It might be worthwhile remembering these two integrals.

7.3.7. EXAMPLE. What can go wrong if we ignore the range of u? Let $f(x) = \dfrac{1}{(x+1)^2}$ for $x \neq -1$. Set $u(x) = x^2 - 2x$. Then $u(0) = 0$ and $u(4) = 8$. Compare

$$\int_0^4 f(u(x))u'(x)\,dx \quad \text{and} \quad \int_0^8 f(y)\,dy.$$

The integral on the right is

$$\int_0^8 \frac{1}{(y+1)^2}\,dy = \frac{-1}{y+1}\bigg|_0^8 = -\frac{1}{9} + 1 = \frac{8}{9}.$$

Since $u'(x) = 2x - 2$, the integral on the left is

$$\int_0^4 f(u(x))u'(x)\,dx = \int_0^4 \frac{2x-2}{(x^2-2x+1)^2}\,dx = \int_0^4 \frac{2}{(x-1)^3}\,dx.$$

This is not integrable because the integrand blows up at $x = 1$. The problem is that $u([0,4]) = [-1,8]$, not $[0,8]$.

7.3.8. EXAMPLE. Sometimes we combine our two new techniques. Consider $\int_0^1 e^{\sin^{-1}(x)}\,dx$. The substitution $x = \sin u$ jumps out. So $\sin^{-1}(\sin u) = u$. Then $dx = \cos u\,du$. Note that u runs from 0 to $\frac{\pi}{2}$. We substitute and then integrate by parts twice, integrating e^u each time.

$$\begin{aligned}\int_0^1 e^{\sin^{-1}(x)}\,dx &= \int_0^{\pi/2} e^u \cos u\,du \\ &= e^u \cos u\Big|_0^{\pi/2} - \int_0^{\pi/2} e^u(-\sin u)\,du \\ &= e^u \cos u + e^u \sin u\Big|_0^{\pi/2} - \int_0^{\pi/2} e^u \cos u\,du.\end{aligned}$$

Therefore

$$2\int_0^{\pi/2} e^u \cos u\, du = e^u(\cos u + \sin u)\Big|_0^{\pi/2} = e^{\pi/2} - 1.$$

Thus

$$\int_0^1 e^{\sin^{-1}(x)}\, dx = \frac{e^{\pi/2} - 1}{2}.$$

If we solve $\cos u = \sqrt{1 - \sin^2 u} = \sqrt{1 - x^2}$, we get

$$\int e^{\sin^{-1}(x)}\, dx = \tfrac{1}{2}e^u(\cos u + \sin u) + c = e^{\sin^{-1}(x)}\Big(\frac{\sqrt{1 - x^2} + x}{2}\Big) + c$$

This can be verified by differentiation.

7.4. Integral Recursion Formulae

Sometimes a method can be repeated by induction to get a whole family of integral formulas. These formulas are called *recursion formulae*. (There are two acceptable plurals for formula, formulas and formulae, which comes from latin. As time goes on, formulas is becoming much more common.)

7.4.1. EXAMPLE. Let $I_n = \displaystyle\int x^n e^x\, dx$ for integers $n \geqslant 0$. We know that $I_0 = e^x + c$. Using integration by parts, we get

$$I_n = \int x^n e^x\, dx = x^n e^x - \int n x^{n-1} e^x\, dx = x^n e^x - nI_{n-1}.$$

Hence

$$\begin{gathered} I_1 = (x-1)e^x + c \\ I_2 = x^2 e^x - 2((x-1)e^x + c) = (x^2 - 2x + 2)e^x + c \\ I_3 = x^3 - 3I_2 = (x^3 - 3x^2 + 6x - 6)e^x + c \end{gathered}$$

Note that since c is an arbitrary constant, it does not change when added or multiplied. By observation, we detect the pattern as

$$\begin{aligned} I_n &= \big(x^n - nx^{n-1} + n(n-1)x^{n-2} - \cdots + (-1)^n n!\big)e^x + c \\ &= \sum_{j=0}^{n} (-1)^j n(n-1)\cdots(n+1-j)x^{n-j} = \sum_{j=0}^{n} (-1)^j \frac{n!}{(n-j)!} x^{n-j}. \end{aligned}$$

The last formula an be verified by induction. It is true for $n = 0, 1, 2, 3$ as shown. Suppose that it holds for n. Then

$$\begin{aligned} I_{n+1} &= x^{n+1}e^x - (n+1)I_n = x^{n+1}e^x - (n+1)\sum_{j=0}^{n}(-1)^j \frac{n!}{(n-j)!}x^{n-j} \\ &= x^{n+1}e^x - \sum_{j=0}^{n}(-1)^j \frac{(n+1)!}{(n-j)!}x^{n-j} \end{aligned}$$

set $k = j + 1$,

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

So the formula holds for all $n \geqslant 0$.

7.4.2. Example. Let $I_n = \int \sin^n x\,dx$ for integers $n \geqslant 0$. Then $I_0 = x + c$ and $I_1 = -\cos x + c$. Again we integrate by parts

$$\begin{aligned} I_n &= \int \sin^n x\,dx = \int \sin^{n-1} x \sin x\,dx \\ &= -\sin^{n-1} x \cos x + \int (n-1)\sin^{n-2}\cos^2 x\,dx \\ &= -\sin^{n-1} x \cos x + (n-1)\int \sin^{n-2}(1 - \sin^2 x)\,dx \\ &= -\sin^{n-1} x \cos x + (n-1)(I_{n-2} - I_n). \end{aligned}$$

Solving, we get

$$I_n = -\frac{1}{n}\sin^{n-1} x \cos x + \frac{n-1}{n} I_{n-2}.$$

Now let $a_n = \int_0^{\pi/2} \sin^n x\,dx$ Then $a_0 = \frac{\pi}{2}$ and $a_1 = 1$. The recursion formula shows that for $n \geqslant 2$,

$$a_n = \frac{n-1}{n} a_{n-2}.$$

We iterate this formula:

$$\begin{aligned} a_{2n} &= \frac{2n-1}{2n} a_{2n-2} = \frac{2n-1}{2n}\frac{2n-3}{2n-2} a_{2n-4} = \dots \\ &= \frac{2n-1}{2n}\;\frac{2n-3}{2n-2}\;\cdots\;\frac{3}{4}\;\frac{1}{2}\;a_0 \\ &= \frac{(2n)!}{(2n)^2(2n-2)^2\cdots 4^2 2^2}\;\frac{\pi}{2} = \frac{(2n)!}{2^{2n}(n!)^2}\;\frac{\pi}{2}. \end{aligned}$$

Similarly

$$\begin{aligned} a_{2n+1} &= \frac{2n}{2n+1} a_{2n-1} = \frac{2n}{2n+1}\frac{2n-2}{2n-1} a_{2n-3} = \dots \\ &= \frac{2n}{2n+1}\frac{2n-2}{2n-1} \cdots \frac{4}{5}\frac{2}{3} a_1 \\ &= \frac{(2n)^2(2n-2)^2 \cdots 4^2 2^2}{(2n+1)!} = \frac{2^{2n}(n!)^2}{(2n+1)!}. \end{aligned}$$

These formulas lead to a famous formula for π.

7.4.3. Wallis Product Formula.

$$\frac{\pi}{2} = \lim_{n\to\infty} \frac{2\cdot 2\cdot 4\cdot 4\cdots}{1\cdot 3\cdot 3\cdot 5\cdots} \frac{(2n)}{(2n-1)} \frac{(2n)}{(2n+1)} = \lim_{n\to\infty} \frac{2^{4n}(n!)^4}{(2n)!(2n+1)!}.$$

Proof. Observe that

$$\frac{a_{2n}}{a_{2n+1}} = \frac{1\cdot 3\cdot 3\cdot 5\cdots}{2\cdot 2\cdot 4\cdot 4\cdots} \frac{(2n-1)}{(2n)} \frac{(2n+1)}{(2n)} \frac{\pi}{2}.$$

Hence

$$\frac{\pi}{2} = \frac{2\cdot 2\cdot 4\cdot 4\cdots}{1\cdot 3\cdot 3\cdot 5\cdots} \frac{(2n)}{(2n-1)} \frac{(2n)}{(2n+1)} \frac{a_{2n}}{a_{2n+1}} = \frac{2^{4n}(n!)^4}{(2n)!(2n+1)!} \frac{a_{2n}}{a_{2n+1}}.$$

Thus Wallis's product formula holds if we can show that $\lim_{n\to\infty} \frac{a_{2n}}{a_{2n+1}} = 1$.

Observe that on $[0, \frac{\pi}{2}]$, $0 \leqslant \sin^{2n+1} x \leqslant \sin^{2n} x \leqslant \sin^{2n-1} x$. Therefore $a_{2n+1} \leqslant a_{2n} \leqslant a_{2n-1}$. Also $a_{2n+1} = \frac{2n-1}{2n} a_{2n-1}$. Hence

$$1 \leqslant \frac{a_{2n}}{a_{2n+1}} \leqslant \frac{2n}{2n-1} \frac{a_{2n}}{a_{2n-1}} \leqslant \frac{2n}{2n-1} \to 1.$$

By the Squeeze Theorem, we obtain $\lim_{n\to\infty} \frac{a_{2n}}{a_{2n+1}} = 1$. ■

7.4.4. Example. Let $I_n = \int \frac{dx}{(x^2+a^2)^n}$. So $I_0 = x + c$ and

$$I_1 = \frac{1}{a} \int \frac{dx/a}{\left(\frac{x}{a}\right)^2 + 1} = \frac{1}{a} \tan^{-1}\left(\frac{x}{a}\right) + c.$$

Integrating by parts,

$$\begin{aligned} I_n &= \int (1)\frac{1}{(x^2+a^2)^n}\,dx \\ &= \frac{x}{(x^2+a^2)^n} - \int x\frac{-2nx}{(x^2+a^2)^{n+1}}\,dx \\ &= \frac{x}{(x^2+a^2)^n} + 2n\int \frac{x^2+a^2-a^2}{(x^2+a^2)^{n+1}}\,dx \\ &= \frac{x}{(x^2+a^2)^n} + 2n(I_n - a^2 I_{n+1}). \end{aligned}$$

Therefore

$$I_{n+1} = \frac{x}{2na^2(x^2+a^2)^n} + \frac{2n-1}{2na^2}I_n.$$

For example,

$$\begin{aligned} \int \frac{dx}{(x^2+4)^4} &= \frac{x}{24(x^2+4)^3} + \frac{5}{24}I_3 \\ &= \frac{x}{24(x^2+4)^3} + \frac{5}{24}\Big(\frac{x}{16(x^2+4)^2} + \frac{3}{16}I_2\Big) \\ &= \frac{x}{24(x^2+4)^3} + \frac{5x}{24(16)(x^2+4)^2} + \frac{15}{24(16)}\Big(\frac{x}{8(x^2+4)} + \frac{1}{8}I_1\Big) \\ &= \frac{x}{24(x^2+4)^3} + \frac{5x}{384(x^2+4)^2} + \frac{15x}{3072(x^2+4)} + \frac{15}{6144}\tan^{-1}\Big(\frac{x}{2}\Big) + c. \end{aligned}$$

Hence

$$\begin{aligned} \int_0^1 \frac{dx}{(x^2+4)^4} &= \frac{1}{24(5)^3} + \frac{5}{384(5)^2} + \frac{15}{3072(5)} + \frac{15}{6144}\tan^{-1}(\tfrac{1}{2}) \\ &= \frac{1}{3000} + \frac{1}{1920} + \frac{1}{1024} + \frac{5}{2048}\tan^{-1}(\tfrac{1}{2}). \end{aligned}$$

7.5. Partial Fractions

A rational function is a function of the form $\dfrac{p(x)}{q(x)}$ for two polynomials p and q. There is a systematic method to integrate any rational function. You can always divide q into p to get a polynomial plus a remainder. Thus we can always assume that $\deg p < \deg q$.

Since $q(x)$ is a real polynomial, it factors into a product of linear and irreducible quadratic terms. Say

$$q(x) = (x-a_1)^{m_1}\dots(x-a_d)^{m_d}\big((x-b_1)^2+c_1^2\big)^{n_1}\dots\big((x-b_e)^2+c_e^2\big)^{n_e}.$$

Of course, this factorization is not always readily available.

First we will deal with some important special cases.

Case 1. Suppose that $q(x) = (x-a_1)(x-a_2)\dots(x-a_d)$ has d simple real roots. Then $p(x) \equiv p(a_i) \pmod{x-a_i}$ for $1 \leqslant i \leqslant d$. Let $q_i(x) = \dfrac{q(x)}{x-a_i}$ and define

$$P(x) = \sum_{i=1}^{d} \frac{p(a_i)}{q_i(a_i)} q_i(x).$$

Then $\deg P \leqslant \max\{\deg q_i : 1 \leqslant i \leqslant n\} = d-1 < d = \deg q$. Moreover

$$P(x) \equiv P(a_i) = p(a_i) \pmod{x-a_i} \quad \text{for} \quad 1 \leqslant i \leqslant d.$$

That means $x - a_i$ divides $P(x) - p(x)$ for $1 \leqslant i \leqslant d$. Hence q divides $P - p$. Since $\deg P - p \leqslant d - 1 < \deg q$, we have $P = p$. Therefore

$$\frac{p(x)}{q(x)} = \sum_{i=1}^{d} \frac{p(a_i)}{q_i(a_i)} \frac{q_i(x)}{q(x)} = \sum_{i=1}^{d} \frac{p(a_i)}{q_i(a_i)} \frac{1}{x-a_i}.$$

Thus

$$\int \frac{p(x)}{q(x)}\,dx = \sum_{i=1}^{d} \frac{p(a_i)}{q_i(a_i)} \ln|x-a_i| + c.$$

7.5.1. Example.

$$\begin{aligned}\int \frac{x^3 - x^2 + x + 2}{x^3 - x}\,dx &= \int 1 + \frac{-x^2 + 2x + 2}{(x+1)x(x-1)}\,dx \\ &= \int 1 + \frac{-1/2}{x+1} - \frac{2}{x} + \frac{3/2}{x-1}\,dx \\ &= x - \frac{1}{2}\ln|x+1| - 2\ln|x| + \frac{3}{2}\ln|x-1| + c \\ &= x + \frac{1}{2}\ln\frac{|x-1|^3}{x^4|x+1|} + c.\end{aligned}$$

Here $p(x) = -x^2 + 2x + 2$, and $\frac{p(-1)}{q_1(-1)} = \frac{-1}{(-1)((-2)} = -\frac{1}{2}$, $\frac{p(0)}{q_2(0)} = \frac{2}{1(-1)} = -2$ and $\frac{p(1)}{q_3(1)} = \frac{3}{(2)(1)} = \frac{3}{2}$.

Case 2. Suppose that $q(x) = (x-a)^d$, and suppose that $\deg p < d$. Write $p(x) = \sum_{k=0}^{d-1} b_k (x-a)^k$. It is easy to see that $\{(x-a)^k : 0 \leqslant k < d\}$ is a basis for the space of polynomials of degree at most $d-1$. Thus the coefficients b_k are uniquely determined. In fact, if we compute the derivatives

$$p^{(i)}(a) = \sum_{k=i}^{d-1} b_k k(k-1)\cdots(k+1-i)(x-a)^{k-i}\Big|_{x=a} = b_i i!.$$

Hence $b_i = \frac{p^{(i)}(a)}{i!}$.

7.5.2. Example. Consider $\int \frac{x^3+x^2-x-1}{(x-2)^4}\,dx$. Then

$$\begin{aligned}
p(x) &= x^3+x^2-x-1(x-2)^2 & b_0 &= \frac{p(2)}{0!} = 9\\
p'(x) &= 3x^2+2x-1 & b_1 &= \frac{p'(2)}{1!} = 15\\
p''(x) &= 6x+2 & b_2 &= \frac{p''(2)}{2!} = 7\\
p^{(3)}(x) &= 6 & b_3 &= \frac{p^{(3)}(2)}{3!} = 1.
\end{aligned}$$

Therefore

$$\begin{aligned}
\int \frac{x^3+x^2-x-1}{(x-2)^4}\,dx &= \int \frac{9}{(x-2)^4} + \frac{15}{(x-2)^3} + \frac{7}{(x-2)^2} + \frac{1}{x-2}\,dx\\
&= -3(x-2)^{-3} - \frac{15}{2}(x-2)^{-2} - 7(x-2)^{-1} + \ln|x-2| + c.
\end{aligned}$$

Case 3. Suppose that $q(x) = \big((x-a)^2+c^2\big)^k$, and that $\deg p < d = 2k$. Then we can write $p(x) = \sum_{i=0}^{k-1} b_i(x)\big((x-a)^2+c^2\big)^i$, where each $b_i(x) = d_i + e_i x$ are linear, in a unique way. It isn't quite as easy as Case 2, but we achieve this by repeated division of $p(x)$ by $r(x) = (x-a)^2+c^2$ and getting the remainders. For example, consider $q(x) = (x^2+2x+5)^3$ and $p(x) = 3x^5-2x^4+7x^3+x^2-1$. Division by x^2+2x+5 yields

$$\begin{aligned}
3x^5-2x^4+7x^3+x^2-1 &= (3x^3-8x^2+8x+25)(x^2+2x+5)-10x-126\\
3x^3-8x^2+8x+25 &= (3x-14)(x^2+2x+5)+21x+45
\end{aligned}$$

Hence

$$3x^5-2x^4+7x^3+x^2-1 = (3x-14)(x^2+2x+5)^2+(21x+45)(x^2+2x+5)-10x-26.$$

Therefore

$$\int \frac{3x^5-2x^4+7x^3+x^2-1}{(x^2+2x+5)^3}\,dx = \int \frac{3x-14}{x^2+2x+5} + \frac{21x+45}{(x^2+2x+5)^2} - \frac{10x+126}{(x^2+2x+5)^3}\,dx.$$

Now we discuss how to integrate $\int \frac{rx+s}{\big((x-a)^2+c^2\big)^k}\,dx$ for $c>0$. We can substitute $cu = x-a$ and $c\,du = dx$ to get

$$\int \frac{cru+ar+s}{\big(c^2u^2+c^2\big)^k}\,c\,du = c^{2-2k}r\int \frac{u}{(u^2+1)^k}\,du + c^{1-2k}(ar+s)\int \frac{du}{(u^2+1)^k}\,du.$$

Now

$$\int \frac{u}{(u^2+1)^k}\,du = \begin{cases} \frac{-1}{2(k-1)}(u^2+1)^{1-k} + c & \text{if } k \geqslant 2 \\ \frac{1}{2}\ln(u^2+1) + c & \text{if } k = 1 \end{cases}$$

Finally for $I_k = \displaystyle\int \frac{1}{(u^2+1)^k}\,du$, use recursion. We know $I_1 = \tan^{-1} u + c$.

$$\begin{aligned} I_k &= \int \frac{1}{(u^2+1)^k}(1)\,du \\ &= \frac{u}{(u^2+1)^k} - \int \frac{-2ku^2}{(u^2+1)^{k+1}}\,du \\ &= \frac{u}{(u^2+1)^k} + 2k\int \frac{u^2+1-1}{(u^2+1)^{k+1}}\,du \\ &= \frac{u}{(u^2+1)^k} + 2k(I_k - I_{k+1}) \end{aligned}$$

Therefore $I_{k+1} = \dfrac{u}{2k(u^2+1)^k} + \frac{2k-1}{2k} I_k$.

7.5.3. Example. Here is example of Case 3 where we use our knowledge of the answer to find unknown coefficients. The analysis of case 3 shows that there are constants a, b, c, d so that

$$\int \frac{2x^3 - 7x^2 + 5}{(2x^2+3x+2)^2}\,dx = \frac{ax+b}{2x^2+3x+2} + \int \frac{cx+d}{2x^2+3x+2}\,dx.$$

Differentiate to get

$$\frac{2x^3-7x^2+5}{(2x^2+3x+2)^2} = \frac{a(2x^2+3x+2) - (ax+b)(4x+3)}{(2x^2+3x+2)^2} + \frac{cx+d}{2x^2+3x+2}$$

Therefore

$$\begin{aligned} &2x^3 - 7x^2 + 5 \\ &= 2ax^2 + 3ax + 2a - 4ax^2 - 3ax - 4bx - 3b + 2cx^3 + 3cx^2 + 2cx + 2dx^2 + 3dx + 2d \\ &= 2cx^3 + (-2a+3c+2d)x^2 + (-4b+2c+3d)x + (2a-3b+2d). \end{aligned}$$

Thus

$$\begin{aligned} 2c &= 2 \\ -2a + 3c + 2d &= -7 \\ -4b + 2c + 3d &= 0 \\ 2a - 3b + 2d &= 5. \end{aligned}$$

The coefficient of x^3 shows that $c = 1$. Setting $x = 1$ (or adding the four equations) yields $-7b + 7 + 7d = 0$, so $b = d + 1$. Setting $x = -1$ yields $b - 1 + d = -4$, so

$d = -2$ and $b = -1$. Hence $a = 3$. Therefore

$$\int \frac{2x^3 - 7x^2 + 5}{(2x^2 + 3x + 2)^2}\,dx = \frac{3x - 1}{2x^2 + 3x + 2} + \int \frac{x - 2}{2x^2 + 3x + 2}\,dx$$
$$= \frac{3x - 1}{2x^2 + 3x + 2} + \frac{1}{4}\int \frac{4x + 3}{2x^2 + 3x + 2}\,dx - \frac{11}{8}\int \frac{1}{(x + \frac{3}{4})^2 + \frac{7}{16}}\,dx$$
$$= \frac{3x - 1}{2x^2 + 3x + 2} + \frac{1}{4}\ln|2x^2 + 3x + 2| + \frac{22}{7}\int \frac{1}{(\frac{4}{\sqrt{7}}x + \frac{3}{\sqrt{7}})^2 + 1}\,dx$$
$$= \frac{3x - 1}{2x^2 + 3x + 2} + \frac{1}{4}\ln|2x^2 + 3x + 2| + \frac{11}{2\sqrt{7}}\tan^{-1}(\tfrac{4}{\sqrt{7}}x + \tfrac{3}{\sqrt{7}}) + c.$$

We now show how to reduce the general case to cases 2 and 3. This method is known as *partial fractions*. The proof uses some polynomial algebra from Math 145.

7.5.4. THEOREM. *Let*

$$q(x) = (x - a_1)^{m_1} \dots (x - a_d)^{m_d}\big((x - b_1)^2 + c_1^2\big)^{n_1} \dots \big((x - b_e)^2 + c_e^2\big)^{n_e},$$

where $c_j > 0$, *and let* $p(x)$ *be a polynomial with* $\deg p < \deg q$. *Then there are unique polynomials* $p_i(x)$ *with* $\deg p_i < m_i$ *for* $1 \leqslant i \leqslant d$ *and* $\deg p_{d+j} < 2n_j$ *for* $1 \leqslant j \leqslant e$ *so that*

$$\frac{p(x)}{q(x)} = \sum_{i=1}^{d} \frac{p_i(x)}{(x - a_i)^{m_i}} + \sum_{j=1}^{s} \frac{p_{d+j}(x)}{((x - b_j)^2 + c_j^2)^{n_j}}.$$

PROOF. Let $q_i(x) = (x - a_i)^{m_i}$ for $1 \leqslant i \leqslant d$ and $q_{d+j}(x) = \big((x - b_j)^2 + c_j^2\big)^{n_j}$ for $1 \leqslant j \leqslant s$. Then q_i and q_j have no common factor for $i \neq j$, and thus $\gcd(q_i, q_j) = 1$. There are unique polynomials $r_i(x)$ so that

$$p(x) \equiv r_i(x) \pmod{q_i(x)} \quad \text{and} \quad \deg r_i < \deg q_i \quad \text{for} \quad 1 \leqslant i \leqslant d + e.$$

Let $Q_i(x) = \dfrac{q(x)}{q_i(x)}$, so that $q(x) = q_i(x)Q_i(x)$.

Suppose that there are polynomials p_i with $\deg p_i < \deg q_i$ so that

$$\frac{p(x)}{q(x)} = \sum_{i=1}^{d+e} \frac{p_i(x)}{q_i(x)} = \frac{1}{q(x)} \sum_{i=1}^{d+e} p_i(x)Q_i(x).$$

Then $p(x) = \sum_{i=1}^{d+e} p_i(x)Q_i(x)$. Since Q_j is a multiple of q_i when $j \neq i$, we have

$$r_i \equiv p \equiv p_i Q_i \pmod{q_i}.$$

Now r_i and Q_i are known, and since $\gcd(Q_i, q_i) = 1$, this congruence equation has a unique solution $p_i \pmod{q_i}$. In particular, there is a unique polynomial p_i with $\deg p_i < \deg q_i$ in this congruence class.

Define $P(x) = \sum_{i=1}^{d+e} p_i(x)Q_i(x)$. Then

$$\deg P \leqslant \max\{\deg p_i + \deg Q_i\} < \deg q.$$

Observe that by construction,

$$P(x) \equiv p_i Q_i \equiv r_i \equiv p(x) \pmod{q_i} \quad \text{for} \quad 1 \leqslant i \leqslant d+e.$$

Hence q_i divides $P - p$. Since the factors q_i are relatively prime, their product q divides $P - p$. However $\deg(P - p) < \deg q$, which forces $P = p$. Therefore $p(x)$ has the desired decomposition. ■

Surprisingly one does not need to solve the congruences to find the decomposition. Instead, one can replace the coefficients by variables and simplify, or plug in some values of x, to get enough linear equations to determine them. Solving linear equations is easily done on a computer, and for small systems, easily done by hand.

7.5.5. EXAMPLE. Consider $\displaystyle\int \frac{dx}{x^4+1}$. The tricky part is factoring $x^4 + 1$ into two quadratics. It has no real roots because it is strictly positive. It isn't obvious, but it is the difference of two perfect squares.

$$x^4 + 1 = (x^4 + 2x^2 + 1) - 2x^2 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1).$$

Thus $\dfrac{1}{x^4+1}$ has a partial fraction decomposition of the form

$$\begin{aligned}\frac{1}{x^4+1} &= \frac{ax+b}{x^2+\sqrt{2}x+1} + \frac{cx+d}{x^2-\sqrt{2}x+1} \\ &= \frac{(a+c)x^3 + (-\sqrt{2}a+b+\sqrt{2}c+d)x^2 + (a-\sqrt{2}b+c+\sqrt{2}d)x+(b+d)}{x^4+1}\end{aligned}$$

So

$$\begin{aligned} a + c &= 0 \\ -\sqrt{2}a + b + \sqrt{2}c + d &= 0 \\ a - \sqrt{2}b + c + \sqrt{2}d &= 0 \\ b + d &= 1 \end{aligned}$$

Thus $c = -a$ and the third equation yields $b = d$, so $b = d = \frac{1}{2}$ by the fourth. Finally the second shows that $2\sqrt{2}a = 1$, so $a = \frac{1}{2\sqrt{2}} = -c$.

Therefore

$$\int \frac{dx}{x^4+1} = \int \frac{\frac{1}{2\sqrt{2}}x + \frac{1}{2}}{x^2 + \sqrt{2}x + 1} + \frac{-\frac{1}{2\sqrt{2}}x + \frac{1}{2}}{x^2 - \sqrt{2}x + 1}\, dx$$

$$= \int \frac{1}{4\sqrt{2}} \frac{2x+\sqrt{2}}{x^2+\sqrt{2}x+1} + \frac{1}{2}\frac{1}{(\sqrt{2}x+1)^2+1} - \frac{1}{4\sqrt{2}}\frac{2x-\sqrt{2}}{x^2-\sqrt{2}x+1} + \frac{1}{2}\frac{1}{(\sqrt{2}x-1)^2+1}\, dx$$

$$= \frac{1}{4\sqrt{2}} \ln\left|\frac{x^2+\sqrt{2}x+1}{x^2-\sqrt{2}x+1}\right| + \frac{1}{2\sqrt{2}}\tan^{-1}(\sqrt{2}x+1) + \frac{1}{2\sqrt{2}}\tan^{-1}(\sqrt{2}x-1) + c$$

7.6. Rationalization Tricks

One of the reason integrating rational functions by partial fractions is important is that there are methods to reduce other more complicated integrals to the rational situation. We will look at two such techniques.

7.6.1 $\int R(\sin x, \cos x)\, dx$ where R is rational.

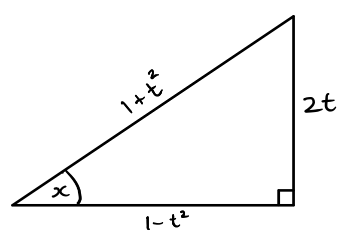
Making the substitution $t = \tan\frac{x}{2}$ *always* converts this to the integral of a rational function. We have the following identities, all of which are rational functions of t:

$$x = 2\tan^{-1} t \quad \text{thus} \quad dx = \frac{2}{1+t^2}dt$$

$$\tan x = \frac{2t}{1-t^2}$$

$$\sin x = \frac{2t}{1+t^2}$$

$$\cos x = \frac{1-t^2}{1+t^2}.$$



7.6.2. EXAMPLE. Make this substitution in the following integral.

$$\int \frac{dx}{\sin x(2+\cos x-2\sin x)} = \int \frac{1}{\frac{2t}{1+t^2}\left(\frac{2+2t^2+1-t^2-4t}{1+t^2}\right)} \frac{2}{1+t^2}\, dt$$

$$= \int \frac{1+t^2}{t(t^2-4t+3)}\, dt = \int \frac{1/3}{t} + \frac{5/3}{t-3} - \frac{1}{t-1}\, dt$$

$$= \tfrac{1}{3}\ln|t| + \tfrac{5}{3}\ln|t-3| - \ln|t-1| + c$$

$$= \tfrac{1}{3}\ln\left|\tan\tfrac{x}{2}\right| + \tfrac{5}{3}\ln\left|\tan\tfrac{x}{2} - 3\right| - \ln\left|\tan\tfrac{x}{2} - 1\right| + c.$$

7.6.3 $\int R(x, \sqrt{ax^2+bx+c})\,dx$ where R is rational.

There are two tricks that might work here.

Case 1. ax^2+bx+c has real roots r_1, r_2, so $ax^2+bx+c = a(x-r_1)(x-r_2)$. Substitute $t = \dfrac{\sqrt{ax^2+bx+c}}{x-r_1}$. Then

$$t^2 = \frac{ax^2+bx+c}{(x-r_1)^2} = \frac{a(x-r_2)}{x-r_1}.$$

$$t^2x - r_1t^2 = ax - ar_2.$$

Thus

$$x = \frac{r_1t^2 - ar_2}{t^2-a}.$$

Notice that

$$x - r_1 = \frac{a(r_1-r_2)}{t^2-a}.$$

Hence

$$dx = \frac{d}{dt}\Big(\frac{a(r_1-r_2)}{t^2-a}\Big)dt = \frac{-2a(r_1-r_2)t}{(t^2-a)^2}\,dt$$

and

$$\sqrt{ax^2+bx+c} = \sqrt{a(x-r_1)(x-r_2)} = (x-r_1)\sqrt{\frac{a(x-r_2)}{x-r_1}}$$
$$= (x-r_1)t = \frac{a(r_1-r_2)t}{t^2-a}.$$

These are all rational substitutions.

7.6.4. EXAMPLE. $\displaystyle\int \frac{x}{(7x-10-x^2)^{3/2}}$.

Then $-x^2+7x-10 = -(x-2)(x-5)$. Set $t = \dfrac{\sqrt{7x-10-x^2}}{x-2}$. The formulae above show that

$$x = \frac{2t^2+5}{t^2+1} \qquad \sqrt{7x-10-x^2} = \frac{3t}{t^2+1} \quad \text{and} \quad dx = \frac{-6t}{(t^2+1)^2}\,dt.$$

$$\begin{aligned}\int \frac{x}{(7x-10-x^2)^{3/2}} &= \int \frac{2t^2+5}{t^2+1}\,\frac{(t^2+1)^3}{(3t)^3}\,\frac{-6t}{(t^2+1)^2}\,dt \\ &= \int \frac{(2t^2+5)(-6)}{27t^2}\,dt = \int \frac{-4t^2-10}{9t^2}\,dt = -\frac{4}{9}t + \frac{10}{9t} + c \\ &= \frac{-4}{9}\frac{\sqrt{7x-10-x^2}}{x-2} + \frac{10}{9}\frac{x-2}{\sqrt{7x-10-x^2}} + c\end{aligned}$$

$$= \frac{-4(7x - 10 - x^2)/(x-2) + 10(x-2)}{9\sqrt{7x - 10 - x^2}} + c$$
$$= \frac{-4(5-x) + 10(x-2)}{9\sqrt{7x - 10 - x^2}} + c = \frac{14x - 40}{9\sqrt{7x - 10 - x^2}} + c$$

Case 2. $c > 0$. Set $t = \dfrac{\sqrt{ax^2 + bx + c} - \sqrt{c}}{x}$. Then

$$(xt + \sqrt{c})^2 = ax^2 + bx + c \quad \text{or} \quad x^2t^2 + 2\sqrt{c}xt = ax^2 + bx.$$

Divide by x and solve

$$x = \frac{b - 2\sqrt{c}t}{t^2 - a} =: h(t).$$

Therefore

$$dx = h'(t)\,dt \quad \text{and} \quad \sqrt{ax^2 + bx + c} = xt + \sqrt{c} = h(t)t + \sqrt{c}.$$

These are all rational functions of t.

7.6.5. Example. $\displaystyle\int \frac{1}{x\sqrt{x^2 + x + 1}}\,dx.$

Set $t = \dfrac{\sqrt{x^2 + x + 1} - 1}{x}$. The formulae above show that $x = h(t) = \dfrac{1 - 2t}{t^2 - 1}$. Thus

$$\sqrt{x^2 + x + 1} = \frac{t - 2t^2}{t^2 - 1} + 1 = \frac{t - t^2 - 1}{t^2 - 1}$$

and

$$dx = h'(t)\,dt = \frac{2(t^2 - t + 1)}{(t^2 - 1)^2}\,dt.$$

$$\int \frac{1}{x\sqrt{x^2 + x + 1}} = \int \frac{t^2 - 1}{1 - 2t}\,\frac{t^2 - 1}{t - t^2 - 1}\,\frac{2(t^2 - t + 1)}{(t^2 - 1)^2}\,dt$$
$$= \int \frac{2}{2t - 1}\,dt = \ln|2t - 1| + c$$
$$= \ln\left|\frac{2\sqrt{x^2 + x + 1} - 2}{x} - 1\right| + c$$
$$= \ln|2\sqrt{x^2 + x + 1} - 2 - x| - \ln|x| + c$$

Exercises for Chapter 7

1. Compute the following integrals:

(a) $\displaystyle\int \frac{\sin^3 x}{\sqrt{\cos x}}\,dx$

(b) $\displaystyle\int x^2 \sin^{-1}(x^3)\,dx$

(c) $\displaystyle\int_0^{63} \frac{dt}{\sqrt{1+t}+\sqrt[3]{1+t}}$

(d) $\displaystyle\int_1^2 \big(\log x\big)^2\,dx$

(e) $\displaystyle\int e^{2x}\cos(3x)\,dx.$

(f) $\displaystyle\int_{-1}^1 x^3 e^{x^4}\cos 2x\,dx$

(g) $\displaystyle\int \frac{5x^2-13x+9}{x^3-3x^2+4}\,dx$

(h) $\displaystyle\int_{-3}^{-2} \frac{x^2+8x+10}{(x^2+6x+10)^2}\,dx$

(i) $\displaystyle\int_{-\pi/2}^{\pi/2} \frac{1}{5+\sin x+7\cos x}\,dx.$

2. Compute a recursion formula for $I_m = \displaystyle\int x^a(\log x)^m\,dx$, $m \geqslant 0$, $a \neq -1$.

Hence obtain an explicit formula for I_3.

3. Compute $\displaystyle\int_0^{\pi} \frac{x\sin x}{1+\cos^2 x}\,dx.$

HINT: Substitute $u = \pi - x$ and combine the two integrals.

4. Suppose that $f(x)$ is a C^2 function on $\mathbb{R}$ such that $|f(x)| \leqslant A$ and $|f''(x)| \leqslant C$ for $x \in \mathbb{R}$. Prove that $|f'(x)| \leqslant \sqrt{2AC}$.

HINT: fix x_0 with $f'(x_0) = b \geqslant 0$. Get a lower bound for $f'(x_0 \pm h)$.

Use this to estimate $\displaystyle\int_{x_0-H}^{x_0+H} f'(x)\,dx$ for a good choice of H.

5. Suppose that $f(0) = 0$ and $0 < f'(x) \leqslant 1$ for all $x \geqslant 0$. Show that

$$\int_0^x f(t)^3\,dt \leqslant \left(\int_0^x f(t)\,dt\right)^2 \qquad \text{for all } x > 0.$$

When does equality hold? HINT: differentiate, factor and differentiate again.

6. Compute $\displaystyle\int \sqrt{\tan x}\,dx$. HINT: try setting $u^2 = \tan x$.

CHAPTER 8

Other Aspects of Integration

8.1. Improper integrals

Sometimes it is not enough to integrate bounded functions on bounded intervals, which is what the Riemann integral accomplishes. When the domain is unbounded or the function is unbounded, there is a way to extend the definition of integral. These are called *improper integrals* to stress the point that they are not Riemann integrable.

8.1.1. DEFINITION. Let $f(x)$ be Riemann integrable on $[a, b]$ for all $b > a$. We define

$$\int_a^\infty f(x)\,dx = \lim_{b\to\infty} \int_a^b f(x)\,dx$$

when the limit exists. Similarly we can define $\displaystyle\int_{-\infty}^b f(x)\,dx$. If f is Riemann integrable on $[a, b]$ for all $a < b \in \mathbb{R}$, we let

$$\int_{-\infty}^\infty f(x)\,dx = \lim_{a\to-\infty} \lim_{b\to\infty} \int_a^b f(x)\,dx.$$

8.1.2. EXAMPLE. Let $f(x) = \dfrac{1}{1+x^2}$. Then

$$\int_0^\infty \frac{1}{1+x^2}\,dx = \lim_{b\to\infty} \int_0^b \frac{1}{1+x^2}\,dx = \lim_{b\to\infty} \tan^{-1}(x)\Big|_0^b = \lim_{b\to\infty} \tan^{-1}(b) = \frac{\pi}{2}.$$

Similarly,

$$\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx = \lim_{a\to-\infty} \lim_{b\to\infty} \tan^{-1}(b) - \tan^{-1}(a) = \pi.$$

8.1.3. EXAMPLE. Let $f(x) = \begin{cases} x & \text{if } |x| \leqslant 1 \\ \frac{1}{x} & \text{if } |x| \geqslant 1 \end{cases}$. Then

$$\int_0^b f(x)\,dx = \frac{1}{2} + \ln b \quad \text{and} \quad \int_a^0 f(x)\,dx = -\frac{1}{2} - \ln|a|$$

for $b > 1$ and $a < -1$. Thus

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{a\to-\infty}\lim_{b\to\infty} \ln b - \ln|a|.$$

This limit is undefined, and the function does not have an integral. Notice that

$$\lim_{b\to\infty}\int_{-b}^{b} f(x)\,dx = 0.$$

However

$$\lim_{b\to\infty}\int_{-b}^{rb} f(x)\,dx = \lim_{b\to\infty} \ln(rb) - \ln b = \ln r$$

for any $r > 0$. Thus this integral might be assigned *any* real value if we specify how we approach the two limits in a synchronized fashion. So this function is not integrable. There are some instances where the limit $\lim\limits_{b\to\infty}\int_{-b}^{b} f(x)\,dx$ is used, and in this case it is called the *principal value* of the integral.

8.1.4. EXAMPLE. Let $f(x) = x^p$ for $p < 0$ and $p \neq -1$. Then

$$\int_{1}^{\infty} x^p\,dx = \lim_{b\to\infty}\int_{0}^{b} x^p\,dx = \lim_{b\to\infty} \frac{x^{p+1}}{p+1}\Big|_0^b$$
$$= \lim_{b\to\infty}\frac{b^{p+1}-1}{p+1} = \begin{cases} +\infty & \text{if } -1 < p < 0 \\ \frac{1}{|p+1|} & \text{if } p < -1 \end{cases}$$

Thus f is integrable only when $p < -1$.

The following is a useful check that will guarantee that an improper integral exists. Later on in our study of series, this can be compared to absolute convergence.

8.1.5. PROPOSITION. *Let $f(x)$ be Riemann integrable on $[a,b]$ for all $b > a$. If $\int_a^{\infty} |f(x)|\,dx < \infty$, then $\int_a^{\infty} f(x)\,dx$ exists.*

PROOF. Proposition 6.5.1 shows that $|f(x)|$ is also Riemann integrable on $[a,b]$ for $b > a$. So we can define

$$F(x) = \int_a^x f(t)\,dt \quad\text{and}\quad G(x) = \int_a^x |f(t)|\,dt \quad\text{for}\quad x > a.$$

Note that $G(x)$ is monotone increasing, and by hypothesis, it is bounded above. Thus $\lim\limits_{x\to\infty} G(x) = \sup_{x>a} G(x) = M < \infty$.

Given $\varepsilon > 0$, choose b_0 so that $G(b_0) > M - \varepsilon$. Then if $b_0 \leqslant b_1 \leqslant b_2$, then

$$\begin{aligned} |F(b_2) - F(b_1)| &= \Big| \int_{b_1}^{b_2} f(t)\,dt \Big| \leqslant \int_{b_1}^{b_2} |f(t)|\,dt \\ &= G(B_2) - G(b_1) \leqslant M - G(b_0) < \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the values $F(b)$ satisfy the Cauchy condition, and therefore $\lim_{b\to\infty} F(b) = \int_a^\infty f(x)\,dx$ exists. ■

8.1.6. EXAMPLE. Let $f(x) = \dfrac{\sin x}{x^2}$ for $x \geqslant \pi$. Then

$$\int_\pi^\infty |f(x)|\,dx = \int_\pi^\infty \frac{|\sin x|}{x^2}\,dx \leqslant \int_\pi^\infty \frac{1}{x^2}\,dx = \frac{-1}{x}\Big|_\pi^\infty = \frac{1}{\pi} < \infty.$$

Therefore Proposition 8.1.5 shows that $\displaystyle\int_\pi^\infty \frac{\sin x}{x^2}\,dx$ exists.

8.1.7. EXAMPLE. Here is a more subtle example where Proposition 8.1.5 does not apply. Let $f(x) = \dfrac{\sin x}{x}$ for $x \geqslant 0$, and consider $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx$. Since $f(x)$ has limit 1 as $x \to 0$, setting $f(0) = 1$ makes $f(x)$ continuous. Therefore it is Riemann integrable on $[0, b]$ for every $b > 0$. Now $\sin x$ changes sign at each multiple of π. Define

$$a_k = \int_{(k-1)\pi}^{k\pi} \frac{\sin x}{x}\,dx \quad \text{for} \quad k \geqslant 1.$$

Then $a_k = (-1)^{k-1}|a_k|$ alternates sign. Observe that for $(k-1)\pi \leqslant x \leqslant k\pi$,

$$\frac{|\sin x|}{k\pi} \leqslant \frac{|\sin x|}{x} \leqslant \frac{|\sin x|}{(k-1)\pi}.$$

Therefore

$$\int_{(k-1)\pi}^{k\pi} \frac{|\sin x|}{k\pi}\,dx \leqslant \int_{(k-1)\pi}^{k\pi} \frac{|\sin x|}{x}\,dx \leqslant \int_{(k-1)\pi}^{k\pi} \frac{|\sin x|}{(k-1)\pi}\,dx$$

Thus

$$\frac{2}{k\pi} \leqslant |a_k| \leqslant \frac{2}{(k-1)\pi}.$$

It follows that

$$\int_0^{n\pi} \frac{|\sin x|}{x}\,dx \geqslant \sum_{k=1}^n \frac{1}{k\pi} = \frac{1}{\pi}\sum_{k=1}^n \frac{1}{k} = +\infty.$$

This is called the *harmonic series*. The reason it diverges is that

$$1 + \tfrac{1}{2} + (\tfrac{1}{3} + \tfrac{1}{4}) + (\tfrac{1}{5} + \dots + \tfrac{1}{8}) + (\tfrac{1}{9} + \dots + \tfrac{1}{16}) + \dots > 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots$$

because $\frac{1}{2^{n-1}+1} + \cdots + \frac{1}{2^n} > \dfrac{2^{n-1}}{2^n} = \dfrac{1}{2}$ for each $n \geqslant 1$.

On the other hand, we have $|a_n| > \frac{1}{n\pi} \geqslant |a_{n+1}|$, so the sequence $|a_n$ decreases monotonely to 0. The Alternating Series Test 9.3.2, which we prove later, shows that the series

$$\sum_{k=1}^{n} a_k = \int_0^{n\pi} \frac{\sin x}{x}\, dx$$

converges. From this it follows that $\displaystyle\int_0^{\infty} \frac{\sin x}{x}\, dx$ is defined.

It takes some sophisticated methods to find the actual limit, which is $\dfrac{\pi}{2}$. We provide one proof in Appendix A.7.

Now we deal with functions which are unbounded as they approach a single point.

8.1.8. DEFINITION. Suppose that $f(x)$ is Riemann integrable on $[a+\varepsilon, b]$ for all $\varepsilon > 0$ but f is unbounded on $[a, b]$. We say that the *improper integral* $\displaystyle\int_a^b f(x)\, dx$ exists if there is a limit

$$\int_a^b f(x)\, dx := \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^b f(x)\, dx.$$

Similarly we define an improper integral if $f(x)$ becomes unbounded as it approaches b. If $f(x)$ becomes unbounded as x approaches an interior point c but is Riemann integrable on $[a, c-\varepsilon]$ and $[c+\delta, b]$ for $\varepsilon, \delta > 0$, then the improper integral exists if there are two limits

$$\lim_{\varepsilon \to 0^+} \int_a^{c-\varepsilon} f(x)\, dx \quad \text{and} \quad \lim_{\delta \to 0^+} \int_{c+\delta}^b f(x)\, dx$$

and then $\displaystyle\int_a^b f(x)\, dx := \lim_{\varepsilon \to 0^+} \int_a^{c-\varepsilon} f(x)\, dx + \lim_{\delta \to 0^+} \int_{c+\delta}^b f(x)\, dx.$

8.1.9. EXAMPLE. For $a < 0$, consider $\int_0^1 x^a\,dx$. First consider $a \neq -1$.

$$\begin{aligned}\int_0^1 x^a\,dx &= \lim_{\varepsilon\to 0^+}\int_\varepsilon^1 x^a\,dx\\ &= \lim_{\varepsilon\to 0^+}\left.\frac{x^{a+1}}{a+1}\right|_\varepsilon^1 = \lim_{\varepsilon\to 0^+}\frac{1-\varepsilon^{a+1}}{a+1}\\ &= \begin{cases}\frac{1}{a+1} & \text{if } -1 < a < 0\\ -\infty & \text{if } a < -1\end{cases}\end{aligned}$$

Similarly the integral $\int_0^1 x^{-1}\,dx$ does not exist. Note that for this reason, $\int_{-1}^1 x^{-1}\,dx$ also does not exist. However as in Example 8.1.3,

$$\lim_{\varepsilon\to 0^+}\int_{-1}^{-\varepsilon}\frac{dx}{x}+\int_\varepsilon^1\frac{dx}{x} = 0 \quad\text{and}\quad \lim_{\varepsilon\to 0^+}\int_{-1}^{-2\varepsilon}\frac{dx}{x}+\int_\varepsilon^1\frac{dx}{x} = \ln 2.$$

8.2. Volumes

8.2.1 Disk method. Consider a volume obtained by rotating a function $f(x)$ about the x-axis from $x = a$ to $x = b$.

The idea is to consider a slice perpendicular to the x-axis at a point x. The cross section is a solid disc of radius $f(x)$. We imagine that it has an infinitesimal thickness dx and integrate. The disc has area $\pi f(x)^2$. Thus the volume is

$$V = \int_a^b \pi f(x)^2\,dx.$$

If we are integrating a solid figure bounded above by $f(x)$ and below by $g(x)$, then we can think of this as the volume of the rotation of $f(x)$ minus the volume obtained by rotating $g(x)$.

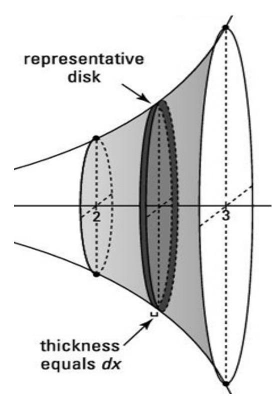


FIGURE 8.1. disc method

8.2.2. EXAMPLE. Compute the volume of a doughnut obtained by taking a disc of radius r and rotating it about an axis which is distance R from the centre of the disc. For convenience, let the disc have centre $(0, R)$. Then the upper and lower arcs of the boundary circle are $f(x) = R + \sqrt{r^2 - x^2}$ and $g(x) = R - \sqrt{r^2 - x^2}$

for $-r \leqslant x \leqslant r$. Thus the volume is

$$\begin{aligned}
V &= \pi \int_{-r}^{r} f(x)^2 - g(x)^2 \, dx = \pi \int_{-r}^{r} \big(R + \sqrt{r^2 - x^2}\big)^2 - \big(R - \sqrt{r^2 - x^2}\big)^2 \, dx \\
&= \pi \int_{-r}^{r} 4R\sqrt{r^2 - x^2} \, dx \qquad \text{substitute } x = r \sin\theta \text{ and } dx = r\cos\theta \, d\theta \\
&= 4\pi R \int_{-\pi/2}^{\pi/2} \sqrt{r^2 - r^2 \sin^2\theta} \, r \cos\theta \, d\theta = 4\pi R r^2 \int_{-\pi/2}^{\pi/2} \cos^2\theta \, d\theta \\
&= 2\pi R r^2 \int_{-\pi/2}^{\pi/2} \cos 2\theta + 1 \, d\theta = 2\pi R r^2 \big(\tfrac{1}{2}\sin 2\theta + \theta\big)\Big|_{-\pi/2}^{\pi/2} \\
&= 2\pi^2 R r^2 = (2\pi R)(\pi r^2).
\end{aligned}$$

8.2.3. EXAMPLE. There is no reason to restrict this technique to circular cross sections. Let's compute the volume of a regular tetrahedron with side length s.

The base of the tetrahedron is an equilateral triangle. We need a formula for the height. When you drop a perpendicular line from the apex to the base, it hits the centroid of the base triangle. Now $\overline{AB} = \frac{s}{2}$. Also $\triangle ABC$ is similar to $\triangle DBA$. so that $\overline{BC} = \frac{1}{2}\overline{AC}$. Let $t = \overline{DC} = \overline{AC} = t$. Then $\overline{DB} = t + \frac{t}{2} = \frac{\sqrt{3}s}{2}$.

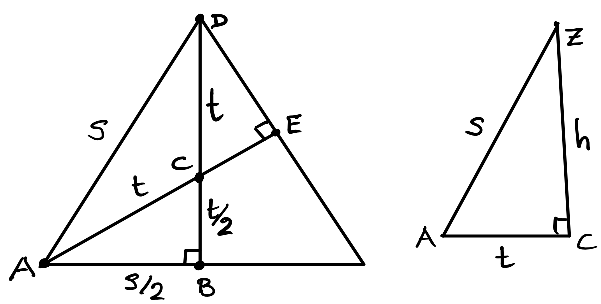


FIGURE 8.2. height of a regular tetrahedron

Thus $t = \frac{s}{\sqrt{3}}$. Let Z be the apex of the tetrahedron, and observe that $\triangle ACZ$ is a right triangle. Therefore the height is $h = \sqrt{s^2 - t^2} = \sqrt{\frac{2}{3}}s$. The cross section at height x is an equilateral triangle with side length proportional to $h - x$ changing linearly from s when $x = 0$ to 0 when $x = h$, so that the side length is $\frac{s}{h}(h - x) = \sqrt{\frac{3}{2}}(h - x)$. The area of an equilateral triangle of side y is $\frac{1}{2}y^2 \sin\frac{\pi}{3} = \frac{\sqrt{3}}{4}y^2$. Thus the volume is

$$V = \int_0^h \frac{\sqrt{3}}{4}\,\frac{3}{2}(h - x)^2 \, dx = -\frac{\sqrt{3}}{8}(h - x)^3\Big|_0^h = \frac{\sqrt{3}}{8}h^3 = \frac{s^3}{6\sqrt{2}}.$$

8.2.4 Cylinder method. Consider a volume obtained by rotating a function $f(x)$ from $x = a$ to $x = b$ about the y-axis. The vertical line segment from $(x, 0)$

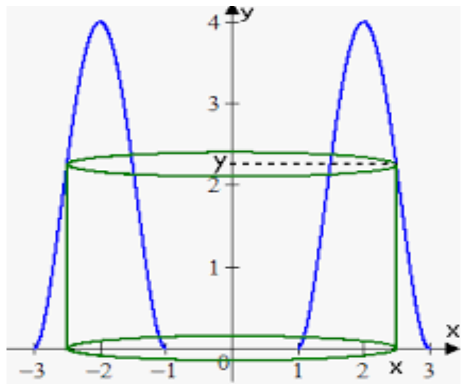


FIGURE 8.3. cylinder method

to $(x, f(x))$ is swept around the y axis to form a cylinder of radius x and height $f(x)$. We introduce an infinitesimal thickness dx and integrate. The volume of this thin cyclinder is the surface area of the cylinder times dx. The surface area is the circumference of the circle times the height, so $2\pi x f(x)\,dx$ is the infinitesimal volume. Hence the volume is

$$V = \int_a^b 2\pi x f(x)\,dx.$$

If the region swept around the y-axis is bounded above by $f(x)$ and below by $g(x)$, then the volume is

$$V = \int_a^b 2\pi x\big(f(x) - g(x)\big)\,dx.$$

8.2.5. EXAMPLE. Compute the volume of a sphere of radius r by sweeping a semicircle with diameter along the y axis around. Put the centre at $(0, 0)$ so that $f(x) = \sqrt{r^2 - x^2}$ and $g(x) = -\sqrt{r^2 - x^2}$ Thus the volume is

$$\begin{aligned} V &= \int_0^r 2\pi x\big(f(x) - g(x)\big)\,dx = 4\pi \int_0^r x\sqrt{r^2 - x^2}\,dx \\ &= -4\pi\,\frac{1}{3}(r^2 - x^2)^{3/2}\Big|_0^r = \frac{4}{3}\pi r^3. \end{aligned}$$

8.2.6. EXAMPLE. The *centroid* or *centre of mass* of a planar object is the point at which it will balance on a pin. The physical information that we need is that mass at distance h from the midpoint will exert a force proportional to h. We assume that the planar object is of uniform density; and that it is bounded above and below by $f(x)$ and $g(x)$ for $a \leqslant x \leqslant b$. We compute the x and y coordinates separately. To compute the x coordinate $\bar{x}$, we imaging the figure balancing on the line $x = \bar{x}$. Let the cross section at x be $l(x) = f(x) - g(x)$. If the figure when sitting on $x = \bar{x}$, the force exerted by a rectangle of length $l(x)$ and infinitesimal width dx is $l(x)(x - \bar{x})\,dx$. The total force should be 0, whence

$$0 = \int_a^b (x - \bar{x})l(x)\,dx = \int_a^b x\big(f(x) - g(x)\big)\,dx - \bar{x}\int_a^b f(x) - g(x)\,dx.$$

Since the area of the figure is $A = \displaystyle\int_a^b f(x) - g(x)\,dx$, we obtain

$$\bar{x} = \frac{1}{A}\int_a^b x\big(f(x) - g(x)\big)\,dx.$$

A similar formula will hold for $\bar{y}$. The centroid is then $(\bar{x}, \bar{y})$.

Now if we assume that this body lies in the right half plane, and we rotate this body around the y axis, the formula for the volume is known as *Pappus's Theorem*.

$$V = \int_a^b 2\pi x\big(f(x) - g(x)\big)\,dx = 2\pi A\bar{x}$$

8.2.7. EXAMPLE. Compute the volume of the intersection of two solid cylinders of the same diameter which meet at right angles. If we align the cylinders along the x and y axes, respectively, then we have

$$C_1 = \{(x, y, z) : y^2 + z^2 \leqslant r^2\} \quad \text{and} \quad C_2 = \{(x, y, z) : x^2 + z^2 \leqslant r^2\}.$$

Fix a value z_0 with $|z_0| \leqslant r$. The intersection of the plane $z = z_0$ is

$$\{(x, y, z_0) : x^2 \leqslant r^2 - z_0^2,\ y^2 \leqslant r^2 - z_0^2\}.$$

This is a square of area $4(r^2 - z_0^2)$. Thus the volume is

$$V = \int_{-r}^{r} 4(r^2 - z^2)\,dz = 4r^2 z - \frac{4}{3}z^3\Big|_{-r}^{r} = 8r^3 - \frac{8}{3}r^3 = \frac{16}{3}r^3.$$

8.2.8. EXAMPLE. Now compute the volume of the intersection of three solid cylinders of the same diameter which meet at right angles. If we align the cylinders along the x, y, z axes, respectively, then we have a third cylinder

$$C_3 = \{(x, y, z) : x^2 + y^2 \leqslant r^2\}.$$

Notice that the intersection $C = C_1 \cap C_2 \cap C_3$ contains the cube

$$D = \{(x, y, z) : |x| \leqslant \frac{r}{\sqrt{2}},\ |y| \leqslant \frac{r}{\sqrt{2}},\ |z| \leqslant \frac{r}{\sqrt{2}}\}$$

In addition, it contains a 'cap' on each of the six faces of the cube. Let us intersect the plane $z = z_0$ for $z_0 \geqslant \frac{r}{\sqrt{2}}$ with C. As in the previous example, we obtain a square $\{(x,y,z_0) : x^2 \leqslant r^2 - z_0^2,\ y^2 \leqslant r^2 - z_0^2\}$ which has area $4(r^2 - z_0^2)$. Thus the volume is

$$\begin{aligned} V = V(D) + 6V(\text{cap}) &= 8\Big(\frac{r}{\sqrt{2}}\Big)^3 + 6\int_{r/\sqrt{2}}^{r} 4(r^2 - z^2)\,dz \\ &= 2\sqrt{2}r^3 + 24(r^2x - \tfrac{1}{3}z^3)\Big|_{r/\sqrt{2}}^{r} \\ &= 2\sqrt{2}r^3 + 24r^3 - 12\sqrt{2}r^3 - 8r^3 + 2\sqrt{2}r^3 \\ &= 16r^3 - 8\sqrt{2}r^3 = 8(2-\sqrt{2})r^3. \end{aligned}$$

The set C is interesting geometrically. The six caps each consist of 4 curvilinear triangles. For example, on the cap computed above with $z \geqslant r/\sqrt{2}$, the four curves defined by

$$x = \pm\sqrt{r^2 - z_0^2} \quad \text{and} \quad y = \pm\sqrt{r^2 - z_0^2} \quad \text{for} \quad \frac{r}{\sqrt{2}} \leqslant z \leqslant r$$

lie on $C_1 \cap C_2$ and together with the four sides of the square face of the cube

$$D_1 = \{(x,y,z) : z = \frac{r}{\sqrt{2}},\ x^2 \leqslant r^2/2,\ y^2 \leqslant r^2/2\}$$

determine the four triangular regions. The 'triangle'

$$T = \{(x,y,z) : x^2 \leqslant r^2 - z^2,\ y = \sqrt{r^2 - z^2},\ \frac{r}{\sqrt{2}} \leqslant z \leqslant r\}$$

lies on the surface of C_1. This triangle fits together with another triangle on the adjacent cap and form one connected rhombus on the surface of C_1.

So instead of 24 triangles, there are actually 12 congruent rhombuses that fit together to make the surface of C. There are 14 vertices, 8 corners of the cube and 6 vertices at the top of each cap. Four rhombuses meet at each cap vertex and three meet at each corner of the cube. There is a semiregular solid called a *rhombic dodecahedron* with twelve congruent rhombic faces. The intersection C is a curvy version of it. We can check the Euler characteristic. We have 12 faces, 14 vertices, and since each rhombus has 4 sides, but each side lies on two faces, there are $12(4)/2 = 24$ edges. The Euler characteristic is $F - E + V = 12 - 24 + 14 = 2$, which is the same for any convex solid polyhedron in 3-space.

8.3. Arc length

In this section, we explain how to compute the length of a curve. Suppose that the curve is $y = f(x)$ from $x = a$ to $x = b$. Let $s(x)$ be the length of the curve from a to x. As in computation of areas, we compute the infinitesimal change in s from x to $x + \Delta x$.

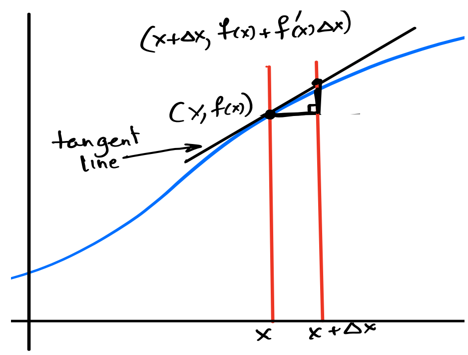


FIGURE 8.4. computing arc length

On the infinitesimal interval from x to $x + \Delta x$, the curve $f = f(x)$ is well approximated by the tangent line $y(t) = f(x) + f'(x)(t - x)$. So the segment of the curve has length

$$\Delta s = \sqrt{\Delta x^2 + \Delta f(x)^2} = \sqrt{\Delta x^2 + (f'(x)\Delta x)^2} = \sqrt{1 + f'(x)^2}\,\Delta x.$$

This leads us to the formula

$$s = s(b) = \int_a^b \sqrt{1 + f'(x)^2}\,dx.$$

8.3.1. EXAMPLE. Compute the circumference of a circle of radius r. It is enough to compute the length of a semicircle and double it. So let $f(x) = \sqrt{r^2 - x^2}$ for $-r \leqslant x \leqslant r$. Thus $f'(x) = \dfrac{-x}{\sqrt{r^2 - x^2}}$. Therefore its length is

$$s = \int_{-r}^{r} \sqrt{1 + \frac{x^2}{r^2 - x^2}}\,dx = \int_{-r}^{r} \sqrt{\frac{r^2}{r^2 - x^2}}\,dx = r\int_{-r}^{r} \frac{1}{\sqrt{r^2 - x^2}}\,dx$$

Substitute $x = r\sin\theta$ for $-\frac{\pi}{2} \leqslant \theta \leqslant \frac{\pi}{2}$; so $dx = r\cos\theta\,d\theta$. Hence

$$s = r\int_{-\pi/2}^{\pi/2} \frac{r\cos\theta}{\sqrt{r^2 - r^2\sin^2\theta}}\,d\theta = \pi r.$$

Therefore a circle of radius r has a circumference of $2\pi r$.

8.3.2. EXAMPLE. Consider a uniform cable suspended between two points (x_1, y_1) and (x_2, y_2). What is the shape and length of this curve?

The shape of a hanging cable is called a *catenary*. Say $y = f(x)$. There will be a point $x_0 \in (x_1, x_2)$ at which the cable attains its minimum height, so $f'(x_0) = 0$.

It always has the form $y = f(x) = a + b\cosh\frac{x-x_0}{b}$. The length of the cable is

$$\begin{aligned} L &= \int_{x_1}^{x_2} \sqrt{1+f'(x)^2}\,dx = \int_{x_1}^{x_2} \sqrt{1+\sinh^2\tfrac{x-x_0}{b}}\,dx \\ &= \int_{x_1}^{x_2} \cosh\tfrac{x-x_0}{b}\,dx = b\sinh\tfrac{x-x_0}{b}\Big|_{x_1}^{x_2} \\ &= b\sinh\tfrac{x_2-x_0}{b} - b\sinh\tfrac{x_1-x_0}{b}. \end{aligned}$$

In particular, if $y_1 = y_2$, we have $x_0 = \frac{x_1+x_2}{2}$ by symmetry, and $L = 2b\sinh\frac{x_2-x_1}{2b}$.

To derive the shape of the curve, let ρ be the density of the cable. Consider the segment of cable between (x_0, y_0) and another point (x, y) on the curve. The length of the cable from x_0 to x is $s(x) = \int_{x_0}^{x} \sqrt{1+f'(t)^2}\,dt$; and thus this section of cable has mass $\rho s(x)$. There are three forces acting on this segment of chain. There is a force T_0 at x_0 tangent to the curve, and thus parallel to the ground provided by the tension on the cable. Likewise there is a tension T_1 at the point x with slope $f'(x)$. Finally there is the force of gravity, $\rho g s(x)$, where g is the force of gravity. Since the cable is in equilibrium, these forces must sum to 0. See figure.

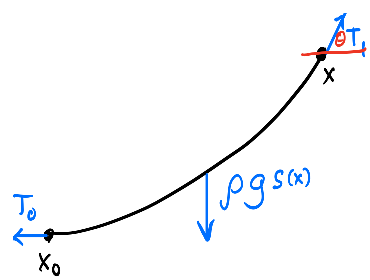


FIGURE 8.5. hanging cable at equilibrium

Let the tangent at $(x, f(x))$ be at an angle θ to the horizontal. Then $f'(x) = \tan\theta$. We get

$$T_0 = T_1\cos\theta \quad \text{and} \quad \rho g s(x) = T_1\sin\theta.$$

Therefore

$$f'(x) = \frac{T_1\sin\theta}{T_1\cos\theta} = \frac{\rho g}{T_0}\int_{x_0}^{x}\sqrt{1+f'(t)^2}\,dt.$$

Let $b = \dfrac{T_0}{\rho g}$, which is a constant. For simplicity of notation, write $p(x) = f'(x)$. Then

$$p(x) = \frac{1}{b}\int_{x_0}^{x}\sqrt{1+p(t)^2}\,dt.$$

Differentiate this using the FTC to get

$$p'(x) = \frac{1}{b}\sqrt{1 + p(t)^2} \quad \text{or} \quad \frac{p'(x)}{\sqrt{1 + p(t)^2}} = \frac{1}{b}.$$

Now integrate this. We recall that $\sinh^2 u + 1 = \cosh^2 u$. Hence we substitute $p = \sinh u$. Then $p'(t)\,dt = \cosh u\,du$. Thus

$$\int \frac{p'(x)}{\sqrt{1 + p(t)^2}}\,dt = \int \frac{\cosh u}{\cosh u}\,du = u = \sinh^{-1}(p(t)).$$

Therefore since $\sinh^{-1}(0) = 0$,

$$\frac{x - x_0}{b} = \int_{x_0}^{x} \frac{1}{b}\,dt = \int_{x_0}^{x} \frac{p'(x)}{\sqrt{1 + p(t)^2}}\,dt = \sinh^{-1}(p(t))\Big|_{x_0}^{x} = \sinh^{-1}(p(x)).$$

That is,

$$f'(x) = p(x) = \sinh \frac{x - x_0}{b}.$$

Integrating again, we get

$$\begin{aligned} f(x) &= f(x_0) + \int_{x_0}^{x} \sinh \tfrac{x-x_0}{b}\,dt = y_0 + b\cosh \tfrac{x-x_0}{b}\Big|_{x_0}^{x} \\ &= (y_0 - b) + b\cosh \tfrac{x-x_0}{b}. \end{aligned}$$

So the shape of a catenary is a hyperbolic cosine as claimed.

This still does not explicitly determine b in terms of known quantities. We can make the following computation.

$$\begin{aligned} L^2 - (y_2 - y_1)^2 &= \big(b\sinh \tfrac{x_2-x_0}{b} - b\sinh \tfrac{x_1-x_0}{b}\big)^2 - \big(b\cosh \tfrac{x_2-x_0}{b} - b\cosh(\tfrac{x_1-x_0}{b}\big)^2 \\ &= b^2\big(\sinh^2 \tfrac{x_2-x_0}{b} - \cosh^2 \tfrac{x_2-x_0}{b} + \sinh^2 \tfrac{x_1-x_0}{b} - \cosh^2 \tfrac{x_1-x_0}{b} \\ &\qquad + 2\cosh(\tfrac{x_2-x_0}{b}\cosh \tfrac{x_1-x_0}{b} - 2\sinh \tfrac{x_2-x_0}{b}\sinh \tfrac{x_1-x_0}{b}\big) \\ &= b^2\big(2\cosh \tfrac{x_2-x_1}{b} - 2\big) = 2b^2\big(1 + 2\sinh^2 \tfrac{x_2-x_1}{2b} - 1\big) \\ &= 4b^2\sinh^2 \tfrac{x_2-x_1}{2b}. \end{aligned}$$

Therefore

$$2b\sinh \tfrac{x_2-x_1}{2b} = \sqrt{L^2 - (y_2 - y_1)^2}.$$

The quantities L, y_1 and y_2 are known, and the LHS, as a function of b, is monotone decreasing on $(0, \infty)$. (Check that $g(t) = t\sinh \frac{c}{t}$ has $g''(t) > 0$ on $(0, \infty)$, and $\lim_{t\to\infty} g'(t) = 0$.) Thus there is a unique solution to this equation. In general, this computation requires numerical techniques such as Newton's Method.

In the special case in which $y_2 = y_1$, we can compute b explicitly as a function of the length $L = 2b\sinh \frac{x_2-x_1}{2b}$ and the sag of the cable,

$$h = f(x_1) - f(x_0) = b\big(\cosh \tfrac{x_2-x_1}{2b} - 1\big).$$

Indeed,

$$L^2 - 4h^2 = 4b^2\big(\sinh^2 \tfrac{x_2-x_1}{2b} - (\cosh^2 \tfrac{x_2-x_1}{2b} - 2\cosh \tfrac{x_2-x_1}{2b} + 1)\big)$$
$$= 4b^2(2\cosh \tfrac{x_2-x_1}{2b} - 2) = 8bh.$$

Thus $b = \dfrac{L^2 - 4h^2}{8h}$.

8.4. Polar coordinates

Polar coordinates is an alternative method of specifying a point in the plane. It starts with the positive real axis including the special point of the origin O.

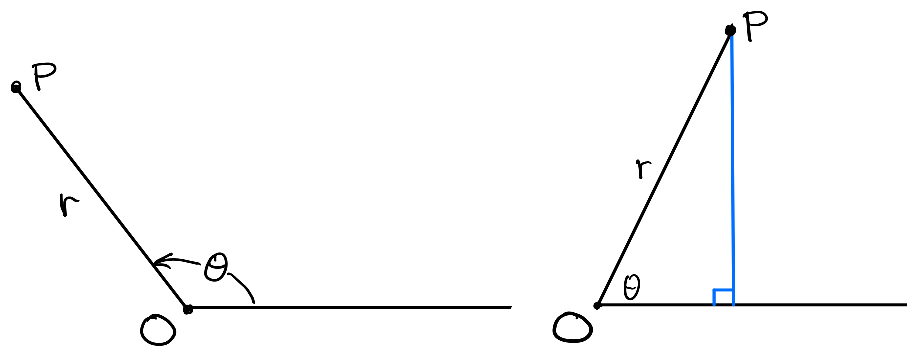


FIGURE 8.6. polar coordinates

A point P in the plane is specified by the distance r from O, which places it on the circle of radius r centred at O, together with the angle θ from the positive real axis in the anticlockwise direction. Of course, θ is in radians because this is calculus!

The angle is only determined up to a multiple of 2π because an angle of 2π is a complete rotation. The point (r,θ) and $(r,\theta+6\pi)$ and $(r,\theta-4\pi)$ all represent the same point. Normally we use $r \geqslant 0$. However should a formula yield a negative value for r, we can interpret this as the opposite direction; i.e., $(-r,\theta) = (r,\theta+\pi)$.

It is not difficult to convert between Cartesian coordinates and polar coordinates. The point (r,θ) corresponds to (x,y) where $x = r\cos\theta$ and $y = r\sin\theta$. Conversely, (x,y) converts to $r = \sqrt{x^2+y^2}$ and $\theta = \cos^{-1}(\frac{x}{r}) \cap \sin^{-1}(\frac{y}{r})$. The point of the intersection is that $\cos\theta = \cos(\pm\theta + 2n\pi)$ while $\sin(\theta + 2n\pi) = \sin(\pi - \theta + 2m\pi)$.

Certain figures are more easily described using polar coordinates. We will see first how to compute area. The small sector of a circle of radius $r(\theta)$ and infinitesimal angle $d\theta$ is $\frac{1}{2}r^2(\theta)\,d\theta$. You can see this from its share of the full circle

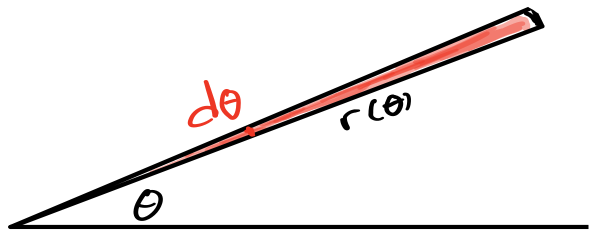


FIGURE 8.7. area in polar coordinates

because the sector of a circle of radius r and angle α is $\frac{1}{2}r^2\alpha$. Thus the area swept out by a curve specified a $r = r(\theta)$ for $\theta_1 \leqslant \theta \leqslant \theta_2$ is

$$A = \int_{\theta_1}^{\theta_2} \frac{1}{2} r^2(\theta)\, d\theta.$$

8.4.1. EXAMPLE. Consider the figure $r = \sqrt{|\sin\theta|}$. This figure has two lobes, one in the upper half plane, and its reflection in the x-axis. The area of both lobes

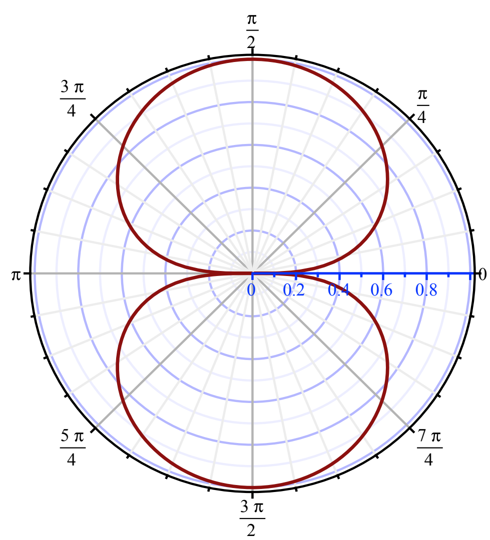


FIGURE 8.8. $r = \sqrt{|\sin\theta|}$

is double the one on top. So

$$A = 2\int_0^{\pi} \frac{1}{2} r^2(\theta)\, d\theta = \int_0^{\pi} \sin\theta\, d\theta = -\cos\theta\Big|_0^{\pi} = 2$$

8.4.2. EXAMPLE. Consider the figure $y^2 = \dfrac{x^2(1+x)}{1-x}$ for $-1 \leqslant x < 1$. It is probably easier to understand this curve in Cartesian coordinates. But for practice, we will convert this to polar coordinates by setting $x = r\cos\theta$ and $y = r\sin\theta$. We get

$$r^2\sin^2\theta = \frac{r^2\cos^2\theta(1+r\cos\theta)}{1-r\cos\theta}$$

or

$$\sin^2\theta - r\sin^2\theta\cos\theta = \cos^2\theta + r\cos^3\theta.$$

Thus

$$r\cos\theta = \sin^2\theta - \cos^2\theta = 1 - 2\cos^2\theta$$

so that

$$r = \sec\theta - 2\cos\theta.$$

When $\theta = 0$, we have $r = -1$, indicating the point $(1,\pi) = (-1,0)$ which is $(-1,0)$ in Cartesian coordinates. We have $r = 0$ when $\cos^2\theta = \frac{1}{2}$, or $\theta = \pm\frac{\pi}{4}$. In the range $[-\frac{\pi}{4}, \frac{\pi}{4}]$, $r(\theta) \leqslant 0$ and the loop of the strophoid is swept out. In the range $[\frac{\pi}{4}, \frac{\pi}{2})$, the radius $r(\theta)$ tends to $+\infty$. It is much easier to see what happens in Cartesian coordinates. As $x \to 1^-$, y^2 tends to infinity. So there is a vertical asymptote at $x = 1$. The range $[\frac{\pi}{4}, \frac{\pi}{2})$ corresponds to the upper part of the curve, while the range $(-\frac{\pi}{2}, -\frac{\pi}{4}]$ corresponds to the lower part of the curve.

Let's compute the area of the loop in two ways. Using polar coordinates,

$$\begin{aligned}
A &= \int_{-\pi/4}^{\pi/4} \frac{1}{2}(\sec\theta - 2\cos\theta)^2\,d\theta = \frac{1}{2}\int_{-\pi/4}^{\pi/4} \sec^2\theta - 4 + 4\cos^2\theta\,d\theta \\
&= \frac{1}{2}\int_{-\pi/4}^{\pi/4} \sec^2\theta - 2 + 2\cos 2\theta\,d\theta = \frac{1}{2}\tan\theta - \theta + \tfrac{1}{2}\sin 2\theta\Big|_{-\pi/4}^{\pi/4} \\
&= \frac{1}{2}(2) - \frac{\pi}{2} + 1 = 2 - \frac{\pi}{2}.
\end{aligned}$$

Using Cartesian coordinates, we see that the area is double the area above the x-axis in the range $[-1,0]$. The best substitution after some manipulation is $x = \sin\theta$

$$\begin{aligned}
A &= 2\int_{-1}^{0} \sqrt{\frac{x^2(1+x)}{1-x}}\,dx = \int_0^1 2x\sqrt{\frac{1-x}{1+x}}\,dx \\
&= \int_0^1 \frac{2x(1-x)}{\sqrt{1-x^2}}\,dx = \int_0^{\pi/2} \frac{2\sin\theta - 2\sin^2\theta}{\cos\theta}\cos\theta\,d\theta \\
&= \int_0^{\pi/2} 2\sin\theta + \cos 2\theta - 1\,d\theta \;= -2\cos\theta + \tfrac{1}{2}\sin 2\theta - \theta\Big|_0^{\pi/2} = 2 - \frac{\pi}{2}.
\end{aligned}$$

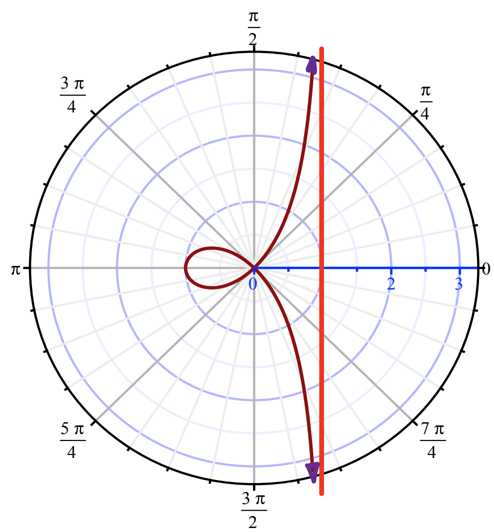


FIGURE 8.9. strophoid

8.5. Parametric equations

A curve in the plane can sometimes be conveniently described as $\gamma(t) = (x(t), y(t))$ for $a \leqslant t \leqslant b$ where x and y are functions of a *parameter* t. Generally x and y will be continuous, and frequently differentiable functions, of t. This can be convenient when a curve is nice, but has singularities as a function of x, or there are two or more y values for some (many) choices of x.

8.5.1. EXAMPLE. A simple example is a circle of radius r and centre (x_0, y_0) given by

$$\gamma(t) = (x_0 + r\cos t, y_0 + r\sin t) \quad \text{for} \quad 0 \leqslant t \leqslant 2\pi.$$

For $x_0 - r < x < x_0 + r$, there are two values of y for each x on the circle. Moreover the function $y = y_0 \pm \sqrt{r^2 - (x - x_0)^2}$ fails to be differentiable when $x = x_0 \pm r$. However $\dfrac{d\gamma}{dt} = (-r\sin t, r\cos t)$ is defined for all $t \in [0, 2\pi]$.

Let's look at various techniques to find $\dfrac{dy}{dx}$. The Cartesian formula for the curve is $(x - x_0)^2 + (y - y_0)^2 = r^2$. Thus

$$y - y_0 = \pm\sqrt{r^2 - (x - x_0)^2} \quad \text{thus} \quad y' = \frac{\mp(x - x_0)}{\sqrt{r^2 - (x - x_0)^2}} = -\frac{x - x_0}{y - y_0}.$$

By implicit differentiation of $(x - x_0)^2 + (y - y_0)^2 = r^2$, we get

$$2(x - x_0) + 2(y - y_0)y' = 0 \quad \text{thus} \quad y' = -\frac{x - x_0}{y - y_0}.$$

Finally with the parametric form $x(t) = x_0 + r\cos t$ and $y(t) = y_0 + r\sin t$,

$$x'(t) = -r\sin t = y_0 - y(t) \quad \text{and} \quad y'(t) = r\cos t = x(t) - x_0,$$

and thus

$$\frac{dy}{dx} = \frac{y'(t)}{x'(t)} = -\cot t = -\frac{x - x_0}{y - y_0}.$$

8.5.2. EXAMPLE. Other conics can be expressed nicely using parameters. The ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

can be expressed as

$$x(t) = a\cos t \quad \text{and} \quad y(t) = b\sin t \quad \text{for} \quad 0 \leqslant t \leqslant 2\pi.$$

The hyperbola

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

can be expressed using the hyberbolic trig functions as

$$x(t) = a\cosh t \quad \text{and} \quad y(t) = b\sinh t \quad \text{for} \quad -\infty < t < \infty.$$

We say that γ is a *closed curve* if $\gamma(a) = \gamma(b)$. The curve γ does not intersect itself if $\gamma(s) = \gamma(t)$ implies that $s = t$ or $\{s, t\} = \{a, b\}$. The curve γ is called C^1 if $x(t)$ are $y(t)$ are C^1 functions and we write $\gamma'(t) = (x'(t), y'(t))$, provided that if γ is a closed curve, then $\gamma'(a) = \gamma'(b)$. We say that γ is *piecewise* C^1 is the derivative is piecewise continuous, meaning that there are at most finitely many jump discontinuities in $x'(t)$ and $y'(t)$.

To compute the area enclosed in a closed curve γ, we use a special case of a result from multivariable calculus.

8.5.3. GREEN'S THEOREM. *Let $\gamma(t) = (x(t), y(t))$ for $a \leqslant t \leqslant b$ be a closed, piecewise C^1, curve that does not intersect itself. Assume that $x'(t) = 0$ for only finitely many values of $x(t)$. Then the area enclosed by γ is*

$$A = \Big| \int_a^b -y(t)x'(t)\, dt \Big|.$$

If the direction of γ is counterclockwise, the integral is positive.

PROOF. For convenience, we will assume that the curve is traversed in the counterclockwise direction, keeping the enclosed area on the left at all times. It is easy to see that if the integral is computed for the contrary direction, this just changes the sign.

The main idea is contained in the following special case. Suppose that $x(t)$ is strictly monotone increasing on $[a, c]$, constant on $[c, d]$, and strictly monotone decreasing on $[d, e]$ and constant on $[e, b]$. Then there are functions $f(x)$ and $g(x)$ on $[x(a), x(c)] = [x(e), x(d)]$ so that $\gamma(t) = (x(t), f(x(t))$ for $a \leqslant t \leqslant c$ and $\gamma(t) = (x(t), g(x(t))$ for $d \leqslant t \leqslant e$, and two vertical segments of the curve on the left and right sides. Since the curve is traversed counterclockwise, the first section

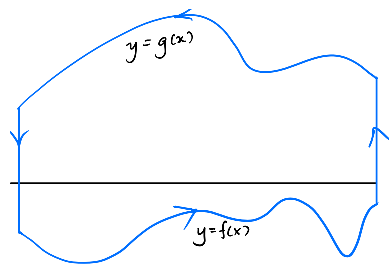


FIGURE 8.10. Special case

of the curve lies below the second section, so that $f(x) \leqslant g(x)$. Hence the area is

$$A = \int_{x(a)}^{x(c)} g(x) - f(x)\,dx.$$

Now compute the other integral, making the substitution $x(t)$ and $dx = x'(t)\,dt$. Note that $x'(t) = 0$ for $t \in [c, d] \cup [e, b]$.

$$\begin{aligned}\int_a^b -y(t)x'(t)\,dt &= -\int_a^c y(t)x'(t)\,dt - 0 - \int_d^e y(t)x'(t)\,dt - 0 \\ &= -\int_{x(a)}^{x(c)} f(x)\,dx - \int_{x(d)}^{x(e)} g(x)\,dx \\ &= \int_{x(a)}^{x(c)} g(x) - f(x)\,dx = A.\end{aligned}$$

Let D be the finite points at which $x'(t)$ is discontinuous. Let

$$C = \{c : x(t) = c \text{ and } x'(t) = 0 \text{ or } t \in D\}.$$

By hypothesis, this is a finite set, so we may write it as $c_0 < c_1 < \cdots < c_n$. Draw vertical lines $x = c_i$ for $0 \leqslant i \leqslant n$. When $c_{i-1} < x(t) < c_i$, we have that $x'(t)$ is continuous and non-zero. Thus $x'(t)$ does not change sign on any interval (u, v) with $x(t)$ in one of these intervals. So any *maximal* interval of this type maps onto an arc of the curve γ for which $x(t)$ is strictly monotone, and necessarily runs from c_{i-1} to c_i or vice versa. Thus the vertical strips cut γ into a finite number of arcs. It is now possible to obtain a finite number of closed curves γ_j which follow an arc from c_{i-1} to c_i, then taking a vertical line segment up to the next segment of the curve, which must run from from c_i to c_{i-1}, and then a vertical segment down to the beginning of the first arc. Altogether, the sum of these curves consists of the original curve γ together with a number of vertical segments. However any vertical

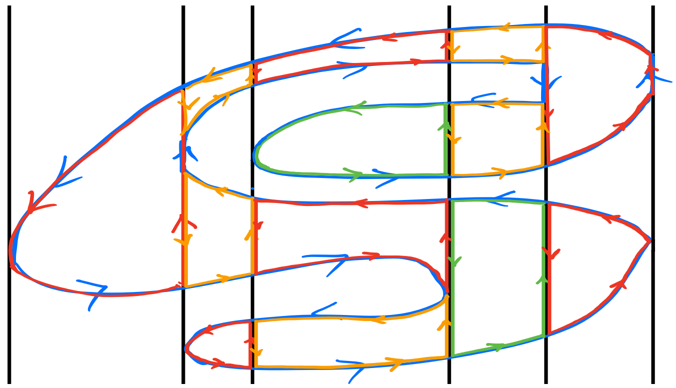


FIGURE 8.11. Cutting the curve

segment that is not part of γ occurs twice with opposite orientations in two of the smaller curves. Thus any integral over one is cancelled by the integral over the other. However in our situation, the integrals are always 0 because $x'(t) = 0$ on these segments.

The integral over γ_j of $-y(t)x'(t)\,dt$ is the area enclosed by γ_j by the first part of our proof. The total area enclosed by γ is the sum of these areas. As noted, the sum of the integrals over each γ_j yields the integral over γ. Thus the formula is verified. ■

8.5.4. EXAMPLE. The ellipse

$$x(t) = a\cos t \quad \text{and} \quad y(t) = b\sin t \quad \text{for} \quad 0 \leqslant t \leqslant 2\pi.$$

has area

$$A = \int_0^{2\pi} -b\sin t(-a\sin t)\,dt = ab\int_0^{2\pi} \sin^2 t\,dt = \pi ab.$$

I'll mention a simple trick here. When integrating $\sin^2 t$ over an interval which is a multiple of $\frac{\pi}{2}$, a simple symmetry argument shows that

$$\int_a^{a+n\pi/2} \sin^2 t\,dt = \int_a^{a+n\pi/2} \cos^2 t\,dt = \frac{1}{2}\int_a^{a+n\pi/2} \sin^2 t + \cos^2 t\,dt = \frac{n\pi}{4}.$$

A related and useful argument is that $\displaystyle\int_a^{a+2\pi} \sin t\,dt = 0.$

8.5.5. EXAMPLE. An *epicycloid* is the curve swept out by a point on the circumference of a small circle of radius r as it rolls around the circumference of a large circle of radius R. Let's set the centre of the large circle as the origin, so

$$\gamma_1(t) = (x_1(t), y(t)) = (R\cos t, R\sin t) \quad \text{for} \quad t \geqslant 0.$$

Let's start the small circle tangent to the point $\gamma(0) = (R, 0)$ with the special point at $(R + 2r, 0)$ corresponding to angle 0 to the positive axis. Roll the small circle counterclockwise around the large circle without slipping. At time t, the small circle is tangent to $\gamma(t)$. The arc of the large circle has length Rt. The small circle has traversed the same distance Rt. So it has rotated through an angle Rt/r relative to the tangent point $\gamma_1(t)$. Thus the angle from the horizontal is the sum of the angle t from the fact that the circle is now tangent at $\gamma_1(t)$ plus Rt/r for a total of $\frac{R+r}{r}t$. The centre of the small circle lies on the circle $\gamma_2(t) = ((R+r)\cos t, (R+r)\sin t)$ or radius $R + r$. Thus the new curve γ is given by

$$\begin{aligned}\gamma(t) &= \gamma_2(t) + (r\cos \tfrac{R+r}{r}t, r\sin \tfrac{R+r}{r}t)\\ &= \big((R+r)\cos t + r\cos \tfrac{R+r}{r}t, (R+r)\sin t + r\sin \tfrac{R+r}{r}t\big).\end{aligned}$$

In general, this curve is not periodic because R is not a rational multiple of r. Let's take $R = 5$ and $r = 1$. Then 5 rotations of the small circle will exactly make

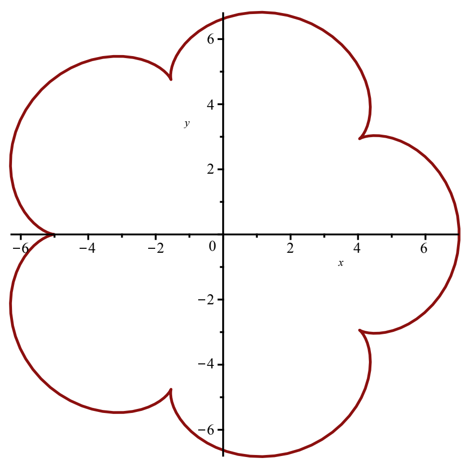


FIGURE 8.12. An epicycloid

a single turn around the larger one. So the curve should consist of 5 lobes. The formula is

$$\gamma(t) = (6\cos t + \cos 6t, 6\sin t + \sin 6t) \quad \text{for} \quad 0 \leqslant t \leqslant 2\pi.$$

The area of this figure is

$$\begin{aligned}A &= -\int_0^{2\pi} y(t)x'(t)\,dt = -\int_0^{2\pi} \big(6\sin t + \sin 6t\big)\big(-6\sin t - 6\sin 6t\big)\,dt\\ &= \int_0^{2\pi} 36\sin^2 t + 42\sin t\sin 6t + 6\sin^2 6t\,dt\\ &= 36\pi + 6\pi + 42\int_0^{2\pi} \cos 5t - \cos 7t\,dt = 42\pi.\end{aligned}$$

The formula for the perimeter of the epicycloid follows from the natural analogue arc length formula

$$\begin{aligned}
P &= \int_0^{2\pi} \sqrt{x'(t)^2 + y'(t)^2}\, dt \\
&= \int_0^{2\pi} \sqrt{(-6\sin t - 6\sin 6t)^2 + (6\cos t + 6\cos 6t)^2}\, dt \\
&= \int_0^{2\pi} \sqrt{36\sin^2 t + 72\sin t\sin 6t + 36\sin^2 6t + 36\cos^2 t + 72\cos t\cos 6t + 36\cos^2 6t}\, dt \\
&= \int_0^{2\pi} \sqrt{72 + 72\cos 5t}\, dt = 6\sqrt{2}\int_0^{2\pi} \sqrt{1+\cos 5t}\, dt \\
&= 6\sqrt{2}\int_0^{2\pi} \sqrt{1 + 2\cos^2 \tfrac{5t}{2} - 1}\, dt = 6\sqrt{2}\int_0^{2\pi} \sqrt{2}\,\big|\cos \tfrac{5t}{2}\big|\, dt \\
&= 120\int_0^{\pi/5} \cos \tfrac{5t}{2}\, dt = (120)\frac{2}{5}\sin \tfrac{5t}{2}\Big|_0^{\pi/5} = 48.
\end{aligned}$$

8.5.6. EXAMPLE. Folium of Descartes. Consider the curve

$$x^3 + y^3 = 3axy$$

where $a > 0$ is a constant. There is no obvious way to solve this equation. It is symmetric about the line $x = y$ because if (x, y) is a solution, so is (y, x).

Note that if $x = 0$ or $y = 0$, then $x = y = 0$. So we may look for solutions of the form $y = tx$. This yields $x^3(1 + t^3) = 3atx^2$. Thus

$$x = \frac{3at}{1+t^3} \quad \text{and} \quad y = \frac{3at^2}{1+t^3} \quad \text{for} \quad t \neq -1.$$

This must be a complete solution since $y/x = t \neq 0$ is defined whenever $x \neq 0$. Also $y = -x$ would yields $3axy = 0$ and thus $x = y = 0$, which corresponds to $t = 0$. This yields a parameterization $\gamma(t)$ of the solution set.

Notice that

$$\gamma\big(\frac{1}{t}\big) = \Big(\frac{3a/t}{(t^3+1)/t^3}, \frac{3a/t^2}{(t^3+1)/t^3}\Big) = \Big(\frac{3at^2}{t^3+1}, \frac{3at}{t^3+1}\Big).$$

This is the reflection of $\gamma(t)$ in the line $y = x$. We will also compute the derivative

$$x'(t) = \frac{3a(1-2t^3)}{(1+t^3)^2} \quad \text{and} \quad y'(t) = \frac{3at(2-t^3)}{(1+t^3)^2}.$$

To understand this solution, we consider three regions for t.

<u>Case 1.</u> $t \geqslant 0$. Then $\gamma(0) = (0, 0)$ and $\gamma'(0) = (3a, 0)$. Thus $\gamma(t)$ is tangent to the x-axis at the origin. When $t > 0$, $\gamma(t)$ remains bounded. The maximum value for $x(t)$ occurs when $x'(t) = 0$ of $t = 2^{-1/3}$ which yields $\gamma(2^{-1/3}) = (4^{1/3}a, 2^{1/3}a)$. Similarly $y'(t) = 0$ when $t = 2^{1/3}$ and $\gamma(2^{1/3}) = (2^{1/3}a, 4^{1/3}a)$ is the reflected

point. In between, γ crosses $y = x$ at $\gamma(1) = (\frac{3a}{2}, \frac{3a}{2})$. As $t \to +\infty$, $\gamma(t) \to (0,0)$ again. But $\frac{dy}{dx} = \frac{y'(t)}{x'(t)} = \frac{t(2-t^3)}{1-2t^3}$ approaches $+\infty$, indicating that the curve has a vertical tangent. Thus this section sweeps out a closed loop.

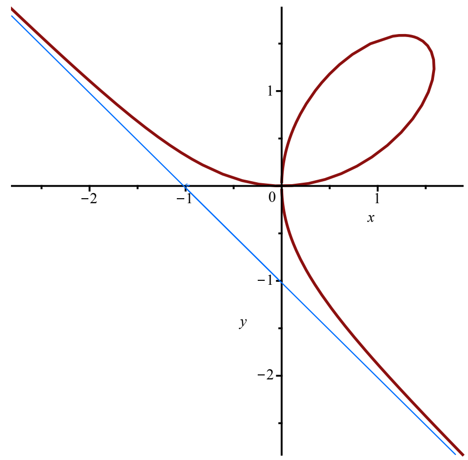


FIGURE 8.13. Folium of Descartes, $a = 1$

Case 2. $-1 < t < 0$. In this range, $x(t) < 0 < y(t)$, so the points lie in the second quadrant. As in Case 1, as $t \to 0^-$, the curve approaches $\gamma(0) = (0,0)$ tangent to the x axis. Moreover

$$\lim_{t \to -1} x(t) = -\infty \quad \text{and} \quad \lim_{t \to -1} y(t) = +\infty.$$

Observe that

$$\lim_{t \to -1^+} x(t) + y(t) = \lim_{t \to -1^+} \frac{3at(1+t)}{1+t^3} = \lim_{t \to -1^+} \frac{3at}{1-t+t^2} = -a.$$

Therefore the curve $\gamma(t)$ is asymptotic to the line $x + y + a = 0$ as $t \to -1^+$.

Case 3. $-\infty < t < -1$. This portion of the curve is the reflection of case 2, so it lies in the fourth quadrant. As $t \to -1^-$, the curve is asymptotic to the line $x + y + a = 0$. And as $t \to -\infty$, the curve approaches $(0,0)$ tangent to the y axis.

Let's compute the area of the closed loop. Substitute $u = t^3$ and $du = 3t^2\,dt$.

$$\begin{aligned} A &= -\int_0^\infty y(t)x'(t)\,dt = -\int_0^\infty \frac{3at^2}{1+t^3}\frac{3a(1-2t^3)}{(1+t^3)^2}\,dt \\ &= -9a^2 \int_0^\infty \frac{1-2t^3}{(1+t^3)^3}\,t^2\,dt = -3a^2 \int_0^\infty \frac{1-2u}{(1+u)^3}\,du \\ &= -3a^2 \int_0^\infty 3(1+u)^{-3} - 2(1+u)^{-2}\,du \\ &= 3a^3\big(\tfrac{3}{2}(1+u)^{-2} - 2(1+u)^{-1}\big)\Big|_0^\infty = \tfrac{3}{2}a^2. \end{aligned}$$

A famous problem of the ancient Greeks was to determine the closed curve of given perimeter which encloses the largest area. The solution, known as the *isoperimetric inequality*, is established in Appendix A.8.

Exercises for Chapter 8

1. Evaluate the following improper integrals when they exist.

(a) $\displaystyle\int_2^\infty \frac{dx}{x(\log x)^a}$ for $a > 0$.

(b) $\displaystyle\int_0^{\pi/2} \log \sin x \, dx$. HINT: substitute $u = \frac{\pi}{2} - x$ and combine.

2. Which of the following improper integrals exist? (Do not try to evaluate them exactly.)

(a) $\displaystyle\int_0^\infty \frac{1}{\sqrt{x}} \sin \tfrac{1}{x} \, dx$ (b) $\displaystyle\int_0^1 \frac{dx}{\ln x}$ (c) $\displaystyle\int_\pi^\infty \frac{\sin x}{\log x} \, dx$

3. Suppose that $f(x)$ and $g(x)$ are bounded continuous functions on $[0, \infty)$. If $\displaystyle\int_0^\infty f(x) \, dx$ exists as an improper integral, does it follow that $\displaystyle\int_0^\infty f(x)g(x) \, dx$ also exists? Give a proof or provide a counterexample.

4. Consider a region R bounded by the curve $y = \dfrac{1}{\sqrt{7x - 10 - x^2}}$ for $2 < x < 5$ together with the lines $x = 2$, $x = 5$ and $y = 0$. Compute the volume of the solid obtained by rotating R about the y-axis.

5. Consider the region S bounded by the curve $y = \log x$ for $0 < x \leqslant 1$ together with the lines $x = 0$ and $y = 0$. Compute the volume of the solid obtained by rotating S about the x-axis.

6. Consider the parabola P given by $y = ax^2$ for $a > 0$. At each point (x_0, y_0) on P, construct the *normal* line through (x_0, y_0) perpendicular to the tangent line, and consider the area of the sector of P cut off by this line.
(a) Find the minimal area of this sector.
(b) What are the slopes of the normal lines that minimize this area?

7. Compute the arc length of the curve $y = x^2$ from $x = 0$ to $x = 1$.

8.* Show that the arc length of the curve $y = x^p$ from $x = 0$ to $x = 1$ is an increasing function of p for $p \geqslant 1$. **Warning:** as far as I know, this cannot be done using the arc length formula. A geometric argument is needed.

9. Consider a curve given in polar coordinates by $r(\theta) = \dfrac{1}{1 + e\cos\theta}$, where $e \geqslant 0$.

(a) Show that the distance of each point on this curve to the line $x = \frac{1}{e}$ is a constant multiple of $r(\theta)$.

(b) When $e > 1$, show that the curve approaches two asymptotes, find them and sketch the curve. HINT: If the critical angles are $\pm\theta_0$, compute the vertical distance of the point of the curve at angle $\theta = \theta_0 + h$ to the line $\theta = \theta_0$, and take a limit.

(c) Observe that the curve is bounded if and only if $e < 1$. Show that the curve is an ellipse as follows: Let a be the midpoint between the two points intersecting the x-axis. Show that $(1 - e^2)(x - a)^2 + y^2$ is constant.

(d) What happens when $e = 1$?

10. A point on the circumference of a bicycle tire of radius R starts touching the road. As the bicycle rides along a straight line, the point on the tire sweeps out copies of a figure called a *cycloid*. Find the arc length of a single loop, and compute the area between the loop and the road.

CHAPTER 9

Series

9.1. Convergence of series

9.1.1. DEFINITION. A *series* is an infinite sum $\sum_{n=1}^{\infty} a_n$. The series *converges* or is *summable* if the sequence of partial sums $s_n = \sum_{i=1}^{n} a_i$ converges as $n \to \infty$. Otherwise the series *diverges*.

9.1.2. EXAMPLE. A *geometric series* has the form $\sum_{n=0}^{\infty} a_0 r^n$, where $a_0 \neq 0$ and $a_n = a_0 r^n$ for $n \geqslant 0$. (It is usual to begin at $n = 0$ here.) Let

$$s_n = \sum_{i=0}^{n-1} a_0 r^i = a_0 \frac{1 - r^n}{1 - r}.$$

This familiar formula comes from

$$(1 - r) \sum_{i=0}^{n-1} r^i = \sum_{i=0}^{n-1} r^i - r^{i+1} = \sum_{i=0}^{n-1} r^i - \sum_{i=1}^{n} r^i = 1 - r^n.$$

If $|r| < 1$, $\lim_{n\to\infty} s_n = \dfrac{a_0}{1-r}$, and the series converges. If $r = 1$, then $s_n = na_0$. This diverges. Likewise if $r = -1$, then $s_{2n+1} = a_0$ and $s_{2n} = 0$; so again the series diverges. If $|r| > 1$, then

$$\lim_{n\to\infty} |s_n| = |a_0| \lim_{n\to\infty} \frac{|r^n - 1|}{|r - 1|} = +\infty.$$

Thus this series diverges.

9.1.3. EXAMPLE. The *harmonic series* $\sum_{n\geqslant 1} \dfrac{1}{n}$ diverges. This follows from

$$s_{2^n} = 1 + \frac{1}{2} + \sum_{k=2}^{n} \sum_{i=2^{k-1}+1}^{2^k} \frac{1}{i} > 1 + \frac{1}{2} + \sum_{k=2}^{n} 2^{k-1} \frac{1}{2^k} = 1 + \frac{n}{2}.$$

In fact this shows that $s_n > \frac{1}{2} \log_2 n$. We will improve on this soon.

9.1.4. EXAMPLE. Consider $\sum_{n\geqslant 1}\dfrac{1}{n(n+3)}$. Observe that $\frac{1}{n(n+3)} = \frac{1}{3}\big(\frac{1}{n} - \frac{1}{n+3}\big)$. Therefore for $n \geqslant 3$,

$$s_n = \frac{1}{3}\sum_{i=1}^{n}\frac{1}{i} - \frac{1}{3}\sum_{i=4}^{n+3}\frac{1}{i} = \frac{1}{3}\Big(1 + \frac{1}{2} + \frac{1}{3} - \frac{1}{n+1} + \frac{1}{n+2} + \frac{1}{n+3}\Big).$$

This is an example of a *telesoping sum* because most terms cancel. Therefore

$$\sum_{n\geqslant 1}\frac{1}{n(n+3)} = \lim_{n\to\infty} s_n = \frac{1}{3}\Big(1 + \frac{1}{2} + \frac{1}{3}\Big) = \frac{11}{18}.$$

A basic result is the following. The harmonic series shows that the converse is false.

9.1.5. PROPOSITION. *If* $\sum\limits_{n=1}^{\infty} a_n$ *converges, then* $\lim\limits_{n\to\infty} a_n = 0$.

PROOF. If the series converges to L, then

$$L = \lim_{n\to\infty} s_n = \lim_{n\to\infty} s_{n+1}.$$

Therefore

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} s_{n+1} - s_n = L - L = 0. \qquad \blacksquare$$

The Cauchy criterion for a convergent sequence readily translates to series.

9.1.6. CAUCHY CRITERION FOR SERIES. *For a series* $\sum\limits_{n=1}^{\infty} a_n$*, the following are equivalent:*

(1) *The series converges.*

(2) *For all* $\varepsilon > 0$*, there is* $N \in \mathbb{N}$ *so that* $\Big|\sum\limits_{i=n+1}^{m} a_i\Big| < \varepsilon$ *for all* $N \leqslant n < m$.

PROOF. Suppose that the series converges to L. Then given $\varepsilon > 0$, there is an N so that if $n \geqslant N$, then $|s_n - L| < \varepsilon/2$. Therefore is $N \leqslant n < m$, we have

$$\Big|\sum_{i=n+1}^{m} a_i\Big| = |s_m - s_n| \leqslant |s_m - L| + |L - s_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Conversely, if (2) holds, for any $\varepsilon > 0$, we have an N so that $|s_m - s_n| < \varepsilon$ if $N \leqslant n < m$. This says that the sequence (s_n) is Cauchy. By completeness of $\mathbb{R}$, this sequence has a limit, say L. Therefore the series converges. $\blacksquare$

There is also a straightforward translation of the Monotone Convergence Theorem to series with positive terms.

9.1.7. Proposition. *If $a_n \geqslant 0$ for $n \geqslant 1$, then $\sum\limits_{n=1}^{\infty} a_n$ converges if and only if* $\sup_{n\geqslant 1} s_n < \infty$.

Proof. Since $a_n \geqslant 0$, the sequence of partial sums s_n are monotone increasing. Thus by the Monotone Convergence Theorem, the sequence converges only if it is bounded above, in which case, it converges to the supremum. Otherwise the series diverges to $+\infty$. ■

9.2. Tests for Convergence

Now we come to some new ideas that apply specifically to series.

9.2.1. Comparison Test. *Suppose that $\sum\limits_{n\geqslant 1} a_n$ and $\sum\limits_{n\geqslant 1} b_n$ are series such that $|a_n| \leqslant b_n$ for $n \geqslant 1$. Then if $\sum\limits_{n\geqslant 1} b_n$ converges, then so does $\sum\limits_{n\geqslant 1} a_n$.*

Proof. Suppose that $\sum\limits_{n\geqslant 1} b_n$ converges. Let $\varepsilon > 0$. By the Cauchy criterion, there is an N so that if $N \leqslant n < m$, then $\sum\limits_{i=n+1}^{m} b_i < \varepsilon$. Therefore

$$\Big|\sum_{i=n+1}^{m} a_i\Big| \leqslant \sum_{i=n+1}^{m} |a_i| \leqslant \sum_{i=n+1}^{m} b_i < \varepsilon.$$

Therefore $\sum\limits_{n\geqslant 1} a_n$ converges by the Cauchy criterion. ■

9.2.2. Example. Consider $\sum\limits_{n=1}^{\infty} (1 - \sqrt[n]{n})^n$. Observe that $f(x) = \ln(x^{1/x}) = \dfrac{\ln x}{x}$ has derivative $f'(x) = \dfrac{1-\ln x}{x^2} < 0$ for $x > e$. Therefore $\sqrt[n]{n} - 1 \leqslant \sqrt[3]{3} - 1 < \frac{1}{2}$ for $n \geqslant 3$. Hence

$$|a_n| = (\sqrt[n]{n} - 1)^n < 2^{-n} \quad \text{for} \quad n \geqslant 3.$$

Set $b_n = 2^{-n}$ for $n \geqslant 3$. Since $\sum\limits_{n=3}^{\infty} 2^{-n} < \infty$, it follows that $\sum\limits_{n=3}^{\infty} (1 - \sqrt[n]{n})^n$ converges by the comparison test. Of course, convergence is unaffected by the first few terms, so the original series converges.

Next we have a continuous version of the comparison test.

9.2.3. INTEGRAL TEST. *Let $f(x)$ be a positive, monotone decreasing function on $[1,\infty)$. Then $\sum\limits_{n=1}^{\infty} f(n)$ converges if and only if $\int_1^{\infty} f(x)\,dx < \infty$. Indeed,*

$$\sum_{n=2}^{\infty} f(n) \leqslant \int_1^{\infty} f(x)\,dx \leqslant \sum_{n=1}^{\infty} f(n).$$

PROOF. Take the integral $\int_1^{n+1} f(x)\,dx$. With partition $\mathcal{P} = \{1,2,3,\dots,n,n{+}1\}$, form the upper and lower Riemann sums. Since f is monotone decreasing, on the interval $[k, k+1]$, we have $f(k+1) \leqslant f(x) \leqslant f(k)$. Therefore

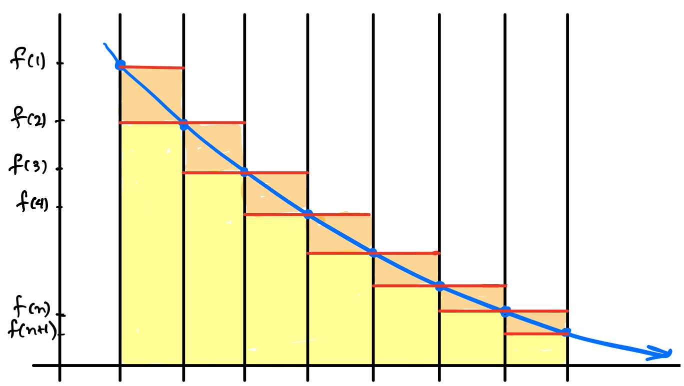


FIGURE 9.1. Integral test

$$L(f,\mathcal{P}) = \sum_{k=2}^{n+1} f(k) \leqslant \int_1^{n+1} f(x)\,dx \leqslant \sum_{k=1}^{n} f(k) = U(f,\mathcal{P}).$$

Now let $n \to \infty$. If the series converges, then

$$\int_1^{\infty} f(x)\,dx = \sup_{n\geqslant 1} \int_1^{n+1} f(x)\,dx \leqslant \sum_{n=1}^{\infty} f(n) < \infty.$$

Thus by the Monotone Convergence Theorem, the integral exists. Similarly if the integral exists,

$$\sum_{k=2}^{\infty} f(k) = \sup_{n\geqslant 1} \sum_{k=2}^{n+1} f(k) \leqslant \int_1^{\infty} f(x)\,dx < \infty.$$

Therefore $\sum\limits_{k=2}^{\infty} f(k)$ converges. Adding one term at the beginning does not affect convergence. Finally the estimates are part of the proof. ■

9.2.4. EXAMPLE. Consider $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ for $\alpha > 0$. The function $f(x) = x^{-\alpha}$ is monotone decreasing, so we can apply the integral test.

$$\int_1^\infty x^{-\alpha}\,dx = \begin{cases} \frac{x^{1-\alpha}}{1-\alpha}\Big|_1^\infty = +\infty & \text{if } 0 < \alpha < 1 \\ \ln x\,\Big|_1^\infty = +\infty & \text{if } \alpha = 1 \\ \frac{x^{1-\alpha}}{1-\alpha}\Big|_1^\infty = \frac{1}{\alpha-1} & \text{if } \alpha > 1. \end{cases}$$

Thus the series converges when $\alpha > 1$. For example, $\sum_{n=1}^{\infty} \dfrac{1}{n^2}$ and $\sum_{n=1}^{\infty} \dfrac{1}{n^{3/2}}$ converge and $\sum_{n=1}^{\infty} \dfrac{1}{\sqrt{n}}$ diverges.

9.2.5. EXAMPLE. Consider $\sum_{n=2}^{\infty} \frac{1}{n \ln n (\ln\ln n)^2}$. The function $f(x) = \frac{1}{x \ln x\,(\ln\ln x)^2}$ is monotone decreasing on $[3, \infty)$. We need to start at 3 because $\ln\ln e = 0$. Hence we can apply the integral test. Substitute $u = \ln\ln x$. Then $du = \frac{1}{x\ln x}\,dx$. Hence

$$\int_3^\infty \frac{1}{x\ln x(\ln\ln x)^2}\,dx = \int_{\ln\ln 3}^\infty u^{-2}\,du = -\frac{1}{u}\Big|_{\ln\ln 3}^\infty = \frac{1}{\ln\ln 3} < \infty.$$

Thus the series converges.

9.2.6. EXAMPLE. Consider the harmonic series again: $\sum_{n=1}^{\infty} \frac{1}{n}$. The argument in the proof of the integral test shows that

$$\sum_{k=2}^{n+1} \frac{1}{k} \leqslant \int_1^{n+1} \frac{1}{x}\,dx = \ln n{+}1 \leqslant \sum_{k=1}^{n} \frac{1}{k}.$$

The difference $\sum_{k=1}^{n} \frac{1}{k} - \ln n{+}1$ can be seen to be the sum of the areas of the regions

$$A_k = \{(x,y) : k \leqslant x \leqslant k+1,\ \ln x \leqslant y \leqslant \ln k\}$$

from 1 to n. Imagine translating these regions to regions B_k in the column between $x = 0$ and $x = 1$. They are disjoint because $B_k \subset [0,1] \times [\frac{1}{k+1}, \frac{1}{k}]$. Thus the total area of $\bigcup_{k\geqslant 1} B_k$ is less than 1. Because of the slope of $y = \frac{1}{x}$, it looks to be approximately half of the area. The limiting area exists by the Comparison test, since the areas $|A_k| = |B_k| < \frac{1}{k} - \frac{1}{k+1}$ and $\sum_{k=1}^{\infty} \frac{1}{k} - \frac{1}{k+1} = 1$ is a telescoping sum. The limiting value is known as *Euler's constant*, which has been determined numerically as

$$\gamma = 0.57721\,56649\,01532\,86060\,65120\,90082\,40243\,10421\,59335\,93992\,\ldots$$

Thus

$$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k} - \ln n = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k} - \ln n+1 = \gamma.$$

So for large n, $\sum\limits_{k=1}^{n} \frac{1}{k} \approx \ln n + \gamma$. It is unknown whether γ is irrational.

9.2.7. RATIO TEST. *Suppose that a series* $\sum\limits_{n=1}^{\infty} a_n$ *of non-zero terms satisfies* $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = r$ *exists. If* $|r| < 1$*, then the series converges. If* $|r| > 1$*, then the series diverges.*

PROOF. If $|r| < 1$, pick $|r| < R < 1$. There is an N so that if $n \geqslant N$,

$$\left|\frac{a_{n+1}}{a_n}\right| \leqslant R.$$

Therefore

$$|a_{N+k}| \leqslant R^k |a_N| \quad \text{or} \quad k \geqslant 1.$$

Since the geometric series $\sum\limits_{k\geqslant 0} |a_N| R^k$ converges, the Comparison test shows that $\sum\limits_{n=N}^{\infty} a_n$ converges. Thus $\sum\limits_{n=1}^{\infty} a_n$ converges.

If $|r| > R > 1$, there is an N so that

$$\left|\frac{a_{n+1}}{a_n}\right| \geqslant R \quad \text{or} \quad |a_{n+1}| \geqslant R|a_n|.$$

In this case, $|a_n|$ is increasing, so does not go to 0. Hence the series diverges. ■

9.2.8. REMARK. If $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = 1$, nothing can be said. For example if $a_n = n^{-\alpha}$, we have $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = 1$. The series converges for $\alpha > 1$ and diverges otherwise.

The following test is harder to use than the ratio test, but it is more powerful.

9.2.9. ROOT TEST. *Given a series* $\sum\limits_{n=1}^{\infty} a_n$*, define* $r = \limsup_{n\to\infty} |a_n|^{1/n}$*. If* $r < 1$*, the series converges; and if* $r > 1$*, the series diverges.*

PROOF. If $r < R < 1$, find N so that $|a_n|^{1/n} \leqslant R$ for all $n \geqslant N$. That is, $|a_n| \leqslant R^n$ and $\sum\limits_{n\geqslant N} R^n$ is a convergent geometric series. Therefore $\sum\limits_{n=1}^{\infty} a_n$

converges by the Comparison test. On the other hand, if $r > 1$, there are infinitely many terms a_{n_i} so that $|a_{n_i}| \geqslant 1$. Thus the series diverges. ■

Here is one more test that sometimes helps.

9.2.10. CAUCHY'S CONDENSATION TEST. *Suppose that a_n is a monotone decreasing sequence. Then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges*

PROOF. Note that

$$\sum_{n=2^k}^{2^{k+1}-1} a_n \leqslant 2^k a_{2^k} \quad \text{and} \quad \sum_{n=2^k+1}^{2^{k+1}} a_n \geqslant 2^k a_{2^{k+1}}.$$

Therefore

$$\sum_{n=1}^{\infty} a_n = \sum_{k=0}^{\infty} \sum_{n=2^k}^{2^{k+1}-1} a_n \leqslant \sum_{k=0}^{\infty} 2^k a_{2^k}$$

and

$$2\sum_{n=1}^{\infty} a_n = 2a_1 + \sum_{k=0}^{\infty} 2 \sum_{n=2^k+1}^{2^{k+1}} a_n \geqslant 2a_1 + \sum_{k=0}^{\infty} 2^{k+1} a_{2^{k+1}} \geqslant \sum_{k=0}^{\infty} 2^k a_{2^k}.$$

Thus one series converges if and only if the other does by the Comparison test. ■

9.2.11. EXAMPLE. Consider $\sum_{n=1}^{\infty} \frac{n!}{n^n}$. With $a_n = \frac{n!}{n^n}$, compute

$$\begin{aligned} \lim_{n\to\infty} \frac{a_{n+1}}{a_n} &= \lim_{n\to\infty} \frac{(n+1)!}{(n+1)^{n+1}} \frac{n^n}{n!} = \lim_{n\to\infty} \frac{(n+1)n^n}{(n+1)^{n+1}} \\ &= \lim_{n\to\infty} \left(\frac{n}{n+1}\right)^n = \frac{1}{e} < 1. \end{aligned}$$

Therefore this series converges by the ratio test.

9.2.12. EXAMPLE. Consider $\sum_{n=1}^{\infty} \frac{x^n}{n!}$. Compute

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{x^{n+1}}{(n+1)!} \frac{n!}{x^n} = \lim_{n\to\infty} \frac{x}{n+1} = 0.$$

Therefore this series converges by the ratio test for all values of x.

9.2.13. EXAMPLE. Consider $\sum\limits_{n=1}^{\infty} \dfrac{n^p}{p^n}$ for $p > 1$. Use the root test.

$$\limsup_{n\to\infty} \left(\frac{n^p}{p^n}\right)^{1/n} = \frac{1}{p}\big(\limsup_{n\to\infty} n^{1/n}\big)^p = \frac{1}{p} < 1.$$

Here we use the fact that $\lim\limits_{n\to\infty} \ln n^{1/n} = \lim\limits_{n\to\infty} \frac{\ln n}{n} = 0$ so that $\lim\limits_{n\to\infty} n^{1/n} = 1$. Therefore this series converges by the root test

9.2.14. EXAMPLE. Consider $\sum\limits_{n=1}^{\infty} \dfrac{1}{n(\ln n)^a}$ for $a > 0$. Use Cauchy's condensation test.

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = \sum_{k=0}^{\infty} 2^k \frac{1}{2^k (k \ln 2)^a} = \sum_{k=0}^{\infty} \frac{1}{(k \ln 2)^a}.$$

We have already analyzed this sum, and it converges if and only if $a > 1$.

Now consider $\sum\limits_{n=1}^{\infty} \dfrac{1}{n \ln n (\ln\ln n)^a}$. Set $b_n = \dfrac{1}{n \ln n (\ln \ln n)^a}$. Then

$$\sum_{k=0}^{\infty} 2^k b_{2^k} = \sum_{k=0}^{\infty} 2^k \frac{1}{2^k (k \ln 2)(\ln(k \ln 2))^a} = \frac{1}{\ln 2} \sum_{k=0}^{\infty} \frac{1}{k(\ln k + \ln 2)^a}.$$

By the previous example, this converges if and only of $a > 1$ as well.

9.2.15. EXAMPLE. Consider $\sum\limits_{n=1}^{\infty} \dfrac{1}{(\ln \ln n)^{\ln n}}$. Use Cauchy's condensation test.

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

Now apply the root test.

$$\limsup_{k\to\infty} \left(2^k \frac{1}{(\ln k + \ln\ln 2)^{k \ln 2}}\right)^{1/k} = \limsup_{k\to\infty} \frac{2}{(\ln k + \ln \ln 2)^{\ln 2}} = 0.$$

Therefore this series converges by Cauchy's condensation test and the root test.

9.3. Absolute and Conditional Convergence

9.3.1. DEFINITION. A series $\sum\limits_{n=1}^{\infty} a_n$ *converges absolutely* if $\sum\limits_{n=1}^{\infty} |a_n| < \infty$. A series which converges, but does not converge absolutely, is said to *converge conditionally*.

By comparing a_n with $|a_n|$, the Comparison test shows that absolutely convergent series converge.

9.3.2. ALTERNATING SERIES TEST. *If* a_n *is monotone decreasing and* $\lim_{n\to\infty} a_n = 0$*, then* $\sum_{n=1}^{\infty} (-1)^n a_n$ *converges, say to* L*. Moreover*

$$s_{2n+1} \leqslant L \leqslant s_{2n} \quad \textit{for} \quad n \geqslant 0.$$

Thus $|L - s_n| \leqslant |a_{n+1}|$.

PROOF. Observe that $s_{2n+1} = s_{2n} - a_{2n+1} \leqslant s_{2n}$. Moreover

$$s_{2n+2} = s_{2n} - (a_{2n+1} - a_{2n+2}) \leqslant s_{2n}$$

and

$$s_{2n-1} \leqslant s_{2n-1} + (a_{2n} - a_{2n+1}) = s_{2n+1}$$

for all $n \geqslant 1$. That is

$$s_1 \leqslant s_3 \leqslant s_5 \cdots \leqslant s_{2n+1} \leqslant s_{2n} \leqslant \cdots \leqslant s_4 \leqslant s_2.$$

The sequence $\{s_{2n-1}\}$ is monotone increasing, and bounded above. Therefore it converges, say to L, by the Monotone Convergence Theorem. Likewise $\{s_{2n}\}$ is monotone decreasing, and bounded below; and thus converges, say to M, also by the Monotone Convergence Theorem. Finally

$$M - L = \lim_{n\to\infty} s_{2n} - s_{2n+1} = \lim_{n\to\infty} -a_{2n+1} = 0.$$

Therefore $\sum_{n=1}^{\infty} (-1)^n a_n$ converges. Finally, $|L - s_n| \leqslant |s_{n+1} - s_n| = |a_{n+1}|$. ■

9.3.3. EXAMPLE. The alternating series test shows that $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ converges. This is known as the *alternating harmonic series*. However we know that $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. So this series converges conditionally, but not absolutely.

9.3.4. DEFINITION. A *rearrangement* of the series $\sum_{n=1}^{\infty} a_n$ is another series with the same terms in a different order, so there is a permutation π of $\mathbb{N}$ (a bijection of $\mathbb{N}$ onto itself) so that the new series is $\sum_{n=1}^{\infty} a_{\pi(n)}$.

9.3.5. EXAMPLE. Consider a rearrangement of the alternating harmonic series:

$$1 - \tfrac{1}{2} - \tfrac{1}{4} + \tfrac{1}{3} - \tfrac{1}{6} - \tfrac{1}{8} + \tfrac{1}{5} - \tfrac{1}{10} - \tfrac{1}{12} + \dots$$

Here we take the positive terms in order, and the negative terms in order, but take twice as many negative terms as positive terms at each stage. We can group the terms as follows

$$(1-\tfrac{1}{2})-\tfrac{1}{4}+(\tfrac{1}{3}-\tfrac{1}{6})-\tfrac{1}{8}+(\tfrac{1}{5}-\tfrac{1}{10})-\tfrac{1}{12}+\cdots=\frac{1}{2}\big(1-\tfrac{1}{2}+\tfrac{1}{3}-\tfrac{1}{4}+\tfrac{1}{5}-\tfrac{1}{6}+\dots\big)$$

From this, we can deduce that this rearrangement converges, but to a different limit, half of the value of the original limit.

9.3.6. THEOREM. *If $\sum_{n=1}^{\infty} a_n$ is absolutely convergent, then every rearrangement converges to the same value.*

PROOF. Let $L = \sum_{k=1}^{\infty} a_k$ and $M = \sum_{k=1}^{\infty} |a_k| < \infty$. Therefore given $\varepsilon > 0$, there is an N so that

$$\sum_{k=1}^{N} |a_k| > M - \frac{\varepsilon}{2} \quad \text{and so} \quad \sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

Hence if $n \geqslant N$,

$$|L - s_n| = \lim_{m\to\infty} |s_m - s_n| \leqslant \lim_{m\to\infty} \sum_{k=n+1}^{m} |a_k| \leqslant \sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

Let π be a permutation of $\mathbb{N}$, and let $K = \max\{\pi^{-1}(1), \pi^{-2}(2), \dots, \pi^{-1}(N)\}$. Suppose that $m \geqslant K$. Then $\{\pi(i) : 1 \leqslant i \leqslant m\} = \{1, 2, \dots, N\} \cup S_m$ for some subset $S_m \subset \{i : i > N\}$. Hence

$$\begin{aligned}\Big|\sum_{i=1}^{m} a_{\pi(i)} - L\Big| &= |s_N + \sum_{i\in S_m} a_i \ - L| \\ &\leqslant |s_N - L| + \sum_{i\in S_m} |a_i| \\ &< \frac{\varepsilon}{2} + \sum_{k=N+1}^{\infty} |a_k| < \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, $\sum_{i=1}^{\infty} a_{\pi(i)} = L$. ∎

9.3.7. EXAMPLE. Consider $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}$. Now $\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$ by the integral test, and therefore our series converges absolutely. Therefore

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} &= \sum_{n \text{ odd}} \frac{1}{n^2} - \sum_{n \text{ even}} \frac{1}{n^2} = \sum_{n\geqslant 1} \frac{1}{n^2} - 2\sum_{n \text{ even}} \frac{1}{n^2} \\ &= \sum_{n\geqslant 1} \frac{1}{n^2} - 2\sum_{n\geqslant 1} \frac{1}{(2n)^2} = \frac{1}{2}\sum_{n\geqslant 1} \frac{1}{n^2}.\end{aligned}$$

It is known that $\sum_{n\geqslant 1}\frac{1}{n^2}=\frac{\pi^2}{6}$. See Appendix A.9. Therefore $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^2}=\frac{\pi^2}{12}$.

Now we see what happens with rearrangements of conditionally convergent series.

9.3.8. LEMMA. *If $\sum_{n=1}^{\infty} a_n$ is a convergent series, let $b_1, b_2, \dots$ be the non-negative terms of the series and let $c_1, c_2, \dots$ be the negative terms of the series, in the order that they appear.*

(1) *If the series converges absolutely, then $\sum_{i\geqslant 1} b_i$ and $\sum_{i\geqslant 1} c_i$ converge absolutely.*
(2) *If the series converges conditionally, then $\sum_{i\geqslant 1} b_i$ and $\sum_{i\geqslant 1} c_i$ both diverge.*

PROOF. If $\sum_{n=1}^{\infty} a_n$ converges absolutely, let $x_n = \max\{a_m, 0\}$. Then $0 \leqslant x_n \leqslant a_n$, so that $\sum_{n=1}^{\infty} x_n$ converges absolutely. However this series is just the series $\sum_{i=1}^{\infty} b_i$ together with some extraneous 0 terms. Hence $\sum_{i\geqslant 1} b_i$ converges absolutely. Similarly, $\sum_{i\geqslant 1} c_i$ converges absolutely. Conversely if both $\sum_{i\geqslant 1} b_i$ and $\sum_{i\geqslant 1} c_i$ converge absolutely, then

$$\sum_{n\geqslant 1}|a_n| = \sum_{i\geqslant 1} b_i + \sum_{i\geqslant 1}|c_i| < \infty.$$

Thus $\sum_{n=1}^{\infty} a_n$ converges absolutely.

Suppose that $\sum_{i\geqslant 1} b_i = B$ converges absolutely but $\sum_{i\geqslant 1} c_i$ diverges to $-\infty$. Therefore for any $M > 0$, there is an J so that if $j \geqslant J$, then $\sum_{i=1}^{j} c_i < -M - B$. Pick N so that $c_J = a_N$. Then for $n \geqslant N$, there is some $j \geqslant J$ so that

$$s_n = \sum_{k=1}^{n} a_k = \sum_{i=1}^{n-j} b_i + \sum_{i=1}^{j} c_i < B - M - B = -M.$$

Since M is arbitrary, $\sum_{n=1}^{\infty} a_n$ diverges to $-\infty$. Similarly, if $\sum_{i\geqslant 1} b_i$ diverges and $\sum_{i\geqslant 1} c_i$ converges absolutely, then $\sum_{n=1}^{\infty} a_n$ diverges to $+\infty$.

Therefore, either both $\sum\limits_{i\geqslant 1} b_i$ and $\sum\limits_{i\geqslant 1} c_i$ converge absolutely or both diverge. By the first paragraph, conditionally convergent series are in the second situation. ■

9.3.9. REARRANGEMENT THEOREM. *If* $\sum\limits_{n=1}^{\infty} a_n$ *is conditionally convergent and* $L \in \mathbb{R}$*, then there is a rearrangement of the series which converges to* L.

PROOF. By the lemma, $\sum\limits_{i\geqslant 1} b_i = +\infty$ and $\sum\limits_{i\geqslant 1} c_i = -\infty$. Since the series converges,

$$0 = \lim_{n\to\infty} a_n = \lim_{i\to\infty} b_i = \lim_{i\to\infty} c_i.$$

Choose $m_0 \geqslant 1$ to be the least positive integer so that $\sum\limits_{i=1}^{m_0} b_i > L$. Then choose the least n_1 so that $\sum\limits_{i=1}^{m_0} b_i + \sum\limits_{i=1}^{n_1} c_i < L$. Since $\sum\limits_{i=1}^{m_0} b_i + \sum\limits_{i=1}^{n_1-1} c_i \geqslant L$, we have

$$L - c_{n_1} \leqslant \sum_{i=1}^{m_0} b_i + \sum_{i=1}^{n_1} c_i < L.$$

Now pick the least $m_1 > m_0$ so that $\sum\limits_{i=1}^{m_1} b_i + \sum\limits_{i=1}^{n_1} c_i > L$. As before

$$L < \sum_{i=1}^{m_1} b_i + \sum_{i=1}^{n_1} c_i \leqslant L + b_{m_1}.$$

Proceed recursively choosing the least integers $n_{j+1} > n_j$ so that

$$L - c_{n_{j+1}} \leqslant \sum_{i=1}^{m_j} b_i + \sum_{i=1}^{n_{j+1}} c_i < L.$$

and $m_{j+1} > m_j$ so that

$$L < \sum_{i=1}^{m_{j+1}} b_i + \sum_{i=1}^{n_{j+1}} c_i \leqslant L + b_{m_{j+1}}.$$

Our rearrangement is

$$b_1, \dots, b_{m_0}, c_1, \dots, c_{n_1}, b_{m_0+1}, \dots, b_{m_1}, c_{n_1+1}, \dots, c_{n_2}, \dots$$

By construction, the partial sums in the range $[m_j + n_j, m_j + n_{j+1}]$ lie in the interval $[L - c_{n_{j+1}}, L + b_{m_j}]$ and partial sums in the range $[m_j + n_{j+1}, m_{j+1}]$ lie in $[L - c_{n_{j+1}}, L + b_{m_{j+1}}]$. Because the terms tend to 0, given any $\varepsilon > 0$, there is a K so that if $i \geqslant K$, then $b_i < \varepsilon$ and $|c_i| < \varepsilon$. Once both m_j and n_j are greater than K, all of the partial sums beyond $m_j + n_j$ lie in $(L - \varepsilon, L + \varepsilon)$. That is, this rearrangement converges to L. ■

9.4. Dirichlet's Test

We prove one more convergence test.

9.4.1. SUMMATION BY PARTS LEMMA. *Let (x_i) and (y_i) be sequences. Define $X_n = \sum_{i=1}^{n} x_i$ and $Y_n = \sum_{i=1}^{n} y_i$. Then*

$$\sum_{i=1}^{n} x_i Y_i + X_i y_{i+1} = X_n Y_{n+1}.$$

PROOF. In the second line, there is a telescoping sum.

$$\begin{aligned} \sum_{i=1}^{n} x_i Y_i + X_i y_{i+1} &= \sum_{i=1}^{n} (X_i - X_{i-1}) Y_i + X_i (Y_{i+1} - Y_i) \\ &= \sum_{i=1}^{n} X_i Y_{i+1} - X_{i-1} Y_i \\ &= X_n Y_{n+1} - X_0 Y_1 = X_n Y_{n+1}. \end{aligned}$$

■

9.4.2. DEFINITION. A series $\sum_{i=1}^{\infty} a_i$ has *bounded partial sums* if there is a constant M so that $\left| \sum_{i=1}^{n} a_i \right| \leqslant M$ for all $n \geqslant 1$.

9.4.3. DIRICHLET'S TEST. *Let $\sum_{i=1}^{\infty} a_i$ be a series with bounded partial sums. Suppose that (b_i) is a monotone decreasing sequence with $\lim_{i\to\infty} b_i = 0$. Then $\sum_{i=1}^{\infty} a_i b_i$ converges.*

PROOF. Define $X_n = \sum_{i=1}^{n} a_i$, $Y_0 = 0$, $Y_n = b_n$ and $y_n = Y_n - Y_{n-1} = b_n$ for $n \geqslant 1$. Then by the Summation by parts Lemma,

$$\sum_{i=1}^{n} a_i b_i = \sum_{i=1}^{n} a_i Y_i = X_n Y_{n+1} - \sum_{i=1}^{n} X_i y_{i+1}.$$

By assumption, $|X_n| \leqslant M$ and thus

$$\lim_{n\to\infty} |X_n Y_{n+1}| \leqslant \lim_{n\to\infty} M b_{n+1} = 0.$$

The series $\sum_{i=1}^{\infty} y_i = \sum_{i=1}^{\infty} b_i - b_{i-1} = b_1$ is absolutely convergent. Therefore

$$\sum_{i=1}^{\infty} |X_i| y_{i+1} \leqslant M \sum_{i=1}^{\infty} y_{i+1} < \infty,$$

and hence $\sum_{i=1}^{\infty} X_i y_{i+1}$ converges absolutely. Hence

$$\begin{aligned}\sum_{i=1}^{\infty} a_i b_i &= \lim_{n\to\infty} \sum_{i=1}^{n} a_i b_i \\ &= \lim_{n\to\infty} X_n Y_{n+1} - \sum_{i=1}^{\infty} X_i y_{i+1} = -\sum_{i=1}^{\infty} X_i y_{i+1}.\end{aligned}$$

Thus this series converges. ■

9.4.4. EXAMPLE. Consider $\sum_{n=1}^{\infty} \frac{\sin n\theta}{n}$. If θ is an integer multiple of π, the sum is 0. Also this series is 2π-periodic and an odd function, so that it suffices to consider $\theta \in (0, \pi)$. Now we use the fact that $e^{i\theta} = \cos\theta + i \sin\theta$.

$$\Big|\sum_{k=1}^{n} \sin k\theta\Big| = \Big|\operatorname{Im} \sum_{k=1}^{n} e^{ik\theta}\Big| = \Big|\operatorname{Im}\Big(\frac{e^{i(n+1)\theta} - e^{i\theta}}{e^{i\theta} - 1}\Big)\Big| \leqslant \frac{2}{|e^{i\theta} - 1|} = \frac{1}{\sin\theta/2}.$$

Thus $\sum_{n=1}^{\infty} \sin n\theta$ has bounded partial sums. The series $b_n = \frac{1}{n}$ decreases monotonely to zero. Therefore by Dirichlet's test, $\sum_{n=1}^{\infty} \frac{\sin n\theta}{n}$ converges for all values of θ.

Since the proof of Dirichlet's test works by comparing the given series with an absolutely convergent one, it might be surprising that this series converges conditionally if θ is not an integer multiple of π. Suppose first that $\theta \in (0, \frac{\pi}{2}]$. Notice that for any k, if $\operatorname{dist}(k\theta, \pi\mathbb{Z}) \leqslant \frac{\theta}{2}$, then $\operatorname{dist}(k\theta, \pi\mathbb{Z}) \geqslant \frac{\theta}{2}$. Hence

$$\frac{|\sin(2k-1)\theta|}{2k-1} + \frac{|\sin(2k)\theta|}{2k} \geqslant \frac{\sin\theta/2}{2k}.$$

Therefore

$$\sum_{n=1}^{\infty} \frac{|\sin n\theta|}{n} = \sum_{k=1}^{\infty} \frac{|\sin(2k-1)\theta|}{2k-1} + \frac{|\sin(2k)\theta|}{2k} \geqslant \sin(\theta/2) \sum_{k=1}^{\infty} \frac{1}{2k} = +\infty$$

Thus this series is not absolutely convergent.

Similarly if $\theta \in (\frac{\pi}{2}, \pi)$, if $\operatorname{dist}(k\theta, \pi\mathbb{Z}) \leqslant \frac{\pi-\theta}{2}$, then $\operatorname{dist}(k\theta, \pi\mathbb{Z}) \geqslant \frac{\pi-\theta}{2}$. This also yields a conditionally convergent series.

Using Fourier series, one can show that the series converges to $\frac{\pi - \theta}{2}$ if θ is not an integer multiple of π.

Exercises for Chapter 9

1. Decide which of the following series converge absolutely, converge conditionally or diverge.

(a) $\sum_{n=1}^{\infty} \frac{\sqrt{n+1}-\sqrt{n}}{n^a}$

(b) $\sum_{n=1}^{\infty} (1-\sqrt[n]{n})^n$

(c) $\sum_{n=1}^{\infty} \frac{(-1)^n \arctan(n)}{n}$

(d) $\sum_{n=1}^{\infty} (-1)^n \frac{n^{42}}{(n+1)!}$

(e) $\sum_{n=1}^{\infty} \frac{1}{n^{1+1/n}}$

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

(g) $\sum_{n=2}^{\infty} \frac{\cos n\theta}{\log n}$

(h) $\sum_{n=3}^{\infty} \frac{(-1)^n}{\sqrt{n}(\log n)^2}$

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

(j) $\sum_{n=2}^{\infty} \frac{1}{(\ln n)^{\ln \ln n}}$

2. It is known that $\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$. Use this fact to compute $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^4}$.

3. Suppose that $a_n \leqslant b_n \leqslant c_n$ for all $n \geqslant 1$. If $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} c_n$ converge, does $\sum_{n=1}^{\infty} b_n$ converge? Prove it or provide a counterexample.

4. Suppose that $\sum_{n=1}^{\infty} a_n = A$ exists and (b_n) is a monotone sequence with limit B. Prove that $\sum_{n=1}^{\infty} a_n b_n$ converges. HINT: Find a way to apply Dirichlet's test.

5. Define an infinite product $\prod_{i=1}^{\infty} 1 + a_i$ as

$$\lim_{n\to\infty} \prod_{i=1}^{n} 1 + a_i = \lim_{n\to\infty} (1+a_1)(1+a_2)\cdots(1+a_n)$$

when this limit exists.

(a) Let $a_i \geqslant 0$. Prove that $\prod_{i=1}^{\infty} 1 + a_i$ converges if and only if $\sum_{i=1}^{\infty} a_i$ converges. HINT: take logs.

(b) Let $0 \leqslant a_i < 1$. Prove that $\prod_{i=1}^{\infty} 1 - a_i > 0$ if and only if $\sum_{i=1}^{\infty} a_i$ converges.

CHAPTER 10

Limits of Functions

10.1. Taylor Polynomials

In this section, we examine whether we can use higher derivatives to get a better approximation to a function by analogy with the tangent line.

10.1.1. DEFINITION. If $f(x)$ has n derivatives at a, the *Taylor polynomial* of degree n for f at a is

$$\begin{aligned} P_{n,a}(x) &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k. \end{aligned}$$

First we see that this polynomial has the same derivatives at a as f does up to the nth order.

10.1.2. LEMMA. *$P_{n,a}^{(k)}(a) = f^{(k)}(a)$ for $0 \leqslant k \leqslant n$.*

PROOF. It is straightforward to check that

$$\frac{d^k}{dx^k}(x-a)^l = \begin{cases} 0 & \text{if } l < k \\ k! & \text{if } l = k \\ l(l-1)\cdots(l+1-k)(x-a)^{l-k} & \text{if } l > k \end{cases}$$

whence

$$\frac{d^k}{dx^k}(x-a)^l\Big|_{x=a} = \begin{cases} 0 & \text{if } l < k \\ k! & \text{if } l = k \\ 0 & \text{if } l > k. \end{cases}.$$

Therefore $P_{n,a}^{(k)}(a) = \dfrac{f^{(k)}(a)}{k!}k! = f^{(k)}(a)$. ∎

In order to decide if the Taylor polynomial is a good approximation to $f(x)$ other than just as one approaches a, we introduce the error function

$$R_{n,a}(x) := f(x) - P_{n,a}(x).$$

Taylor's Theorem is a higher order Mean Value Theorem.

10.1.3. TAYLOR'S THEOREM. *Suppose that $f(x)$ has $n+1$ derivatives on $[a,b]$. Then there is an $x_0 \in (a,b)$ so that*

$$R_{n,a}(b) = f(b) - P_{n,a}(b) = \frac{f^{(n+1)}(x_0)(b-a)^{n+1}}{(n+1)!}.$$

PROOF. For each $t \in [a,b]$, let $P_{n,t}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k$ be the Taylor polynomial for f about t. Set

$$R(t) = f(b) - P_{n,t}(b) = f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(b-t)^k.$$

Note that $R(a) = R_{n,a}(b)$ and $R(b) = 0$. The following computation involves a *telescoping sum.*

$$\begin{aligned} R'(t) &= -\sum_{k=0}^{n} \frac{f^{(k+1)}(t)}{k!}(b-t)^k - \frac{f^{(k)}(t)}{k!}k(b-t)^{k-1} \\ &= -\sum_{k=0}^{n} \frac{f^{(k+1)}(t)}{k!}(b-t)^k + \sum_{k=1}^{n} \frac{f^{(k)}(t)}{(k-1)!}(b-t)^{k-1} \\ &= -\sum_{k=1}^{n+1} \frac{f^{(k)}(t)}{(k-1)!}(b-t)^{k-1} + \sum_{k=1}^{n} \frac{f^{(k)}(t)}{(k-1)!}(b-t)^{k-1} \\ &= -\frac{f^{(n+1)}(t)}{n!}(b-t)^n \end{aligned}$$

Now let $G(t) = R(t) - \left(\frac{b-t}{b-a}\right)^{n+1} R(a)$. Then $G(a) = R(a) - R(a) = 0$ and $G(b) = R(b) - 0 = 0$. By Rolle's Theorem, there is an $x_0 \in (a,b)$ so that

$$0 = G'(x_0) = -\frac{f^{(n+1)}(x_0)(b-x_0)^n}{n!} - \frac{(n+1)(b-x_0)^n}{(b-a)^{n+1}}R(a).$$

Solve for $R(a)$:

$$R_{n,a}(b) = R_n(a) = \frac{f^{(n+1)}(x_0)(b-a)^{n+1}}{(n+1)!}.$$ ■

10.1.4. COROLLARY. *If $f \in C^{n+1}[a,b]$, then $|R_{n,a}(x)| \leqslant C|x-a|^{n+1}$. And if $q(x)$ is a polynomial of degree at most n so that $|f(x) - q(x)| \leqslant C'|x-a|^{n+1}$ for some contant C', then $q(x) = P_{n,a}(x)$.*

PROOF. By hypothesis, $f^{(n+1)}$ is continuous on $[a,b]$. By the Extreme Value Theorem, $\max_{a\leqslant x\leqslant b}|f(x)| = M < \infty$. By Taylor's Theorem (with x in place of b), $C = \frac{M}{(n+1)!}$ works.

Now if $q(x)$ is another polynomial of degree at most n satisfying a similar inequality, then

$$|q(x) - P_{n,a}(x)| \leqslant |q(x) - f(x)| + |f(x) - P_{n,a}(x)| \leqslant (C' + c)|x-a|^{n+1}.$$

Write $q(x) = b_0 + b_1(x-a) + b_2(x-a)^2 + \cdots + b_n(x-a)^n$. Let k be the smallest integer at which the coefficients differ from the coefficients of $P_{n,a}(x)$, namely $c_k := \frac{1}{k!}f^{(k)}(a)$.. Then $q(x) - P_{n,a}(x) = (x-a)^k(b_k - c_k) + \cdots + (b_n - c_n)(x-a)^{n-k})$. Thus

$$\lim_{x\to a}\frac{|q(x)-P_{n,a}(x)|}{|x-a|^{n+1}} = \lim_{x\to a}\frac{|(b_k-c_k)+\cdots+(b_n-c_n)(x-a)^{n-k}|}{|x-a|^{n+1-k}} = +\infty.$$

This contradicts the estimate above. Hence $q(x) = P_{n,a}(x)$. ■

10.1.5. EXAMPLE. Let $f(x) = e^x$ and $a = 0$. Then $f^{(n)}(x) = e^x$ for all $n \geqslant 1$. So $f^{(n)}(0) = 1$. Therefore $P_{n,0}(x) = \sum_{k=0}^{n}\frac{x^k}{k!}$. For any $x \in \mathbb{R}$, Taylor's Theorem provides an x_0 between 0 and x so that

$$\Big|e^x - \sum_{k=0}^{n}\frac{x^k}{k!}\Big| = \frac{|f^{(n+1)}(x_0)||x|^{n+1}}{(n+1)!} \leqslant \frac{\max\{e^x,1\}|x|^{n+1}}{(n+1)!}.$$

Now $\lim_{n\to\infty}\frac{|x|^{n+1}}{(n+1)!} = 0$ because for $|x| \leqslant N$ and $n = 2N + k$,

$$\frac{|x|^n}{n!} = \frac{|x|^{2N}}{(2N)!}\frac{|x|}{2N+1}\cdots\frac{|x|}{2N+k} < \frac{|x|^{2N}}{(2N)!}\frac{1}{2^k} \to 0.$$

Therefore this sum converges to e^x as $n \to \infty$, so that

$$e^x = \sum_{k=0}^{\infty}\frac{x^n}{n!} \quad \text{and} \quad e = \sum_{k=0}^{\infty}\frac{1}{n!}.$$

If we take $n = 13$, we get $\left|e - \sum_{k=0}^{13}\frac{1}{n!}\right| < \frac{e}{14!} < 4\cdot 10^{-11}$, which yields 10 decimals accuracy.

This is a poor way to find e^x if x is large, and even for $x = 1$. Notice that

$$\big|e^{1/16} - \sum_{k=0}^{10}\frac{1}{n!2^{4k}}\big| < \frac{e^{1/16}}{11!2^{44}} = \varepsilon < 1.6\cdot 10^{-21}.$$

Set $a = \sum_{k=0}^{10}\frac{1}{n!2^{4k}}$. Then $e^{1/16} - \varepsilon < a < e^{1/16}$, so that

$$e > a^{16} = \big(((a^2)^2)^2\big)^2 > (e^{1/16} - \varepsilon)^{16} > e - 16e^{15/16}\varepsilon > e - 7\cdot 10^{-20}.$$

This has the same number of computations, but has 19 decimals of accuracy.

10.1.6. EXAMPLE. Let $f(x) = \sin x$, $g(x) = \cos x$ and $a = 0$. The derivatives are periodic:

$$\begin{aligned} f'(x) &= \cos x \quad \text{and} \quad g'(x) = -\sin x \\ f^{(2)}(x) &= -\sin x \quad \text{and} \quad g^{(2)}(x) = -\cos x \\ f^{(2n)}(x) &= (-1)^n \sin x \quad \text{and} \quad g^{(2n)}(x) = (-1)^n \cos x \\ f^{(2n+1)}(x) &= (-1)^n \cos x \quad \text{and} \quad g^{(2n+1)}(x) = (-1)^{n+1} \sin x. \end{aligned}$$

So $f^{(2n)}(0) = 0 = g^{(2n+1)}(0)$; and $f^{(2n+1)}(0) = (-1)^n = g^{(2n)}(0)$. The Taylor polynomials for $\sin x$ are $P_{2n-1,0}(x) = P_{2n,0}(x)$ where

$$P_{2n,0}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!} = \sum_{k=0}^{n-1}(-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Likewise the Taylor polynomials for $\cos x$ are $Q_{2n,0}(x) = Q_{2n+1,0}(x)$ where

$$Q_{2n+1,0}(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} = \sum_{k=0}^{n}(-1)^k \frac{x^{2k}}{(2k)!}.$$

By Taylor's Theorem, the remainders have the form

$$R_{2n,0}(x) = \sin x - P_{2n,0}(x) = \frac{f^{(2n+1)}(x_0)x^{2n+1}}{(2n+1)!} = (-1)^{n+1}\cos x_0 \frac{x^{2n+1}}{(2n+1)!}.$$

Therefore $|R_{2n,0}(x)| \leqslant \frac{|x|^{2n+1}}{(2n+1)!}$. As in the previous example, we know that this converges to 0. Therefore

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Convergence is slow for a while if x is large, so you should always use trig identities to manipulate things so that x is small. For example, if x is 1°, which is $\frac{\pi}{180}$ in radians, then

$$\sin \tfrac{\pi}{180} \approx \tfrac{\pi}{180} - \tfrac{\pi^3}{6(180)^3} + \tfrac{\pi^5}{120(180)^5} - \tfrac{\pi^7}{7!(180)^7} + \tfrac{\pi^9}{9!(180)^9}$$

with an error of at most $\frac{\pi^{11}}{11!(180)^{11}} < 5 \cdot 10^{-22}$. The real issue will be computing powers of π.

Similarly, the error for $\cos x$ is given as

$$\cos x - Q_{2n+1,0} = \frac{f^{(2n+2)}(x_0)x^{2n+2}}{(2n+2)!} = (-1)^{n+1}\cos x_0 \frac{x^{2n+2}}{(2n+2)!}.$$

Thus $|\cos x - Q_{2n+1,0}| \leqslant \frac{|x|^{2n+2}}{(2n+2)!}$. Again we conclude that the error tends to 0 for any value of x. Thus

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}.$$

If we take $f(x) = \sin x$ and $a = \pi/6$, then

$$P_{3,\pi/6}(x) = \frac{1}{2} + \frac{\sqrt{3}}{2}(x - \frac{\pi}{6}) - \frac{1}{2}\frac{1}{2}(x - \frac{\pi}{6})^2 - \frac{\sqrt{3}}{2}\frac{1}{6}(x - \frac{\pi}{6})^3.$$

This is a better starting point if you want to estimate $\sin(31^\circ) = \sin\left(\frac{\pi}{6} + \frac{\pi}{180}\right)$.

10.1.7. Example. These examples give a false impression about how well Taylor polynomials work. Consider $f(x) = \tan^{-1}(x)$. The derivatives get progressively more complicated. But there is a way around the problem. Note that f is an odd function, and

$$f'(x) = \frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + x^8 - \cdots = \sum_{k=0}^{\infty} (-1)^k x^{2k}.$$

This is a geometric series, and it converges when $|x| < 1$, and diverges if $|x| \geqslant 1$. Let

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

Therefore

$$|f'(x) - Q(x)| = \frac{|x|^{2n+2}}{1+x^2} \leqslant |x|^{2n+2}.$$

By Corollary 10.1.4, $Q(x)$ is the $2n+1$st Taylor polynomial of $f'(x)$ at $a = 0$.

It follows immediately from the definition that if $P_{n,a}(x)$ is the nth Taylor polynomial for $f(x)$, then $P'_{n,a}(x)$ is the Taylor polynomial for $f'(x)$ of degree $n-1$. Since $f(x) = 0$, it follows that

$$P_{2n+2,0} = \sum_{k=0}^{n} \frac{(-1)^k}{2k+1} x^{2k+1} = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots + \frac{(-1)^n}{2n+1} x^{2n+1}.$$

Rather than applying Taylor's estimate for the error, which we can't do easily since we don't know the higher derivatives, we instead apply the Mean Value Theorem to $f(x) - P_{2n+2,0}(x)$ on $[0, x]$. There is an $x_0 \in (0, x)$ so that

$$\begin{aligned}\frac{|f(x) - P_{2n+2,0}(x)|}{|x|} &= \frac{|(f(x) - P_{2n+2,0}(x)) - (f(0) - P_{2n+2,0}(0))|}{|x|} \\ &= |f'(x_0) - P'_{2n+2,0}(x_0)| = |f'(x_0) - Q(x_0)| < |x|^{2n+2}.\end{aligned}$$

For example, $\tan^{-1}(\frac{1}{10}) = \frac{1}{10} - \frac{1}{3000} + \frac{1}{500000} - \frac{1}{70000000}$ is within $\frac{1}{9 \cdot 10^9}$. In your homework, you will be asked to verify that

$$\frac{\pi}{4} = 4\tan^{-1}\frac{1}{5} - \tan^{-1}\frac{1}{239}.$$

Using this, you can get a formula for π.

The function $\tan^{-1}(x)$ is defined on the whole real line, but the Taylor polynomials only approximate $f(x)$ when $|x| < 1$. This is fairly typical behaviour.

10.1.8. DEFINITION Big O and little o Notation.. We say that $f(x)$ is $O(g)$ as $x \to a$ if there is a constant C so that $|f(x)| \leqslant C|g(x)|$ for $x \in (a-\delta, a+\delta)$. We write $f(x) = O(g(x))$. Also $f(x)$ is $o(g)$ as $x \to a$ if $\lim_{x\to a} \frac{f(x)}{g(x)} = 0$. We write $f(x) = o(g(x))$.

Normally $g(x)$ will go to 0 as $x \to a$, and $f = O(g)$ means that it goes to 0 at the same rate or faster. And $f = o(g)$ means that f goes to 0 faster than g.

It is not hard to show that if $f_i = O(g_i)$ near $x = a$, then $f_1 f_2 = O(g_1 g_2)$ and $f_1 + f_2 = O(\max\{g_1, g_2\}$. Sometimes division is possible if the denominator is closely related to the numerator. For example $\frac{O((x-a)^n)}{(x-a)^k} = O((x-a)^{n-k})$.

The Corollary 10.1.4 says that $f(x) = P_{n,a}(x) + O((x-a)^{n+1})$.

10.1.9. EXAMPLE. Let $f(x) = \tan x$ and $a = 0$. This is another function whose derivatives get complicated quickly. Note that f is odd, so that $f^{(2n)}(0) = 0$ for $n \geqslant 1$.

$$\begin{aligned}
f'(x) &= \sec^2 x &&\text{and}\quad f'(0) = 1\\
f''(x) &= 2\sec^2 x \tan x &&\text{and}\quad f''(0) = 0\\
f^{(3)}(x) &= 4\sec^2 x \tan^2 x + 2\sec^4 x &&\text{and}\quad f^{(3)}(0) = 2\\
f^{(4)}(x) &= 8\sec^2 x \tan^3 x + 16\sec^4 x \tan x &&\text{and}\quad f^{(4)}(0) = 0\\
f^{(5)}(x) &= 16\sec^2 x \tan^4 x + 88\sec^4 x \tan^2 x + 16\sec^6 x &&\text{and}\quad f^{(5)}(0) = 16.
\end{aligned}$$

Therefore $P_{5,0(x)} = P_{6,0}(x) = x + \frac{1}{3}x^3 - \frac{2}{15}x^5$.

We will illustrate how to use big O arithmetic to find the Taylor polynomials.

$$\tan x = \frac{\sin x}{\cos x} = \frac{x - \frac{1}{3}x^3 + \frac{1}{120}x^5 - \frac{1}{7!}x^7 + O(x^9)}{1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 + O(x^8)}$$

In order to invert the expression for $\cos x$, we find the polynomial which multiplies it to get $1 + O(x^8)$, which can be done by long division. We get

$$\big(1 - \tfrac{1}{2}x^2 + \tfrac{1}{24}x^4 - \tfrac{1}{720}x^6 + O(x^8)\big)\big(1 + \tfrac{1}{2}x^2 + \tfrac{5}{24}x^4 + \tfrac{61}{720}x^6 + O(x^8)\big) = 1 + O(x^8).$$

Therefore

$$\begin{aligned}\tan x &= \left(x - \tfrac{1}{3}x^3 + \tfrac{1}{120}x^5 - \tfrac{1}{7!}x^7 + O(x^9)\right)\left(1 + \tfrac{1}{2}x^2 + \tfrac{5}{24}x^4 + \tfrac{61}{720}x^6 + O(x^8)\right)\\ &= x + \tfrac{1}{3}x^3 + \tfrac{2}{15}x^5 + \tfrac{17}{315}x^7 + O(x^9).\end{aligned}$$

Note that when doing the multiplication, we only need to keep track of terms of order at most 7 (since 8 never occurs). By Corollary 10.1.4, we have

$$P_{8,0}(x) = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \frac{17}{315}x^7.$$

10.1.10. Example. Compute $\lim_{x\to 0} \cot^2 x - \frac{1}{x^2}$.

$$\begin{aligned}\lim_{x\to 0} \cot^2 x - \frac{1}{x^2} &= \lim_{x\to 0} \frac{1}{\tan^2 x} - \frac{1}{x^2}\\ &= \lim_{x\to 0} \frac{1}{(x + \frac{1}{3}x^3 + O(x^5))^2} - \frac{1}{x^2}\\ &= \lim_{x\to 0} \frac{1}{x^2 + \frac{2}{3}x^4 + O(x^6)} - \frac{1}{x^2}\\ &= \lim_{x\to 0} \frac{1 - (1 + \frac{2}{3}x^2 + O(x^4))}{x^2(1 + \frac{2}{3}x^2 + O(x^4))}\\ &= \lim_{x\to 0} \frac{-\frac{2}{3} + O(x^2)}{1 + \frac{2}{3}x^2 + O(x^4)} = -\frac{2}{3}.\end{aligned}$$

10.1.11. Example. Compute $\lim_{x\to 0} \dfrac{(1+x)^{1/x} - e}{x}$.

First $(1+x)^{1/x} = e^{\frac{\ln(1+x)}{x}}$. We know that $e^u = 1 + u + \frac{1}{2}u^2 + O(u^3)$. We find the Taylor polynomial for $f(x) = \ln(1+x)$ at $a = 0$. Then $f(0) = 0$, $f'(x) = \frac{1}{1+x}$ and $f'(0) = 1$ and $f''(x) = \frac{-1}{(1+x)^2}$ and $f''(0) = -1$. So $\ln(1+x) = x - \frac{1}{2}x^2 + O(x^3)$, and thus $\dfrac{\ln(1+x)}{x} = 1 - \frac{1}{2}x + O(x^2)$. Therefore

$$\begin{aligned}(1+x)^{1/x} &= e^{\frac{\ln(1+x)}{x}} = ee^{-\frac{1}{2}x + O(x^2)}\\ &= e\Big(1 + (-\frac{1}{2}x + O(x^2)) + O\big((-\frac{1}{2}x + O(x^2))^2\big)\Big)\\ &= e(1 - \frac{1}{2}x + O(x^2)) = e - \frac{e}{2}x + O(x^2).\end{aligned}$$

Consequently.

$$\lim_{x\to 0} \frac{(1+x)^{1/x} - e}{x} = \lim_{x\to 0} \frac{(e - \frac{e}{2}x + O(x^2)) - e}{x} = \lim_{x\to 0} -\frac{e}{2} + O(x) = -\frac{e}{2}.$$

These two limits are much easier with Taylor polynomials than by L'Hôpital's rule.

10.1.12. EXAMPLE. Consider $f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$. This is a rather strange example. Observe that for $n \geqslant 1$, with the substitution $u = -1/x^2$,

$$\lim_{x\to 0} \frac{f(x)}{x^{2n}} = \lim_{x\to 0} \frac{e^{-1/x^2}}{x^{2n}} = \lim_{u\to -\infty} e^u u^n = 0.$$

Therefore $f(x) = o(x^{2n})$ as $x \to 0$. By Corollary 10.1.4, $P_{2n-1,0}(x) = 0$. In particular, we have that $f^{(n)}(0) = 0$ for all $n \geqslant 0$. The function $f(x)$ is *extremely flat* at $x = 0$.

Notice that the Taylor series $\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} 0x^n = 0$ converges quickly for *all* $x \in \mathbb{R}$. However the sum equals $f(x)$ *only* at the point $x = 0$.

10.2. Uniform limits

In the section on Taylor polynomials, we were sometimes able to show that the infinite series of functions converges to our function. Exactly how this happens can be delicate, and like the previous section, it is better when the estimates for convergence are uniform over the domain. Note that in the ε–N version, there is an interchange of when x and δ are determined.

10.2.1. DEFINITION. Suppose that $f, f_n : [a, b] \to \mathbb{R}$ are functions.

We say that f_n *converges pointwise* to a function $f(x)$ if for each $x \in [a, b]$, $\lim_{n\to\infty} f_n(x) = f(x)$. This means that for any $\varepsilon > 0$ and $x \in [a, b]$, there is an N so that if $n \geqslant N$, then $|f_n(x) - f(x)| < \varepsilon$.

We say that f_n *converges uniformly* to a function $f(x)$ if

$$\lim_{n\to\infty} \sup_{a\leqslant x\leqslant b} |f_n(x) - f(x)| = 0.$$

This means that for any $\varepsilon > 0$, there is an N so that if $x \in [a, b]$ and $n \geqslant N$, then $|f_n(x) - f(x)| < \varepsilon$.

10.2.2. EXAMPLE. Let $f_n(x) = x^n$ for $x \in [0, 1]$. Then

$$\lim_{n\to\infty} x^n = f(x) = \begin{cases} 0 & \text{if } 0 \leqslant x < 1 \\ 1 & \text{if } x = 1. \end{cases}$$

So f_n converge pointwise to f. The convergence is not uniform because

$$\sup_{0\leqslant x\leqslant 1} |f_n(x) - f(x)| = \sup_{0\leqslant x< 1} x^n = 1$$

for every $n \geqslant 1$. Notice that each $f_n(x)$ is continuous, but the limit function is not.

10.2.3. EXAMPLE. Let $f_n(x) = \frac{1}{n}\sin nx$ on $[0, 2\pi]$. Then

$$\max_{0 \leqslant x \leqslant 2\pi} |f_n(x)| = \frac{1}{n}.$$

Therefore $f_n(x) \to 0$ uniformly on $[0, 2\pi]$. So $f = 0$ is the uniform limit. However $f_n'(x) = \cos nx$, and $\|f_n' - f'\|_\infty = 1$ for all $n \geqslant 1$. Hence the derivatives of a uniformly convergent sequence need not be well-behaved.

10.2.4. EXAMPLE. Let $f_n(x) = \begin{cases} n^2 x & \text{if } 0 \leqslant x \leqslant \frac{1}{n} \\ n^2(\frac{2}{n} - x) & \text{if } \frac{1}{n} \leqslant x \leqslant \frac{2}{n} \\ 0 & \text{if } \frac{2}{n} \leqslant x \leqslant 1. \end{cases}$ for $n \geqslant 2$.

Then f_n are continuous, and

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{if } x = 0 \\ 0 & \text{if } x > 0 \quad \text{because } x > \frac{2}{n} \text{ for large } n. \end{cases}$$

Therefore f_n converges pointwise to $f(x) = 0$. This limit is continuous. Nevertheless,

$$\sup_{0 \leqslant x \leqslant 1} |f_n(x) - f(x)| = \sup_{0 \leqslant x \leqslant 1} f_n(x) = f_n(\tfrac{1}{n}) = n.$$

Therefore the convergence is not uniform.

Now consider $\displaystyle\lim_{n\to\infty} \int_0^1 f_n(x)\,dx$. The integral computes the area of a triangle of height n and base $\frac{2}{n}$. Thus

$$\int_0^1 f_n(t)\,dt = \frac{1}{2}(n)\frac{2}{n} = 1.$$

Therefore

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,dx = 1 \neq 0 = \int_0^1 f(x)\,dx$$

Thus the limit of the integrals is not equal to the integral of the limit.

An easy variation is $g_n(x) = \begin{cases} n^3 x & \text{if } 0 \leqslant x \leqslant \frac{1}{n} \\ n^3(\frac{2}{n} - x) & \text{if } \frac{1}{n} \leqslant x \leqslant \frac{2}{n} \\ 0 & \text{if } \frac{2}{n} \leqslant x \leqslant 1. \end{cases}$ for $n \geqslant 2$. This sequence still converges pointwise to $f = 0$, but now the integrals diverge.

The advantage of uniform convergence is explained in the next result.

10.2.5. THEOREM. *Suppose that $f_n : [a, b] \to \mathbb{R}$ are continuous functions which converge uniformly to a function $f(x)$. Then f is continuous.*

PROOF. Let $\varepsilon > 0$. Pick N so that $\sup_{a \leqslant x \leqslant b} |f_n(x) - f(x)| \leqslant \frac{\varepsilon}{3}$. Since f_n is continuous on a closed bounded interval, it is uniformly continuous by Theorem 4.6.3. Thus there is some $\delta > 0$ so that $x, y \in [a, b]$ and $|x - y| < \delta$ implies $|f_n(x) - f_n(y)| < \frac{\varepsilon}{3}$. Compute

$$\begin{aligned} |f(x) - f(y)| &= |f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y)| \\ &\leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Therefore $f(x)$ is (uniformly) continuous. ■

10.2.6. EXAMPLE. We shows in Example 10.1.5 that $e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$. This is pointwise convergence of the partial sums $f_n(x) = \sum_{k=0}^{n} \frac{1}{k!} x^k$. But in fact, we showed more. We had estimates from Taylor's Theorem

$$|e^x - f_n(x)| = \left| e^x - \sum_{k=0}^{n} \frac{1}{k!} x^k \right| \leqslant \frac{e|x|^{n+1}}{(n+1)!}.$$

The convergence is not uniform on $(-\infty, \infty)$ because the terms $\frac{1}{k!} x^k$ are not uniformly small if x can be arbitrarily large. However if we restrict the domain to $[-R, R]$ for any R, we obtain

$$\sup_{|x| \leqslant R} |e^x - f_n(x)| \leqslant \frac{eR^{n+1}}{(n+1)!} \to 0$$

as $n \to \infty$. Thus the convergence is uniform on $[-R, R]$. Frequently this is the best one can do on an unbounded domain.

10.2.7. EXAMPLE. Let $f_n(x) = \sum_{k=0}^{n} (-1)^k x^{2k}$ for $x \in (-1, 1)$. We can sum this geometric series as in Example 10.1.7 to get

$$f_n(x) = \frac{1 + (-1)^n x^{2n+2}}{1 + x^2}.$$

This converges pointwise to $f(x) = \dfrac{1}{1 + x^2}$.

$$\sup_{-1 < x < 1} |f_n(x) - f(x)| = \sup_{-1 < x < 1} \frac{|x|^{2n+2}}{1 + x^2} = \frac{1}{2}.$$

This does not go to 0, so the convergence is not uniform. However, if $0 < r < 1$, and we restrict our domain to $[-r, r]$, then

$$\sup_{-r \leqslant x \leqslant r} |f_n(x) - f(x)| = \sup_{-r \leqslant x \leqslant r} \frac{|x|^{2n+2}}{1 + x^2} \leqslant r^{2n+2} \to 0.$$

Therefore the convergence is uniform on $[-r, r]$. Frequently this is the best one can do on an open domain.

10.3. Norm and Completeness

10.3.1. DEFINITION. Define the *uniform norm* on $C[a, b]$ by

$$\|f\|_\infty = \sup_{a \leqslant x \leqslant b} |f(x)| = \max_{a \leqslant x \leqslant b} |f(x)|.$$

Note that the supremum is a maximum by the Extreme Value Theorem, and thus this is a finite value. The properties that make it a *norm* are contained in the following proposition. The proof is left as an exercise for the reader.

10.3.2. PROPOSITION. *The uniform norm is a function from $C[a, b]$ into $[0, \infty)$ such that for $f, g \in C[a, b]$,*

(1) $\|f\|_\infty = 0$ *if and only if* $f = 0$ (positive definite);

(2) $\|tf\|_\infty = |t|\, \|f\|_\infty$ *for all* $t \in \mathbb{R}$ (positive homogeneous);

(3) $\|f + g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty$ (triangle inequality).

10.3.3. OBSERVATION. If (f_n) is a sequence of functions in $C[a, b]$, then $f_n(x)$ converges uniformly to $f(x)$ on $[a, b]$ if and only if $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$. The quantity $\|f - g\|_\infty$ is a distance function that satisfies the triangle inequality and measures uniform convergence.

10.3.4. DEFINITION. A sequence $(f_n)_{n \geqslant 1}$ in $C[a, b]$ is a *Cauchy sequence* if for all $\varepsilon > 0$, there is an N so that if $N \leqslant n < m$, then $\|f_n - f_m\|_\infty < \varepsilon$.

10.3.5. THEOREM. *$C[a, b]$ is complete. That is, every Cauchy sequence (f_n) of functions in $C[a, b]$ converges.*

PROOF. Let $(f_n)_{n \geqslant 1}$ be a Cauchy sequence in $C[a, b]$. Then for each $x \in [a, b]$, the scalar sequence $(f_n(x))$ is a Cauchy sequence of real numbers since given $\varepsilon > 0$, use the N provided and observe that

$$|f_n(x) - f_m(x)| \leqslant \|f_n - f_m\|_\infty < \varepsilon \quad \text{for} \quad N \leqslant n < m.$$

Therefore

$$f(x) := \lim_{n\to\infty} f_n(x)$$

exists as a pointwise limit. Moreover, from the estimate above,

$$\|f - f_n\|_\infty = \sup_{x\in[a,b]} |f(x) - f_n(x)| \leqslant \limsup_{m\to\infty} \|f_m - f_n\|_\infty \leqslant \varepsilon \quad \text{for} \quad N \leqslant n.$$

That means that f_n converges uniformly to f. By Theorem 10.2.5, f is continuous. Thus f_n converges to f in $C[a,b]$. Therefore $C[a,b]$ is complete. ∎

10.4. Uniform convergence and integration

Now we explore the relationship between uniform convergence and integration. Since integration is defined as a limiting procedure using Riemann sums, the interchange of limits and integrals is the interchange of two limits. This is something that always needs careful consideration.

We saw in Example 10.2.4 that if a sequence converges pointwise to f, their integrals may not converge to $\int_0^1 f(t)\,dt$.

10.4.1. Integral Convergence Theorem. *Suppose that (f_n) is a sequence of (continuous) functions in $C[a,b]$ which converges uniformly to $f(x)$. Then $F_n(x) = \int_a^x f_n(t)\,dt$ converges uniformly to $F(x) = \int_a^x f(t)\,dt$.*

Proof. We compute

$$\begin{aligned}|F_n(x) - F(x)| = \Big| \int_a^x f_n(t) - f(t)\,dt \Big| &\leqslant \int_a^x |f_n(t) - f(t)|\,dt \\ &\leqslant |x-a|\,\|f_n - f\|_\infty \leqslant (b-a)\|f_n - f\|_\infty.\end{aligned}$$

Therefore $\|F_n - F\|_\infty \leqslant (b-a)\|f_n - f\|_\infty$, which goes to 0 as $n \to \infty$. Thus F_n converges uniformly to F. ∎

10.4.2. Corollary. *Suppose that (f_n) is a sequence of continuous functions on $[a,b]$ which converges uniformly to $f(x)$. Then*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx$$

We saw in Example 10.2.3 that the derivatives of a uniformly convergent sequence need not converge. However if we control the derivatives, we can apply the Integral Convergence Theorem.

10.4.3. Corollary. *Suppose that f_n are C^1 functions on $[a, b]$ such that $f_n'(x)$ converges uniformly to a function $g(x)$. If there is a point $c \in [a, b]$ so that $\lim_{n\to\infty} f_n(c) = \gamma$ exists, then f_n converges uniformly to $G(x) = \gamma + \int_c^x g(t)\,dt$.*

Proof. By the FTC, $f_n(x) = f_n(c) + \int_c^x f_n'(t)\,dt$. By the Integral Convergence Theorem, $\int_c^x f_n'(t)\,dt$ converges uniformly to $\int_c^x g(t)\,dt$. Therefore $f_n(x)$ converges uniformly to $\gamma + \int_c^x g(t)\,dt = G(x)$. ■

One useful consequence of this is the following.

10.4.4. Corollary. *Suppose that $f_n(x)$ are C^1 functions on $[a, b]$ such that $f_n(x)$ converges uniformly to $f(x)$ and $f_n'(x)$ converges uniformly to $g(x)$. Then f is differentiable and $f' = g$.*

10.5. Series of functions

10.5.1. Definition. A series of functions $\sum_{n=1}^{\infty} f_n(x)$ in $C[a, b]$ *converges uniformly* on $[a, b]$ if the sequence of partial sums $s_n(x) = \sum_{k=1}^{n} f_k(x)$ converges uniformly.

A handy tool for verifying uniform convergence of series is the following test.

10.5.2. Weierstrass M-test. *Suppose that $f_n \in C[a, b]$ and there are constants $M_n \geqslant \|f_n\|_\infty$ such that $\sum_{n\geqslant 1} M_n < \infty$. Then $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly.*

Proof. For each $x \in [a, b]$, $\sum_{n\geqslant 1} |f_n(x)| \leqslant \sum_{n\geqslant 1} M_n < \infty$. Therefore this series converges absolutely to a function $f(x)$ pointwise. Moreover

$$\|f - s_n\|_\infty = \sup_{a\leqslant x\leqslant b} |f(x) - s_n(x)| = \max_{a\leqslant x\leqslant b} \Big| \sum_{k=n+1}^{\infty} f_k(x) \Big|$$
$$\leqslant \sum_{k=n+1}^{\infty} \|f_n\|_\infty \leqslant \sum_{k=n+1}^{\infty} M_k.$$

The right hand side converges to 0 as $n \to \infty$. Therefore $s_n(x)$ converges uniformly to $f(x)$. ■

10.5.3. EXAMPLE. Let's take another look at Example 10.1.7. Consider the series $\sum_{n=0}^{\infty} (-x^2)^n$. This is a geometric series, and it converges if and only if $|x| < 1$. In this case, the sum is

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

Now $\sup_{|x|<1} |(-x^2)^n| = 1$ for each $n \geqslant 0$. Thus this convergence is not uniform on $(-1, 1)$.

We fix some $r \in (0, 1)$. Then

$$\|(-x^2)^n\|_{C[-r,r]} = \max_{|x|\leqslant r} |(-x^2)^n| = r^{2n}.$$

Since $\sum_{n=0}^{\infty} r^{2n} = \dfrac{1}{1-r^2} < \infty$, the Weierstrass M-test applies to show that the series $\sum_{n=0}^{\infty} (-x^2)^n$ converges uniformly on $[-r, r]$ to $\dfrac{1}{1+x^2}$.

Let

$$F_n(x) = \int_0^x \sum_{k=0}^{n} (-t^2)^k \, dt = \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{2k+1}.$$

By the Integral Convergence Theorem, this sequence of functions converges uniformly on $[-r, r]$ to

$$F(x) = \int_0^x \frac{1}{1+t^2} \, dt = \tan^{-1}(x).$$

Therefore

$$\tan^{-1}(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1} \quad \text{for} \quad |x| < 1.$$

The convergence is uniform on $[-r, r]$ for each $r < 1$, so this convergence is valid pointwise on $(-1, 1)$. At $x = \pm 1$, we have the series $\pm \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$. This converges by the Alternating series test. Indeed, this is an alternating series for every value of $x \in [-1, 1]$. The error estimate from the Alternating series test shows that

$$\left| \tan^{-1}(x) - \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{2k+1} \right| \leqslant \left| \frac{(-1)^{n+1} x^{2n+3}}{2k+3} \right| \leqslant \frac{1}{2n+3}.$$

This shows that

$$\| \tan^{-1}(x) - s_n(x) \|_{C[-1,1]} \leqslant \frac{1}{2n+3}.$$

So the Taylor series for $\tan^{-1}(x)$ converges uniformly on $[-1,1]$ even though the series for its derivative does not. However it converges very slowly at $x = 1$. The famous conditionally convergent series

$$\frac{\pi}{4} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = 1 - \tfrac{1}{3} + \tfrac{1}{5} - \tfrac{1}{7} - \tfrac{1}{9} + \dots$$

is correct, but of limited use.

10.5.4. EXAMPLE. *Weierstrass's Nowhere Differentiable Function.* There exist continuous functions that are not differentiable at any point. It is not easy to write one down. One of the first examples is due to Weierstrass. Define

$$f(x) = \sum_{k\geqslant 1} 2^{-k}\cos(10^k \pi x) = \sum_{k\geqslant 1} f_k(x) \quad \text{for} \quad x \in \mathbb{R}.$$

Since $\|f_k\|_\infty = 2^{-k}$, the Weierstrass M-test shows that this series converges uniformly to a continuous function on $\mathbb{R}$. Moreover each f_k is 1-periodic, so f has period 1. Thus we need only consider $x \in [0,1]$.

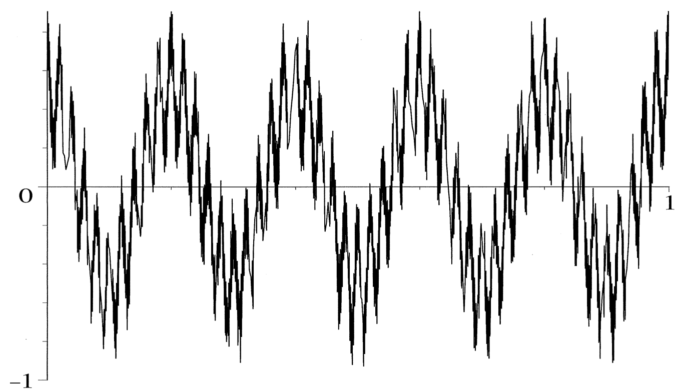


FIGURE 10.1. Weierstrass Nowhere differentiable function

Let $x = 0.x_1x_2x_3\cdots \in [0,1]$. For each $n \geqslant 1$, let $a_n = 0.x_1x_2x_3\dots x_n$ and $b_n = a_n + 10^{-n}$. Notice that $10^n a_n$ is an integer and $10^n b_n = 10^n a_n + 1$; so

$$f_n(a_n) = 2^{-n}\cos(10^n \pi a_n) = 2^{-n}(-1)^{10^n a_n}$$
$$f_n(b_n) = 2^{-n}\cos(10^n \pi b_n) = 2^{-n}(-1)^{10^n a_n + 1}.$$

Therefore $|f_n(a_n) - f_n(b_n)| = 2^{1-n}$.

If $k > n$, $10^k a_n$ and $10^k b_n$ are both even integers, so that $f_k(a_n) = f_k(b_n)$. If $1 \leqslant k < n$, the Mean Value Theorem shows that

$$|f_k(a_n) - f_k(b_n)| \leqslant \|f_k'\|_\infty\,(b_n - a_n) = (2^{-k}10^k\pi)10^{-n} = 2^{-n}5^{k-n}\pi.$$

Therefore

$$\begin{aligned}|f(a_n) - f(b_n)| &= \Big|\sum_{k=1}^{\infty} f_k(a_n) - f_k(b_n)\Big| \\ &\geqslant |f_n(a_n) - f_n(b_n)| - \sum_{k=1}^{n-1} |f_k(a_n) - f_k(b_n)| \\ &\geqslant 2^{1-n} - 2^{-n}\pi \sum_{k=1}^{n-1} 5^{k-n} \\ &> 2^{-n}(2 - \frac{\pi}{4}) > 2^{-n}.\end{aligned}$$

It follows that choosing the endpoint $y_n \in \{a_n, b_n\}$ judiciously, we can arrange that $|f(y_n) - f(x)| > 2^{-n-1}$. However $|y_n - x| \leqslant 10^{-n}$. Therefore

$$\left|\frac{f(y_n) - f(x)}{y_n - x}\right| > \frac{2^{-n-1}}{10^{-n}} = \frac{5^n}{2}.$$

This tends to ∞, from which we deduce that f is not differentiable at x.

10.6. Power series

In this section, we study a special kind of series of functions that plays a central role in Taylor series.

10.6.1. Definition. A *power series* about $x = x_0$ is a series of functions of the form $\sum\limits_{n=0}^{\infty} a_n(x - x_0)^n$.

The first important result about power series is that convergence occurs in an interval centred at x_0 that can be explicitly computed. Note that the theorem does not say what happens at the endpoints of this interval.

10.6.2. Hadamard's Theorem. *Given a power series,* $\sum\limits_{n=0}^{\infty} a_n(x - x_0)^n$, *define*

$$\alpha = \limsup_{n\to\infty} |a_n|^{1/n} \quad \textit{and} \quad R = \begin{cases} +\infty & \textit{if } \alpha = 0 \\ \frac{1}{\alpha} & \textit{if } 0 < \alpha < \infty. \\ 0 & \textit{if } \alpha = +\infty \end{cases}$$

Then

(1) *The series converges absolutely for each x such that $|x - x_0| < R$.*

(2) *The series diverges for all x such that $|x - x_0| > R$.*

(3) *If* $0 \leqslant r < R$*, the series converges uniformly on* $[x_0 - r, x_0 + r]$.

PROOF. We use the root test:

$$\limsup_{n\to\infty} \left|a_n(x-x_0)^n\right|^{1/n} = \alpha|x - x_0|.$$

Therefore this converges absolutely if $\alpha|x - x_0| < 1$, or $|x - x_0| < R$; and it diverges if $\alpha|x - x_0| > 1$, or $|x - x_0| > R$. Endpoints have to be checked separately.

If $0 \leqslant r < R$, then $\sup_{|x-x_0|\leqslant r} |a_n(x - x_0)|^n = |a_n|r^n$. By (1), we have $\sum_{n\geqslant 0} |a_n|r^n < \infty$. Thus by the Weierstrass M-test, the series converges uniformly on $[x_0 - r, x_0 + r]$. ∎

10.6.3. DEFINITION. The value R in Hadamard's Theorem is called the *radius of convergence* of the power series.

10.6.4. EXAMPLE. Consider the series $\sum_{n\geqslant 0} \frac{x^n}{2^n n^a}$ for $a > 0$. Then

$$\limsup_{n\to\infty} \left|\frac{1}{2^n n^a}\right|^{1/n} = \frac{1}{2} \limsup_{n\to\infty} \left(\frac{1}{n^{1/n}}\right)^a = \frac{1}{2}.$$

Therefore the radius of converges is $R = 2$. So the series converges on $(-2, 2)$ and diverges if $|x| > 2$.

We check the endpoints separately. For $x = 2$, we get the series $\sum_{n\geqslant 0} \frac{1}{n^a}$. This converges absolutely if $a > 1$ and diverges if $a \leqslant 1$. Now consider $x = -2$. We get the series $\sum_{n\geqslant 0} \frac{(-1)^n}{n^a}$. This converges absolutely if $a > 1$ and converges conditionally by the Alternating series test if $0 < a \leqslant 1$. So we see that the series can converge at both endpoints, or one endpoint.

10.6.5. EXAMPLE. In Example 10.5.3, we saw that the power series $\sum_{n\geqslant 0} (-x^2)^n$ converges on $(-1, 1)$ to $\frac{1}{1 + x^2}$, and diverges for $|x| \geqslant 1$. So this series has radius of convergence 1, and fails to converge at either endpoint. However the series $\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n + 1}$ also has radius of convergence 1, and converges on $[-1, 1]$ to $\tan^{-1}(x)$. In fact the convergence is uniform on $[-1, 1]$.

10.6.6. EXAMPLE. In Examples 10.1.5 and 10.2.6, we saw that the power series $\sum_{n\geqslant 0} \frac{x^n}{n!}$ converges absolutely for all $x \in \mathbb{R}$. Thus the radius of convergence is

∞. We also showed that convergence is uniform on $[-r, r]$ for any $r < \infty$, but does not converge uniformly on the whole real line.

10.6.7. EXAMPLE. Consider the power series $\sum_{n\geqslant 0} n!x^n$. By the ratio test, we have

$$\lim_{n\to\infty}\left|\frac{(n+1)!x^{n+1}}{n!x^n}\right| = \lim_{n\to\infty}(n+1)|x| = \begin{cases} +\infty & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Thus the series diverges if $x \neq 0$. So the radius of convergence is 0.

10.6.8. EXAMPLE. Consider the power series $\sum_{n\geqslant 0} \dfrac{x^{2n}}{2^n}$. Then

$$\alpha = \limsup_{n\to\infty}\left|\frac{1}{2^n}\right|^{1/2n} = \frac{1}{\sqrt{2}} \quad \text{and} \quad R = \sqrt{2}.$$

You could also use the ratio test here.

10.7. Differentiation and integration of power series

While the derivative of a series is often not the series of derivatives, things work out well for power series.

10.7.1. THEOREM. Term by term differentiation of power series.
Suppose that $f(x) = \sum_{n\geqslant 0} a_n(x-x_0)^n$ *has a radius of convergence* $R > 0$. *Then* $g(x) = \sum_{n\geqslant 0} na_n(x-x_0)^{n-1}$ *has a radius of convergence* R*; and* $f'(x) = g(x)$ *for* $|x - x_0| < R$.

PROOF. The radius of convergence for the derived series is given by the reciprocal of

$$\alpha = \limsup_{n\to\infty}|na_n|^{\frac{1}{n-1}} = \limsup_{n\to\infty} n^{\frac{1}{n-1}} \limsup_{n\to\infty}\left(|a_n|^{1/n}\right)^{\frac{n}{n-1}} = \frac{1}{R}.$$

Thus the radius of convergence is also R. We see that $s_N(x) = \sum_{n=0}^{N} a_n(x-x_0)^n$ converges uniformly to $f(x)$ and $s_N'(x) = \sum_{n=1}^{N} na_n(x-x_0)^{n-1}$ converges uniformly to $g(x)$ on $[x_0 - r, x_0 + r]$ for $0 < r < R$. By Corollary 10.4.4 to the Integral convergence theorem, f is differentiable and $f'(x) = g(x)$ on $[x_0 - r, x_0 + r]$ for $0 < r < R$. Therefore $f'(x) = g(x)$ on $(x_0 - R, x_0 + R)$. ∎

The following consequence for integration is immediate.

10.7.2. COROLLARY. Term by term integration of power series. *Suppose that* $f(x) = \sum_{n\geqslant 0} a_n(x-x_0)^n$ *has a radius of convergence* $R > 0$. *Then* $F(x) = \sum_{n\geqslant 0} \frac{a_n}{n+1}(x-x_0)^{n+1}$ *has a radius of convergence* R; *and*

$$F(x) = \int_{x_0}^{x} f(t)\,dt \quad \textit{for} \quad |x-x_0| < R.$$

We also obtain the following more powerful consequence.

10.7.3. COROLLARY. *Suppose that* $f(x) = \sum_{n\geqslant 0} a_n(x-x_0)^n$ *has a radius of convergence* $R > 0$. *Then* f *is* C^∞ *on* (x_0-R, x_0+R). *Moreover* $a_n = \dfrac{f^{(n)}(x_0)}{n!}$ *for* $n \geqslant 0$.

PROOF. This follows from repeated application of term by term differentiation, so that f has derivatives of all orders. The constant term in the series for $f^{(n)}(x)$ is $n!a_n = f^{(n)}(x_0)$. ■

10.7.4. EXAMPLE. Let $f(x) = \sum_{n\geqslant 0} \frac{x^n}{n!}$. We have already shown that $f(x) = e^x$, but we establish this here in a different way. By the ratio test,

$$\lim_{n\to\infty} \frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}} = \lim_{n\to\infty} \frac{x}{n+1} = 0.$$

Therefore this converges for all $x \in \mathbb{R}$ and so $R = \infty$. The derivative is

$$f'(x) = \sum_{n\geqslant 1} \frac{nx^{n-1}}{n!} = \sum_{n\geqslant 1} \frac{x^{n-1}}{(n-1)!} = \sum_{n\geqslant 0} \frac{x^n}{n!} = f(x).$$

Hence as long as $f(x) \neq 0$ (which by continuity includes an interval around 0 where $f(0) = 1$,

$$1 = \frac{f'(x)}{f(x)} = \frac{d}{dx}\big(\ln f(x)\big).$$

Integrating, we obtain

$$x = \int_0^x 1\,dt = \int_0^x \frac{d}{dx}\big(\ln f(t)\big)\,dt = \ln f(t)\Big|_0^x = \ln f(x).$$

Exponentiating, we have $f(x) = e^x$ on an interval around 0.

Now let $g(x) = e^{-x}f(x)$ for $x \in \mathbb{R}$. Then

$$g'(x) = -e^{-x}f(x) + e^{x}f'(x) = e^{-x}(f'(x) - f(x)) = 0.$$

Therefore $g(x)$ is constant and $g(0) = 1$, so that $f(x) = e^x$ everywhere.

10.7.5. EXAMPLE. Let $f(x) = \sum\limits_{n=1}^{\infty} n^2x^n$. Since $\limsup_{n\to\infty}(n^2)^{1/n} = 1$, we have $R = 1$. Evidently the series diverges at $x = \pm 1$, but it converges on $(-1, 1)$. Now $\dfrac{f(x)}{x} = \sum\limits_{n=0}^{\infty}(n+1)^2x^n$. Integrate:

$$g(x) = \int_0^x \frac{f(t)}{t}\,dt = \sum_{n=0}^{\infty}(n+1)x^{n+1}$$

is also valid on $(-1, 1)$. Similarly $\dfrac{g(x)}{x} = \sum\limits_{n=0}^{\infty}(n+1)x^n$. Integrating again, we can sum a geometric series

$$h(x) = \int_0^x \frac{g(t)}{t}\,dt = \sum_{n=0}^{\infty} x^{n+1} = \frac{x}{1-x}.$$

Therefore

$$\frac{g(x)}{x} = h'(x) = \frac{(1-x)+x}{(1-x)^2} = \frac{1}{(1-x)^2}.$$

Therefore $g(x) = \dfrac{x}{(1-x)^2}$ and

$$\frac{f(x)}{x} = g'(x) = \frac{(1-x)^2 + 2x}{(1-x)^3} = \frac{1+x}{(1-x)^3}.$$

So finally we have $f(x) = \dfrac{x(1+x)}{(1-x)^3}$ on $(-1, 1)$.

We can plug in values of x to get interesting sums. Take $x = \pm\frac{1}{2}$:

$$\sum_{n=1}^{\infty}\frac{n^2}{2^n} = f(\tfrac{1}{2}) = \frac{\frac{1}{2}\frac{3}{2}}{\frac{1}{8}} = 6 \quad\text{and}\quad \sum_{n=1}^{\infty}\frac{(-1)^2n^2}{2^n} = f(-\tfrac{1}{2}) = \frac{-\frac{1}{4}}{\frac{27}{8}} = \frac{-2}{27}.$$

We need the following observation about the uniqueness of a power series for a function.

10.7.6. PROPOSITION. *If two convergent power series* $\sum\limits_{n\geqslant 0} a_n(x-x_0)^n$ *and* $\sum\limits_{n\geqslant 0} b_n(x-x_0)^n$ *agree on* $(x_0 - r, x_0 + r)$ *for some* $r > 0$*, then* $b_n = a_n$ *for* $n \geqslant 0$*. That is, the power series for a function is unique (if it has one).*

PROOF. Let $f(x) = \sum_{n \geqslant 0} a_n(x - x_0)^n$ and $g(x) = \sum_{n \geqslant 0} b_n(x - x_0)^n$. If they agree on an interval around x_0, then all of their derivatives agree on the interval as well. Hence by Corollary 10.7.3,

$$a_n = \frac{f^{(n)}(x_0)}{n!} = \frac{g^{(n)}(x_0)}{n!} = b_n \quad \text{for} \quad n \geqslant 0.$$

Thus the power series for $f(x)$ is unique. ■

10.7.7. EXAMPLE. Let $\alpha \neq 0$ be any real number. Look for a power series for $f(x) = (1 + x)^\alpha$ near $x = 0$. Then $f'(x) = \alpha(1 + x)^{\alpha - 1}$; and therefore $\alpha f(x) = (1 + x) f'(x)$. Suppose that there is a power series $f(x) = \sum_{n \geqslant 0} a_n x^n$ valid on $(-r, r)$ for some $r > 0$. Then $f'(x) = \sum_{n \geqslant 1} n a_n x^{n-1}$. Therefore

$$\alpha \sum_{n \geqslant 0} a_n x^n = (1 + x) \sum_{n \geqslant 1} n a_n x^{n-1} = a_1 + \sum_{n \geqslant 1} (n a_n + (n + 1) a_{n+1}) x^n.$$

By Proposition 10.7.6, the coefficients are equal:

$$\alpha a_0 = a_1 \quad \text{and} \quad \alpha a_n = n a_n + (n + 1) a_{n+1} \quad \text{for } n \geqslant 1.$$

Now $a_0 = f(0) = 1$. Thus

$$a_1 = \alpha \quad \text{and} \quad a_{n+1} = \frac{\alpha - n}{n + 1} a_n \quad \text{for } n \geqslant 1.$$

The next few terms are

$$a_2 = \frac{\alpha(\alpha - 1)}{2}, \quad a_3 = \frac{\alpha(\alpha - 1)(\alpha - 2)}{6}, \quad a_4 = \frac{\alpha(\alpha - 1)(\alpha - 2)(\alpha - 3)}{24}.$$

The pattern, which is readily verified by induction, is

$$a_n = \frac{\alpha(\alpha - 1) \cdots (\alpha + 1 - n)}{n!} =: \binom{\alpha}{n}.$$

We call this a binomial coefficient by analogy with the positive integer case.

Let's verify that this is correct. Define $g(x) = \sum_{n \geqslant 0} \binom{\alpha}{n} x^n$. By the ratio test, we have

$$\lim_{n \to \infty} \left| \frac{\binom{\alpha}{n+1} x^{n+1}}{\binom{\alpha}{n} x^n} \right| = \lim_{n \to \infty} \frac{|(\alpha - n)x|}{n + 1} = |x|.$$

Therefore the series converges for $|x| < 1$ and diverges for $|x| > 1$. Term by term differentiation shows that

$$\begin{aligned} g'(x) &= \sum_{n\geqslant 1} \binom{\alpha}{n} n x^{n-1} = \sum_{n\geqslant 0} (n+1) \binom{\alpha}{n+1} x^n \\ &= \sum_{n\geqslant 0} (n+1) \frac{\alpha(\alpha-1)\cdots(\alpha-n)}{(n+1)!} x^n \\ &= \sum_{n\geqslant 0} (\alpha-n) \frac{\alpha(\alpha-1)\cdots(\alpha+1-n)}{n!} x^n \\ &= \alpha \sum_{n\geqslant 0} \binom{\alpha}{n} x^n - x \sum_{n\geqslant 1} \binom{\alpha}{n} n x^{n-1} = \alpha g(x) - x g'(x). \end{aligned}$$

Hence $(1+x)g'(x) = \alpha g(x)$.

Now we have $g(0) = 1$ and $\dfrac{g'(x)}{g(x)} = \dfrac{\alpha}{1+x}$. Integrating with $|x| < 1$,

$$\ln g(x) = \int_0^x \frac{g'(t)}{g(t)}\,dt = \int_0^x \frac{\alpha}{1+t}\,dt = \alpha \ln|1+t|\Big|_0^x = \alpha \ln(1+x).$$

Therefore $g(x) = (1+x)^\alpha$. That is,

$$(1+x)^\alpha = \sum_{n\geqslant 0} \binom{\alpha}{n} x^n \quad \text{for} \quad |x| < 1 \text{ and } \alpha \neq 0.$$

10.8. Abel's Theorem

In this section, we show that if the power series converges at an endpoint, then it takes the right value there.

10.8.1. ABEL'S THEOREM. *Suppose that* $f(x) = \sum_{n\geqslant 0} a_n(x-x_0)^n$ *has radius of convergence* $0 < R < \infty$. *If* $\sum_{n\geqslant 0} a_n R^n$ *converges, then the series converges uniformly on* $[x_0, x_0+R]$ *and*

$$\sum_{n\geqslant 0} a_n R^n = \lim_{x\to(x_0+R)^-} f(x).$$

Similarly if $\sum_{n\geqslant 0} a_n(-R)^n$ *converges, then the series converges uniformly on the interval* $[x_0-R, x_0]$ *and*

$$\sum_{n\geqslant 0} a_n(-R)^n = \lim_{x\to(x_0-R)^+} f(x).$$

PROOF. Let $b_n = a_n R^n$. Then $\sum_{n\geqslant 0} b_n$ converges. Given $\varepsilon > 0$, there is an N so that if $N \leqslant n < m$, then $\Big|\sum_{i=n+1}^{m} b_i\Big| < \varepsilon$. Fix $x_1 \in [x_0, x_0 + R)$. Then $c_n = \left(\frac{x_1 - x_0}{R}\right)^n$ converges monotonely to 0. Using the Summation by parts lemma, we have

$$\begin{aligned}\Big|\sum_{i=n+1}^{m} a_i(x_1 - x_0)^n\Big| = \Big|\sum_{i=n+1}^{m} b_i c_i\Big| &= \Big|c_m \sum_{i=n+1}^{m} b_i + \sum_{i=n+1}^{m} b_i \sum_{j=i}^{m-1} c_j - c_{j+1}\Big| \\ &\leqslant c_m \Big|\sum_{i=n+1}^{m} b_i\Big| + \sum_{j=n}^{m-1} c_j - c_{j+1} \Big|\sum_{i=n+1}^{j} b_i\Big| \\ &< \varepsilon(c_m + \sum_{j=n}^{m-1} c_j - c_{j+1}) = c_n\varepsilon \leqslant \varepsilon.\end{aligned}$$

Therefore

$$\sup_{x_0 \leqslant x \leqslant x_0 + R} \Big|\sum_{i=n+1}^{m} a_i(x_1 - x_0)^n\Big| \leqslant \varepsilon \quad \text{for all} \quad N \leqslant n < m.$$

That is, the series is uniformly Cauchy, so converges uniformly. Therefore the limit function $f(x)$ is continuous on $[x_0, x_0 + R]$. Thus

$$\sum_{n\geqslant 0} a_n R^n = f(x_0 + R) = \lim_{x \to (x_0+R)^-} f(x).$$

The other case follows by symmetry. ∎

10.8.2. EXAMPLE. Consider the alternating harmonic series $\sum_{n=1}^{\infty} \dfrac{(-1)^{n+1}}{n}$. This is a limiting case of the power series $f(x) = \sum_{n=1}^{\infty} \dfrac{x^n}{n}$. This series has radius of convergence 1 because $\lim_{n\to\infty} \dfrac{1}{n^{1/n}} = 1$. The series fails to converge at $x = 1$, but does converge at $x = -1$ by the alternating series test. So Abel's Theorem applies to say that the limit function is continuous, and is the uniform limit of the partial sums on $[-1, 0]$.

Using term by term differentiation, we have

$$f'(x) = \sum_{n\geqslant 1} x^{n-1} = \sum_{n\geqslant 0} x^n = \frac{1}{1-x}.$$

Since $f(0) = 0$, we have

$$f(x) = \int_0^x \frac{1}{1-t}\,dt = -\ln(1-x).$$

By Abel's Theorem,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = -f(-1) = \ln 2.$$

Exercises for Chapter 10

1. Use Taylor polynomials to compute these limits, not L'Hôpital's Rule.

(a) $\lim_{x\to 0} \dfrac{e^x + e^{-x} - 2}{x^2}$.

(b) $\lim_{x\to 0} \dfrac{\sinh x - \sin x}{x^3}$.

(c) $\lim_{x\to 0} \left(\dfrac{\sin x}{x}\right)^{1/x^2}$. HINT: Use a 3rd order polynomial for $\sin(x)$ at $x = 0$ and a 2nd order polynomial for $\ln(x)$ at $x = 1$.

2. (a) Verify that $4\tan^{-1}(\frac{1}{5}) - \tan^{-1}(\frac{1}{239}) = \dfrac{\pi}{4}$.

(b) Use (a) and Taylor polynomials to calculate π to 6 decimals of accuracy with error estimates.

3. (a) Find the Taylor polynomials $P_{n,1}(x)$ for $f(x) = \ln x$ about $a = 1$, and give the error estimates.

(b) Compare the errors for the following methods for computing $\ln 2$. Which is best?

(i) $P_{n,1}(2)$ (ii) $-P_{n,1}(.5)$ (iii) $P_{n,1}(\frac{4}{3}) - P_{n,1}(\frac{2}{3})$.

4. Let $f(x) = (1 + x)^{-1/2}$.

(a) Find a formula for $f^{(k)}(x)$. Hence show that

$$\frac{f^{(k)}(0)}{k!} = \binom{\frac{-1}{2}}{k} := \frac{\frac{-1}{2}(\frac{-1}{2} - 1)\cdots(\frac{-1}{2} + 1 - k)}{k!} = \frac{(-1)^k}{4^k}\binom{2k}{k}$$

(b) Find the Taylor polynomial $P_{n,0}(x)$ for f, and give the error estimate.

(c) Show that $\sqrt{2} = 1.4 f(-.02)$. Use this to compute $\sqrt{2}$ to 6 decimal places with error estimates.

5. Let $f_n(x) = xne^{-nx}$ for $x \geqslant 0$ and $n \geqslant 1$. Find $\lim_{n\to\infty} f_n(x)$. Is this limit uniform on $[0, \infty)$?

6. Let $f_n(x) = \dfrac{x}{1 + nx^2}$ for $x \in \mathbb{R}$ and $n \geqslant 1$. Find $\lim_{n\to\infty} f_n(x)$. Is this limit uniform on $\mathbb{R}$?

7. Let f_n and g_n be continuous functions on $[a, b]$ for $n \geqslant 1$. Suppose that f_n converges uniformly to f and g_n converges uniformly to g on $[a, b]$. Prove that $f_n g_n$ converges uniformly to fg on $[a, b]$.

8. Let $f_n(x) = \dfrac{x^2}{(1 + x^2)^n}$ for $x \in \mathbb{R}$, and let $s_k(x) = \sum_{n=0}^{k} f_n(x)$.
(a) Find $\lim_{k\to\infty} s_k(x)$.
(b) For which values $a < b$ does this series converge uniformly on $[a, b]$?

9. For $n \geqslant 1$, define functions f_n on $[0, \infty)$ by

$$f_n(x) = \begin{cases} e^{-x} & \text{for} \quad 0 \leqslant x \leqslant n \\ e^{-2n}(e^n + n - x) & \text{for} \quad n \leqslant x \leqslant n + e^n \\ 0 & \text{for} \quad x \geqslant n + e^n. \end{cases}$$

(a) Find the pointwise limit f of f_n. Show that the convergence is uniform on $[0, \infty)$.
(b) Compute $\int_0^\infty f(x)\,dx$ and $\lim_{n\to\infty} \int_0^\infty f_n(x)\,dx$.
(c) Why does this not contradict Integral Convergence Theorem?

10. (a) Suppose that $f : \mathbb{R} \to \mathbb{R}$ is uniformly continuous. Let $f_n(x) = f(x + 1/n)$. Prove that f_n converges uniformly to f on $\mathbb{R}$.
(b) Does this remain true if f is just continuous? Prove it or provide a counterexample.

11. For which values of $x \geqslant 1$ does the expression $x^{x^{x^{x^{\cdot^{\cdot^{\cdot}}}}}}$ make sense?
To tackle this, define $f_1(x) = x$ and $f_{n+1}(x) = x^{f_n(x)}$ for $x \geqslant 1$ and $n \geqslant 1$.
(a) Show that $f_{n+1}(x) \geqslant f_n(x)$ for all $n \geqslant 1$.
(b) When $L(x) = \lim_{n\to\infty} f_n(x)$ exists, find an equation for $L(x)$.
Use it to find an upper bound for x.
(c) For these values of x, show by induction that $f_n(x)$ is bounded above by e for all $n \geqslant 1$. What can you conclude?
(d) What happens for larger x?

12. Find the radius of convergence for the following series, and evaluate the function.
(a) $f(x) = \sum_{n=0}^{\infty} (n^2 + n)x^n$
(b) $g(x) = \sum_{n=0}^{\infty} \dfrac{x^{2n+1}}{(2n + 1)!}$
(c) $h(x) = \sum_{n=2}^{\infty} \dfrac{(-1)^n x^n}{n^2 - n}$
(d) Evaluate $\sum_{n=2}^{\infty} \dfrac{(-1)^n}{n^2 - n}$. Justify!

13. Suppose that $f(x) = \sum_{j=0}^{\infty} a_j x^j$ and $g(x) = \sum_{k=0}^{\infty} b_k x^k$ have positive radii of convergence R_1 and R_2 respectively. Let $c_n = \sum_{j=0}^{n} a_j b_{n-j}$ for $n \geqslant 0$; and let $R = \min\{R_1, R_2\}$.

(a) Define $h(x) = \sum_{n=0}^{\infty} c_n x^n$. Prove that $h(x) = f(x)g(x)$ on $(-R, R)$.

(b) Give an example where h has radius of convergence strictly greater than R.

CHAPTER 11

Differential Equations

11.1. Examples of DEs

11.1.1. DEFINITION. A *first order DE* is a relation between the independent variable x, a function $y(x)$ and its derivative $y'(x)$ of the form $f(x, y, y') = 0$. It is in *standard form* if $y' = g(x, y)$.

An *nth order DE* has the form $f(x, y, y', \dots, y^{(n)}) = 0$. It is in *standard form* if $y^{(n)} = g(x, y, y', \dots, y^{(n-1)})$.

11.1.2. EXAMPLE. Consider $x + yy' = 0$. In standard form, we have $y' = -\dfrac{x}{y}$ provided that $y'(x) \neq 0$. By inspection, we can notice that

$$\frac{d}{dx}(x^2 + y^2) = 2(x + yy') = 0.$$

Therefore $x^2 + y^2 = c$ is constant. Clearly we need $c \geqslant 0$. So the solution curves appear to be circles of radius $\sqrt{c}$ and centre $(0, 0)$. However a circle does not yield a function $y(x)$, but rather there are two values of y for most x. To get a function, we have two solutions for each $c > 0$, namely

$$y(x) = \sqrt{c - x^2} \quad \text{and} \quad y(x) = -\sqrt{c - x^2} \quad \text{for} \quad |x| < \sqrt{c}.$$

Note that $y'(\pm\sqrt{c})$ is not defined, so that the endpoints are not part of the solution. This provides two one parameter families of solutions.

The usual way to decide which of these various solutions is applicable is to provide extra data, known as *initial value conditions*. A DE of order n requires n pieces of data, often the values at some point a of $y(a), \dots y^{n-1)}(a)$, but other choices arise.

Suppose that in this case, we are told that $y(0) = r \in \mathbb{R}$. Then we can determine the solution as

$$y(x) = \sqrt{r^2 - x^2} \text{ if } r > 0 \quad \text{and} \quad y(x) = -\sqrt{r^2 - x^2} \text{ if } r < 0.$$

11.1.3. EXAMPLE. Consider $y' = f(x)$ and $y(a) = \gamma$. By the FTC, this has a unique solution

$$y(x) = \gamma + \int_a^x f(t)\, dt.$$

11.1.4. EXAMPLE. *Radioactive decay.* A radioactive element will decay (i.e., lose an electron to become a more stable isotope) at a rate proportional to the amount of radioactive material. Suppose that $\rho(t)$ is the amount of material at time t for $t \geqslant t_0$ and that $\rho(t_0) = \rho_0$ is known. We interpret the first sentence as saying that there is a constant $k > 0$ (unknown so far) so that

$$\rho'(t) = -k\rho(t).$$

In this case, we can separate variables, getting all of the ρ's on one side and x's on the other:

$$\frac{\rho'}{\rho} = -k.$$

Therefore

$$k(t_0 - t) = \int_{t_0}^{t} -k\,ds = \int_{t_0}^{t} \frac{\rho'(s)}{\rho}(s)\,ds = \ln \rho(s)\Big|_{t_0}^{t} = \ln \frac{\rho(t)}{\rho(t_0)}.$$

Hence

$$\rho(t) = \rho(t_0)e^{-k(t-t_0)}.$$

This is called exponential decay.

The half life of uranium 235 is 4.5 billion years, and uranium 238 has a half life of 700 million years. Radio carbon or carbon 14, has a half life of 5730 ± 40 years. It is created all of the time in the atmosphere and incorporated into plant material until the plant dies. Animals eat plants and take in carbon 14 as well. The percentage of material at the time of death is predictable, so one can date the age of ancient plants and animals based on measurement of the current percentage of carbon 14. For carbon 14, the constant k can be determined from

$$\frac{1}{2} = e^{-5730k} \quad \text{or} \quad k = \frac{\ln 2}{5730} \approx 1.21 \cdot 10^{-4}$$

if time is measured in years.

11.1.5. EXAMPLE. Consider a DE $y' = f(x, y)$ where f is homogeneous of order 0, meaning that $f(tx, ty) = f(x, y)$ for all $t \neq 0$. To solve this, we make a substitution $z = \dfrac{y}{x}$ or $y = xz$. Then

$$y' = z + xz' = f(x, y) = f(1, \tfrac{y}{x}) = f(1, z).$$

Again we can separate variable

$$\frac{z'}{f(1, z) - z} = \frac{1}{x}.$$

You can now integrate and solve for z.

For a specific example of this type, consider

$$(x + y) + (x - y)y' = 0 \quad \text{where} \quad f(x, y) = \frac{x + y}{y - x} = \frac{1 + z}{z - 1}.$$

Thus

$$\frac{1}{x} = \frac{z'}{\frac{1+z}{z-1} - z} = \frac{(z-1)z'}{1+2z-z^2}.$$

Integrating, we obtain

$$\ln|x| + c = -\frac{1}{2}\ln|z^2 - 2z - 1|$$

Exponentiating, we obtain

$$x^2(z^2 - 2z - 1) = C.$$

Replacing z by $\dfrac{y}{x}$ again, we get

$$C = y^2 - 2xy - x^2 = (y-x)^2 - 2x^2.$$

The parameter C yields a family of hyperbolae which are asymptotic to the lines $y = (1 \pm \sqrt{2})x$. The blue lines are for positive values of C and the red for negative values.

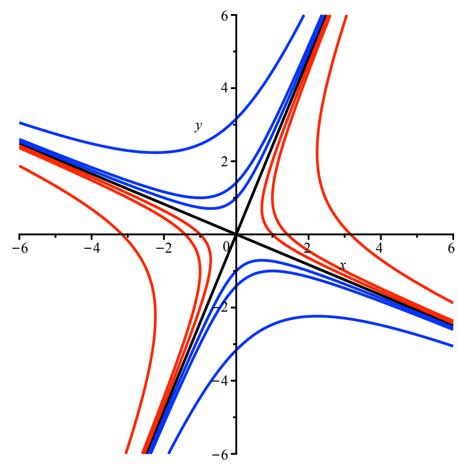


FIGURE 11.1. Solution curves

11.1.6. EXAMPLE. *Pursuit curves.* A rabbit R starts at $(a, 0)$ and runs up the line $x = a$ with speed ρ m/s. A dog D starts at $(0, 0)$ and runs straight at the rabbit at speed $k\rho$, where $k \geqslant 1$. Find the path of the dog. How long does it take him to catch the rabbit? If $k = 1$, how close does the dog get?

Let the dog's path be $\gamma(t) = (x(t), y(t))$. So $\gamma(0) = (0, 0)$. The rabbit's position at time t is $(a, \rho t)$. Thus at time t the slope of the dog's trajectory is

$$\frac{dy}{dx} = \frac{\rho t - y}{a - x}.$$

Using the arc length formula, we see that at time t, the dog has travelled a distance of

$$k\rho t = \int_0^{x(t)} \sqrt{1+\left(\tfrac{dy}{dx}\right)^2}\,dx$$

Solving the first equation for t and writing p for $\frac{dy}{dx}$, we get $\rho t = (a-x)p + y$. So

$$\frac{1}{k}\int_0^x \sqrt{1+p^2}\,dx = \rho t = (a-x)p + y.$$

Differentiating with respect to x yields

$$\frac{1}{k}\sqrt{1+p^2} = (a-x)p' - p + p = (a-x)p'.$$

At time $t = 0$, we have $y(0) = 0$ and $p(0) = y'(0) = 0$. The latter is because the initial direction of the dog is along the x-axis.

Rewrite the DE as

$$\frac{1}{k(a-x)} = \frac{p'}{\sqrt{1+p^2}} \quad \text{and} \quad p(0) = 0.$$

Integrating, we find

$$-\frac{1}{k}\ln(a-x) + c = \sinh^{-1}(p).$$

Plugging in the initial datum $p(0) = 0 = x(0)$, we get

$$0 = \sinh^{-1}(0) = -\frac{1}{k}\ln a + c \quad \text{or} \quad c = \frac{\ln a}{k}.$$

Thus

$$\begin{aligned} p(x) &= \sinh\left(\frac{\ln \frac{a}{a-x}}{k}\right) \\ &= \frac{1}{2}\left(\left(\frac{a}{a-x}\right)^{1/k} - \left(\frac{a}{a-x}\right)^{-1/k}\right) \\ &= \frac{1}{2}\left(\left(1-\tfrac{x}{a}\right)^{-1/k} - \left(1-\tfrac{x}{a}\right)^{1/k}\right). \end{aligned}$$

Finally we integrate to get y. First we deal with $k > 1$.

$$\begin{aligned} y(x) &= y(0) + \int_0^x p(x)\,dx \\ &= \frac{1}{2}\int_0^x \left(1-\tfrac{u}{a}\right)^{-1/k} - \left(1-\tfrac{u}{a}\right)^{1/k}\,du \\ &= \frac{-ka}{2(k-1)}\left(1-\tfrac{u}{a}\right)^{1-\frac{1}{k}} + \frac{ka}{2(k+1)}\left(1-\tfrac{u}{a}\right)^{1+\frac{1}{k}}\Big|_0^x \\ &= \frac{-ka}{2(k-1)}\left(1-\tfrac{x}{a}\right)^{1-\frac{1}{k}} + \frac{ka}{2(k+1)}\left(1-\tfrac{x}{a}\right)^{1+\frac{1}{k}} + \frac{ka}{k^2-1}. \end{aligned}$$

While this is complicated, we are interested in when the dog catches the rabbit, which occurs when $x = a$. At this point, $y(a) = \frac{ka}{k^2-1}$. The rabbit covers this distance in $\frac{ka}{\rho(k^2-1)}$ seconds. So the dog runs $\frac{k^2 a}{k^2-1}$ metres.

Now consider the case $k = 1$. Then

$$\begin{aligned} y(x) &= y(0) + \int_0^x p(x)\,dx = \frac{1}{2}\int_0^x \left(1 - \tfrac{u}{a}\right)^{-1} - \left(1 - \tfrac{u}{a}\right) du \\ &= -\frac{a}{2}\ln\left(1 - \tfrac{u}{a}\right) - \frac{u}{2} + \frac{u^2}{4a}\Big|_0^x \\ &= \frac{x^2}{4a} - \frac{x}{2} - \frac{a}{2}\ln\left(1 - \tfrac{x}{a}\right). \end{aligned}$$

Thus

$$\lim_{x \to a^-} y(x) = +\infty.$$

Finally we compute

$$\begin{aligned} \mathrm{dist}(D,R)^2 &= (a-x)^2 + (\rho t - y)^2 = (a-x)^2 + (a-x)^2 p^2 \\ &= (a-x)^2\Big(1 + \frac{1}{4}\Big(\left(1 - \tfrac{x}{a}\right)^{-1} - \left(1 - \tfrac{x}{a}\right)\Big)^2\Big) \\ &= (a-x)^2\Big(\frac{1}{4}\Big(\left(1 - \tfrac{x}{a}\right)^{-1} + \left(1 - \tfrac{x}{a}\right)\Big)^2\Big). \end{aligned}$$

Therefore

$$\mathrm{dist}(D,R) = \frac{1}{2}(a-x)\Big(\left(1 - \tfrac{x}{a}\right)^{-1} + \left(1 - \tfrac{x}{a}\right)\Big) = \frac{a}{2} + \frac{1}{2a}(a-x)^2.$$

Therefore, in the limit, the dog approaches within $\frac{a}{2}$ metres of the rabbit.

11.2. First Order Linear DEs

In the remaining sections, we consider a special class of DEs called linear DEs.

11.2.1. Definition. A *linear DE of order* n has the form

$$\begin{aligned} y^{(n)} &= a_0(x)y + p_1(x)y' + \dots p_{n-1}(x)y^{(n-1)} + q(x) \\ &= \sum_{i=0}^{n-1} p_i(x)y^{(i)} + q(x) \end{aligned}$$

where $p_0, \dots, p_{n-1}$ and q are in $C[a,b]$. The DE is called linear because the equation is linear in y and its derivatives, although it is not linear in x. The function $q(x)$ is called the *forcing term*. If $q = 0$, the DE is *homogeneous*. The initial data requires n pieces of information. Normally it is given as

$$\begin{aligned} y(a) &= \gamma_0 \\ y'(a) &= \gamma_1 \\ &\vdots \\ y^{(n-1)}(a) &= \gamma_{n-1}. \end{aligned}$$

11.2.2 First order linear DEs. In this section, we study DEs of the form

$$y' = p(x)y + q(x) \quad \text{and} \quad y(a) = \gamma.$$

First solve the homogeneous DE with no forcing term: $y' = p(x)y$. Thus

$$\begin{aligned} \frac{y'}{y} &= p(x) \\ \ln y &= \int p(x)\,dx + c =: P(x) + c. \\ y &= Ce^{P(x)}. \end{aligned}$$

Here C is an arbitrary constant. It can be a negative number, even though it comes from the previous line, so that it appears $C = e^c$ should be positive. Evaluating this at $x = a$ yields $\gamma = Ce^{P(a)}$, so $C = e^{-P(a)}\gamma$.

This exhibits one solution. Why is it unique? If $y(x)$ is a solution, compute

$$\big(e^{-P(x)}y\big)' = e^{-P(x)}y' - e^{-P(x)}p(x)y = e^{-P(x)}\big(y' - p(x)y\big) = 0.$$

Therefore $e^{-P(x)}y$ is constant, and so $y = Ce^{P(x)}$.

Now let y be a solution to the original DE with forcing term. As above, compute

$$\big(e^{-P(x)}y\big)' = e^{-P(x)}(y' - p(x)y) = e^{-P(x)}q(x).$$

Thus

$$e^{-P(x)}y = \int e^{-P(x)}q(x)\,dx + c.$$

Therefore

$$y = e^{P(x)}\Big(\int e^{-P(x)}q(x)\,dx + c\Big) = e^{P(x)}\int e^{-P(x)}q(x)\,dx + ce^{P(x)}.$$

In other words, the general solution to the original DE is the sum of a *particular solution* $y_p(x) = e^{P(x)}\int e^{-P(x)}q(x)\,dx$ and an arbitrary solution $ce^{P(x)}$ of the homogeneous solution. If there is also an initial condition $y(a) = \gamma$, then we can determine the constant c as

$$c = e^{-P(a)}(\gamma - y_p(a)).$$

Suppose that $y_1(x)$ and $y_2(x)$ are solutions to this DE with initial data. Define $y(x) = y_1(x) - y_2(x)$. Then

$$y' = y_1' - y_2' = p(x)y_1 + q(x) - p(x)y_2 - q(x) = p(x)y$$

and

$$y(a) = y_1(a) - y_2(a) = \gamma - \gamma = 0.$$

From the analysis of the homogeneous case, we see that $y = 0$ is the unique solution. Hence $y_1 = y_2$. This shows that the inhomogeneous DE with initial data also has a unique solution.

11.2.3. EXAMPLE. Consider $xy' - 3y = x^6$ and $y(1) = 1$. Rewrite this in standard form as $y' = \frac{3}{x}y + x^5$. Then

$$P(x) = \int \frac{3}{x}\,dx = 3\ln x + c \quad \text{or} \quad e^{P(x)} = Cx^3.$$

The forcing term yields a particular solution

$$y_p(x) = e^{P(x)} \int e^{-P(x)} q(x)\,dx = x^3 \int x^{-3}x^5\,dx = x^3\bigl(\tfrac{1}{3}x^3 + c\bigr) = \tfrac{1}{3}x^6 + cx^3.$$

Now $1 = y(1) = \frac{1}{3} + c$. Therefore the solution is $y(x) = \frac{1}{3}(x^6 + 2x^3)$.

11.2.4. EXAMPLE. *Falling bodies*. Near the surface of the earth, gravitation exerts a force $F = -mg$ in a vertical direction on a particle of mass m. Newton's Law says that $F = ma$, where a is the acceleration of the particle. Thus if $y(t)$ represents the vertical position of the particle, then $v = y'(t)$ is the velocity and the acceleration is $y'' = -g$. Suppose the body is dropped from a height H with initial velocity 0. Then the initial data is $y(0) = H$ and $y'(0) = 0$. Integrate twice:

$$y'(t) = \int_0^t y''(s)\,ds = \int_0^t -g\,ds = -gt$$
$$y(t) - H = \int_0^t y'(s)\,ds = -\int_0^t gs\,ds = -\frac{g}{2}t^2.$$

That is, $y(t) = H - \frac{g}{2}t^2$. Notice that we can eliminate t to find a direct relationship the distance fallen, $d = H - y(t) = \frac{g}{2}t^2$, and velocity, $v = y'(t) = -gt$, to get the formula $v = \sqrt{2gd}$.

However if the body is falling through air, the air adds resistance proportional to the velocity:

$$y'' = -g - cy'.$$

Substitute $v = y'$ to get a first order linear DE in v: $v' = -g - cv$. Therefore

$$\frac{-v'}{g + cv} = 1$$
$$-\frac{1}{c}\ln|g + cv| = \int \frac{-v'(t)}{g + cv(t)}\,dt = \int dt = t + c.$$
$$g + cv = c_2 e^{-ct}.$$

Thus $v = \frac{g}{c}(c_2 e^{-ct} - 1)$. Since $v(0) = 0$, we get $v = \frac{g}{c}(e^{-ct} - 1)$. The quantity

$$\lim_{t\to\infty} v(t) = \lim_{t\to\infty} \frac{g}{c}(e^{-ct} - 1) = -\frac{g}{c}$$

is called the *terminal velocity*.

A skydiver, when spread out wide, reaches a terminal velocity of about 55 m/s. Once the parachute is opened, the drag reduces the velocity to about 5 or 6 m/s.

11.3. Second Order Linear DEs

In this section, we consider DEs of the form

$$y'' = p_0(x)y + p_1(x)y' + q(x) \quad \text{and} \quad y(a) = \gamma_0,\ \ y'(a) = \gamma_1.$$

The following lemma and its corollaries explain why these equations are called linear.

11.3.1. Lemma. *Suppose that y_1 and y_2 are solutions of*

$$y'' = p_0(x)y + p_1(x)y' + q_i(x) \quad \textit{for} \quad j = 1, 2.$$

Then if $c_i \in \mathbb{R}$, $y = c_1y_1 + c_2y_2$ is a solution of

$$y'' = p_0(x)y + p_1(x)y' + c_1q_1(x) + c_2q_2(x).$$

Proof. This is a straightforward calculation.

$$\begin{aligned} y'' &= c_1y_1'' + c_2y_2'' \\ &= c_1\Big(p_0(x)y + p_1(x)y' + q_1(x)\Big) + c_2\Big(p_0(x)y + p_1(x)y' + q_2(x)\Big) \\ &= p_0(x)(c_1y_1 + c_2y_2) + p_1(x)(c_1y_1 + c_2y_2)' + \big(c_1q_1(x) + c_2q_2(x)\big). \end{aligned}$$ ■

11.3.2. Corollary. *Suppose that y_1 and y_2 are solutions of the homogeneous DE $y'' = p_0(x)y + p_1(x)y'$ for $a \leqslant x \leqslant b$. Then $c_1y_1 + c_2y_2$ is also a solution for $c_i \in \mathbb{R}$. Thus the set of solutions is a subspace of $C[a, b]$.*

PROOF. A subspace of the vector space $C[a,b]$ is a subset containing $y = 0$ with the property that it is closed under taking linear combinations. Now $y = 0$ is always a solution of any homogeneous linear DE. Lemma 11.3.1 shows that if y_1 and y_2 are solutions, then so is $c_1y_1 + c_2y_2$. A simple induction argument shows that you can take linear combinations of more solutions. ■

11.3.3. REMARK. Second order linear homogeneous DEs always have a 2-dimensional space of solutions. When combined with the initial conditions $y(a) = \gamma_0$ and $y'(a) = \gamma_1$, there is a unique solution. We do not establish existence of solutions here, but we will prove uniqueness in section 11.5. Thus if we manage to exhibit two linearly independent solutions, then we know that they span the entire solution space.

11.3.4. COROLLARY. *Suppose that y_p is a solution of*

$$y'' = p_0(x)y + p_1(x)y' + q(x).$$

Then every solution has the form $y = y_p + y_h$ where y_h is a solution of the homogeneous DE $y'' = p_0(x)y + p_1(x)y'$.

PROOF. If y is another solution of the DE, then Lemma 11.3.1 shows that $y - y_p$ is a solution of the homogeneous equation. It also shows that if y_h is a solution of the homogeneous equation, then $y = y_p + y_h$ is a solution of our DE. ■

11.3.5 Reduction of order. There is no formula for solving a second order linear DE. However if one can find a single non-trivial solution of the homogeneous DE, the situation is quite different.

Suppose that y_1 is a solution of $y'' = p_0(x)y + p_1(x)y'$. Look for a solution of the form $y = c(x)y_1(x)$.

$$\begin{aligned} y &= cy_1 \\ y' &= cy_1' + c'y_1 \\ y'' &= cy_1'' + 2c'y_1' + c''y_1 \end{aligned}$$

Therefore

$$\begin{aligned} 0 = y'' - p_1y' - p_0y &= c(y_1'' - p_1y_1' - p_0y_1) + c'(2y_1' - p_1y_1) + c''y_1 \\ &= c'(2y_1' - p_1y_1) + c''y_1. \end{aligned}$$

This is a first order separable DE in c'. So let $z = c'$. We have

$$\frac{z'}{z} = -\frac{2y_1' - p_1y_1}{y_1} = -2\frac{y_1'}{y_1} + p_1.$$

Integrating we obtain

$$\ln|z| = -2\ln|y_1(x)| + \int p_1(x)\,dx.$$

If we set $P(x) + c = \int p_1(x)\,dx$, we obtain

$$c' = z = Cy_1(x)^{-2}e^{P(x)}.$$

Thus

$$c(x) = \int \frac{e^{P(x)}}{y_1(x)^2}\,dx.$$

This yields a second solution $y_2 = c(x)y_1$.

11.3.6. EXAMPLE. Consider $x^2y'' + xy' - y = 0$. In standard form, we have $y'' = x^{-2}y - x^{-1}y$. We observe (by inspection) that $y_1 = x$ is a solution. Then $p_1(x) = -\frac{1}{x}$, so $P(x) = -\ln x$. We get

$$c(x) = C\int \frac{x^{-1}}{x^2}\,dx = \frac{-C}{2x^2}.$$

So $y_2(x) = \frac{-C}{2x^2}y_1(x) = \dfrac{-C}{2x}$. Thus we can choose C so that $y_2 = \dfrac{1}{x}$. If you plug this in, we see

$$x^2y_2'' + xy_2' - y_2 = x^2\Big(\frac{2}{x^3}\Big) + x\Big(\frac{-1}{x^2}\Big) - \frac{1}{x} = 0.$$

Thus the general homogeneous solution is

$$y(x) = c_1x + \frac{c_2}{x}.$$

11.3.7 Variation of parameters. Once we have two linearly independent solutions y_1 and y_2 for the homogeneous DE, there is a method to solve the DE with forcing term.

$$y'' = p_0(x)y + p_1(x)y' + q(x).$$

We search for a function of the form $y = c_1y_1 + c_2y_2$ where c_i are unknown functions. Then

$$y' = (c_1y_1' + c_2y_2') + (c_1'y_1 + c_2'y_2).$$

At this stage, we specify $c_1'y_1 + c_2'y_2 = 0$. We will have to satisfy this equation. Then

$$y'' = (c_1y_1'' + c_2y_2'') + (c_1'y_1' + c_2'y_2').$$

Therefore

$$\begin{aligned} &y'' - p_1(x)y' - p_0(x)y \\ &\quad = c_1(y_1''-p_1(x)y_1'-p_0(x)y_1)+c_2(y_2''-p_1(x)y_2'-p_0(x)y_2)+(c_1'y_1'+c_2'y_2') \\ &\quad = c_1'y_1' + c_2'y_2'. \end{aligned}$$

Hence we want $c_1'y_1' + c_2'y_2' = q(x)$. This leaves us with a *linear system* of DEs to solve:

$$\begin{aligned} c_1'y_1 + c_2'y_2 &= 0 \\ c_1'y_1' + c_2'y_2' &= q(x). \end{aligned}$$

We rewrite this as

$$\begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix} \begin{bmatrix} c_1' \\ c_2' \end{bmatrix} = \begin{bmatrix} 0 \\ q(x) \end{bmatrix}.$$

Such a 2×2 linear system is easy to solve for the unknowns c_1' and c_2'. Let

$$W(x) = \det \begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix} = y_1y_2' - y_1'y_2.$$

This is called the *Wronskian*. The solution is

$$\begin{bmatrix} c_1' \\ c_2' \end{bmatrix} = \frac{1}{W(x)} \begin{bmatrix} y_2' & -y_2 \\ -y_1' & y_1 \end{bmatrix} \begin{bmatrix} 0 \\ q(x) \end{bmatrix} = \begin{bmatrix} -y_2q(x)/W(x) \\ y_1q(x)/W(x) \end{bmatrix}.$$

Therefore a solution of the DE is given by

$$y(x) = -y_1(x) \int \frac{y_2q(x)}{W(x)}\,dx + y_2(x) \int \frac{y_1q(x)}{W(x)}\,dx.$$

11.3.8. EXAMPLE. Consider $y'' = x^{-2}y - x^{-1}y + e^x$ with boundary conditions $y(1) = 2 + e$ and $y'(1) = 2$. We saw in Example 11.3.6 that $y_1 = x$ and $y_2 = x^{-1}$ are solutions to the homogeneous DE. The Wronskian is

$$W(x) = x(-x^{-2}) - 1x^{-1} = -2x^{-1}.$$

Thus

$$\begin{aligned} y(x) &= -x \int \frac{x^{-1}e^x}{-2x^{-1}}\,dx + \frac{1}{x} \int \frac{xe^x}{-2x^{-1}}\,dx \\ &= \frac{x}{2} \int e^x\,dx - \frac{1}{2x} \int x^2e^x\,dx \\ &= \tfrac{1}{2}xe^x + c_1x - \frac{1}{2x}(x^2 - 2x + 2)e^x + c_2x^{-1} \\ &= \Big(1 - \frac{1}{x}\Big)e^x + c_1y_1 + c_2y_2. \end{aligned}$$

You can check by hand that this is indeed the solution. Now if we apply the boundary conditions, we have

$$0 + c_1 + c_2 = 2 + e$$

and

$$y'(1) = \Big(1 - \frac{1}{x} + \frac{1}{x^2}\Big)e^x + c_1 - \frac{c_2}{x^2}\Big|_{x=1} = e + c_1 - c_2 = 2.$$

Thus $c_1 = 2$ and $c_2 = e$. The solution is

$$y(x) = \Big(1 - \frac{1}{x}\Big)e^x + 2x + \frac{e}{x}.$$

11.4. Linear DEs with constant coefficients

A linear DE has *constant coefficients* if $p_0(x) = a_0$ and $p_1(x) = a_1$ are constants.

$$y'' = a_0 y + a_1 y' + q(x).$$

We first deal with the homogeneous case. Consider the quadratic polynomial

$$p(t) = t^2 - a_1 t - a_0.$$

11.4.1 Two distinct real roots. Suppose that $p(t) = (t - r_1)(t - r_2)$ where $r_1 \neq r_2 \in \mathbb{R}$. We define a linear map $D : C^1[a, b] \to C[a, b]$ by $Df = f'$. Then define L by

$$L(y) := y'' - a_1 y' - a_0 y = (D^2 - a_1 D - a_0 I)y$$

where $If = f$ is the identity map. The idea is to factor

$$D^2 - a_1 D - a_0 I = (D - r_1 I)(D - r_2 I).$$

We can solve the homogeneous DE by solving $(D - r_1 I)y = 0$ and $(D - r_2 I)y = 0$. Note that $y_1' = r_1 y_1$ implies that $y_1(x) = ce^{r_1 x}$. Thus

$$y_1'' - a_1 y_1' - a_0 y_1 = ce^{r_1 x}(r_1^2 - a_1 r_1 - a_0) = 0.$$

Similarly $y_2' = r_2 y_2$ implies that $y_2(x) = ce^{r_2 x}$; and this is also a solution of the homogeneous DE. Since y_1 and y_2 are not linearly dependent, they span the set of solutions for the homogeneous DE. The most general solution is

$$y = c_1 e^{r_1 x} + c_2 e^{r_2 x}.$$

11.4.2 Double real root. Suppose that $p(t) = (t - r)^2 = t^2 - 2rt + r^2$. Again we see that $y_1 = e^{rx}$ is a solution. To find another one, we look for a solution of the form $y = c(x)y_1(x)$. Then

$$\begin{aligned} 0 = y'' - 2ry' + r^2 y &= (c'' y_1 + 2c' y_1' + c y_1'') - 2r(c' y_1 + c y_1') + r^2 c y_1 \\ &= c'' y_1 + 2c'(y_1' - r y_1) + c(y_1'' - 2r y_1' + r^2 y_1) \\ &= e^{rx} c''. \end{aligned}$$

Therefore $c'' = 0$, and thus $c(x) = c_1 x + c_0$. So the general solution is

$$y(x) = c_1 x e^{rx} + c_0 e^{rx}$$

This forms a two dimensional subspace of solutions.

11.4.3 Complex roots. Suppose that $p(t) = t^2 - 2rt + (r^2 + s^2)$ for $s > 0$. Then the roots of $p(t)$ are $r \pm is$. The first case suggests that $y(x) = e^{(r\pm is)x}$ should be solutions, but these are not real functions. However $\dfrac{e^{isx} + e^{-isx}}{2} = \cos sx$ and $\dfrac{e^{isx} - e^{-isx}}{2i} = \sin sx$ are real functions. Compute for $y_1 = e^{rx} \sin sx$, that $y_1' = e^{rx}(r \sin sx + s \cos sx)$ and $y_1'' = e^{rx}(r^2 \sin sx + 2rs \cos sx - s^2 \sin sx)$. Thus

$$\begin{aligned}
y_1''-2ry_1'+(r^2+s^2)y_1 &= e^{rx}(r^2 \sin sx+2rs \cos sx-s^2 \sin sx)\\
&\quad -2re^{rx}(r \sin sx+s \cos sx)+r^2e^{rx} \sin sx\\
&= e^{rx} \sin sx(r^2-s^2-2r^2+r^2)+e^{rx} \cos sx(2rs-2rs)=0.
\end{aligned}$$

Similarly, $y_2 = e^{rx} \cos sx$ is a solution. Thus the general solution has the form

$$y = c_1e^{rx} \sin sx + c_2e^{rx} \cos sx.$$

Sometimes it is more useful to write this as

$$y = ce^{rx} \cos(sx - \varphi)$$

where $c = \sqrt{c_1^2 + c_2^2}$ and $\sin \varphi = \frac{c_1}{c}$ and $\cos \varphi = \frac{c_2}{c}$.

11.4.4. REMARK. The method for solving a homogeneous DE with constant coefficients of higher order works in a similar manner.

11.4.5. EXAMPLE. Consider $y'' + y = \csc x$ and $y(\frac{\pi}{2}) = y'(\frac{\pi}{2}) = 0$ on $(0, \pi)$. The homogeneous DE has linearly independent solutions $y_1 = \cos x$ and $y_2 = \sin x$. We use the variation of parameters method to find the solution with the forcing term. The Wronskian is $W(x) = \cos^2 x - (-\sin^2 x) = 1$. Therefore

$$\begin{aligned}
y(x) &= -\cos x \int \sin x \csc x \, dx + \sin x \int \cos x \csc x \, dx\\
&= -x \cos x + \sin x \ln \sin x + c_1 \cos x + c_2 \sin x.
\end{aligned}$$

Thus

$$\begin{aligned}
y' &= -\cos x + x \sin x + \cos x \ln \sin x - \cos x - c_1 \sin x + c_2 \cos x\\
&= x \sin x + \cos x \ln \sin x - c_1 \sin x + c_2 \cos x.
\end{aligned}$$

Plugging in $x = \frac{\pi}{2}$, we obtain

$$\begin{aligned}
c_2 &= 0\\
\tfrac{\pi}{2} - c_1 &= 0.
\end{aligned}$$

Therefore the solution is $y(x) = \left(\frac{\pi}{2} - x\right) \cos x + \sin x \ln \sin x$ on $(0, \pi)$.

11.4.6 Undetermined coefficients. When the forcing term for a linear DE with constant coefficients is an exponential, sine, cosine or polynomial, it is often possible to 'guess' a particular solution using the method of *undetermined coefficients*. It works because the derivatives have a similar form. Consider the DE

$$y'' - a_1 y' - a_0 y = e^{bx}.$$

Set $y(x) = ce^{bx}$. Then

$$y'' - a_1 y' - a_0 y = c(b^2 - a_1 b - a_0)e^{bt} = cp(b)e^{bt}.$$

Therefore taking $y = \dfrac{1}{p(b)}e^{bt}$ yields a particular solution. The general solution is obtained by adding the general solution to the homogeneous equation. This fails to work if b is a root of $p(t)$. But you can still use variation of parameters here.

11.4.7. EXAMPLE. In this example, we consider an extremely common form of motion, *damped harmonic oscillation*. We consider a weight such as a trolley attached to a stiff spring. When the trolley is moved away from the equilibrium position, *Hooke's law* states that the force of the spring acting on the trolley to return it to equilibrium is proportional to the distance from equilibrium. In addition, we consider a damping force from friction, which is generally assumed to be proportional to velocity.

We choose coordinates so that equilibrium occurs at $x = 0$ and that the position of the trolley at time t is $x(t)$. Then *Newton's law* $F = ma$ leads to

$$mx'' = -kx - dx'.$$

The negative signs are chosen so that we may assume that k and d are positive, and both forces try to restore the trolley to equilibrium. Let's also suppose that at time $t = 0$, the trolley is moved to a position $x = x_0$ and $x'(0) = 0$. This is a homogeneous second order linear DE. We set $p(t) = mt^2 + dt + k$. It has roots $\dfrac{-d + \sqrt{d^2 - 4mk}}{2m}$. We split the analysis of the solution into three cases:

Case 1. $d^2 > 4mk$. In this case the two roots are real and negative, say r_1 and r_2. The general solution is

$$x(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}.$$

At $t = 0$, we have $x(0) = c_1 + c_2 = x_0$ and $x'(0) = r_1 c_1 + r_2 c_2 = 0$. This system of two linear equations in the unknowns c_1 and c_2 has solution

$$c_1 = \frac{r_2}{r_2 - r_1} x_0 \quad \text{and} \quad c_2 = \frac{-r_1}{r_2 - r_1} x_0.$$

Therefore

$$x(t) = \frac{x_0}{r_2 - r_1}(r_2 e^{r_1 t} - r_1 e^{r_2 t}).$$

In this case, the trolley just returns asymptotically to equilibrium. This is known as an overdamped system.

<u>Case 2.</u> $d^2 = 4mk$. Then p has a double real root $r = \frac{-d}{2m} = -\sqrt{\frac{k}{m}} < 0$. The analysis is similar: the general solution is

$$x(t) = c_1 e^{rt} + c_2 t e^{rt}.$$

At $t = 0$, we have $x(0) = c_1 = x_0$ and $x'(0) = rc_1 + c_2 = 0$. This system has solution $c_1 = x_0$ and $c_2 = -rx_0 = \dfrac{dx_0}{2m} = \sqrt{\frac{k}{m}}x_0$. Therefore

$$x(t) = x_0\big(1 + \sqrt{\tfrac{k}{m}}\,t\big)e^{-\sqrt{k/m}\,t}.$$

Thus

$$x'(t) = -\frac{kx_0}{m}e^{-\sqrt{k/m}\,t} < 0.$$

So in this case, the trolley also returns asymptotically to equilibrium.

<u>Case 3.</u> $d^2 < 4mk$. This is the most interesting case, and explains the name oscillator. Here p has two complex roots: $r \pm is$ where $r = \frac{-d}{2m} < 0$ and $s = \frac{1}{2m}\sqrt{4km - d^2}$. The general solution is

$$x(t) = ce^{rt}\cos(st - \varphi) = ce^{-dt/2m}\cos(st - \varphi).$$

The initial conditions yield $x_0 = x(0) = c\cos\varphi$ and

$$0 = x'(0) = c(-\frac{d}{2m}\cos\varphi + s\sin\varphi).$$

Thus $\tan\varphi = \frac{d}{2ms} = \frac{d}{\sqrt{4mk-d^2}}$ and $c = x_0\sec\varphi = x_0\sqrt{\frac{4km}{4mk-d^2}}$. The solution oscillates above and below the equilibrium with period $\dfrac{2\pi}{s} = \dfrac{4\pi m}{\sqrt{4km - d^2}}$. The damping factor $e^{-dt/2m}$ shows that the trolley approaches equilibrium asymptotically.

Now let's suppose we are in Case 3, and that there is another forcing term which is periodic, such as $f\cos\omega t$. Then the DE becomes

$$mx'' + dx' + kx = f\cos\omega t.$$

This is a candidate for the technique of undetermined coefficients. Try for a solution $y = a\cos\omega t + b\sin\omega t$. Then

$$y' = -a\omega\sin\omega t + b\omega\cos\omega t \quad \text{and} \quad y'' = -a\omega^2\cos\omega t - b\omega^2\sin\omega t.$$

Thus

$$\begin{aligned} my'' + dy' + ky &= (-ma\omega^2 + db\omega + ka)\cos\omega t + (-mb\omega^2 - da\omega + kb)\sin\omega t \\ &= f\cos\omega t. \end{aligned}$$

This yields the linear system

$$\begin{bmatrix} k - m\omega^2 & d\omega \\ -d\omega & k - m\omega^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}.$$

We get

$$y = \frac{f(k-m\omega^2)}{(k-m\omega^2)^2+d^2\omega^2}\cos\omega t + \frac{fd\omega}{(k-m\omega^2)^2+d^2\omega^2}\sin\omega t + ce^{-dt/2m}\cos(st-\varphi).$$

Now we solve for c and φ to satisfy the initial conditions. What we want to point out does not depend on the exact constants. Notice that over time, the homogeneous part of the solution is damped out, while the periodic motion from the external force continues unabated. An example of this behaviour is a child on a swing. The pumping action is a periodic force. The initial start is a bit irregular, but the pumping action quickly becomes dominant. As long as the pumping continues, the swing will move back and forth in a regular motion.

11.5. Uniqueness of solutions for 2nd order linear DEs

We will not establish a general existence theorem for linear DEs. A more general result known as *Picard's Theorem* is generally taught is a theoretical DE course or in a good real analysis course. We can prove uniqueness though.

11.5.1. GRONWALL'S INEQUALITY. *Suppose that a differentiable function $f(x)$ satisfies $f'(x) \leqslant Cf(x)$ on $[a,b]$ for some constant C. Then*

$$f(x) \leqslant f(a)e^{C(x-a)} \quad \textit{for} \quad a \leqslant x \leqslant b.$$

PROOF. Let $g(x) = f(x)e^{-Cx}$. Then

$$g'(x) = e^{-Cx}(f'(x) - Cf(x)) \leqslant 0 \quad \text{for} \quad a \leqslant x \leqslant b.$$

Therefore g is a decreasing function on $[a,b]$. Thus,

$$f(x)e^{-Cx} \leqslant f(a)e^{-Ca} \quad \text{or} \quad f(x) \leqslant f(a)e^{C(x-a)}. \qquad \blacksquare$$

11.5.2. THEOREM. *The DE $y'' = p_0(x)y + p_1(x)y' + q(x)$ with $y(a) = \gamma_0$ and $y'(a) = \gamma_1$ has at most one solution.*

PROOF. Suppose that y_1 and y_2 are two solutions. Then by Lemma 11.3.1, $y = y_1 - y_2$ satisfies $y'' = p_0(x)y + p_1(x)y'$ and $y(a) = y'(a) = 0$. Define $s(x) = y^2 + (y')^2$; so that $s \geqslant 0$ and $s(a) = 0$. Compute

$$\begin{aligned} s'(x) &= 2yy' + 2y'y'' = 2y'(y + y'') \\ &= 2y'(y + p_0(x)y + p_1(x)y') \\ &= 2p_1(x)(y')^2 + 2(1 + p_0(x))yy' \\ &\leqslant 2|p_1(x)|(y')^2 + (1 + |p_0(x)|)(y^2 + (y')^2). \end{aligned}$$

The last inequality uses the AMGM inequality $|2yy'| \leqslant y^2 + (y')^2$. Let

$$C = \max\{2|p_1(x)|, 1 + |p_0(x)| : a \leqslant x \leqslant b\}.$$

Then $s'(x) \leqslant Cs(x)$. By Gronwall's inequality, $s(x) \leqslant s(a)e^{C(x-a)} = 0$. Therefore $s = 0$ and hence $y = 0$. That is, $y_1 = y_2$; the solution is unique. ■

11.5.3. COROLLARY. *If the homogeneous linear DE $y'' = p_0(x)y + p_1(x)y'$ has two solutions y_1 and y_2 such that $(y_1(a), y_1'(a))$ and $(y_2(a), y_2'(a))$ are linearly independent, then every solution of the DE is a linear combination of y_1 and y_2. So the vector space of solutions is at most 2 dimensional.*

PROOF. Let y_0 be a solution of the homogeneous DE. Since $(y_1(a), y_1'(a))$ and $(y_2(a), y_2'(a))$ are linearly independent, they span $\mathbb{R}^2$. Therefore there are constants c_1 and c_2 so that

$$(y_0(a), y_0'(a)) = c_1(y_1(a), y_1'(a)) + c_2(y_2(a), y_2'(a)).$$

Then by Lemma 11.3.1, $y_3 = c_1y_1 + c_2y_2$ satisfies

$$y_3'' = p_0(x)y_3 + p_1(x)y_3' \quad \text{and} \quad y_3(a) = y(a), \;\; y_3'(a) = y'(a).$$

By Theorem 11.5.2, this DE has a unique solution, and thus $y_0 = y_3 = c_1y_1 + c_2y_2$. Therefore every solution is a linear combination of y_1 and y_2. ■

Exercises for Chapter 11

1. (a) Solve the DE $(x^2 - y^2)y' - 2xy = 0$.
(b) Sketch the set of solution curves.
(c) If $(1, 2)$ is a point on the solution curve, find the solution.

2. (a) Solve the DE $(1 + x^2)y' + 2xy = 0$.
(b) Solve the DE $(1 + x^2)y' + 2xy = \cot x$.

3. (a) Consider the DE $y' + p(x)y = q(x)y^a$ where $a \notin \{0, 1\}$ is a real number. Set $z = y^{1-a}$. Turn this DE into a linear DE in z.
(b) Use this to solve $xy^2y' + y^3 = x\cos x$.

4. A function defined on $\mathbb{R}$ satisfies the DE $f'(x) = 2xf(x) + 4x$ and $f(0) = 1$. Assume that $f(x)$ has a power series about $x = 0$ and solve for the Taylor coefficients.

5. Torricelli's Law for fluid flowing out the bottom of a tank states that the velocity is calculated as if the fluid dropped from the surface of the water. A hemispherical container of radius R is completely full of water. A small round hole of radius r at the bottom is unplugged. (You can ignore the small difference in the height caused by the hole.)

(a) Use Newtonian mechanics to deduce how long it takes a drop of water to fall distance h from a position of rest, and use this to compute the velocity at that point.
(b) Compute the volume $V(h)$ in the bowl when the water depth is h.
(c) How long does it take the bowl to empty?
HINT: compute $\frac{dV}{dt}$ in two ways. Get a DE in h.

6. Consider the homogeneous linear DE $y'' + p(x)y' + q(x)y = 0$. Suppose that y_1 and y_2 are two solutions on $[a, b]$.
(a) Find a first order DE satisfied by the Wronskian $W(x)$ and solve it.
(b) Prove that if the vectors $(y_1(c), y_1'(c))$ and $(y_2(c), y_2'(c))$ are linearly independent for some $c \in [a, b]$, then $W(x)$ never vanishes. Hence show that $(y_1(x), y_1'(x))$ and $(y_2(x), y_2'(x))$ are linearly independent vectors for every $x \in [a, b]$.

APPENDICES

A.1. Equivalence Relations

We introduce a basic mathematical construction known as an equivalence relation. Equivalence relations occur frequently in mathematics and have appeared occasionally in these notes.

A.1.1. DEFINITION. Let X be a set, and let R be a subset of $X \times X$. Then R is a *relation* on X. Let us write $x \sim y$ if $(x, y) \in R$. We say that R or $\sim$ is an *equivalence relation* if it is

(1) **(reflexive)** $x \sim x$ for all $x \in X$.

(2) **(symmetric)** if $x \sim y$ for $x, y \in X$, then $y \sim x$.

(3) **(transitive)** if $x \sim y$ and $y \sim z$ for y $x, y, z \in X$, then $x \sim z$.

If $\sim$ is an equivalence relation on X and $x \in X$, then the *equivalence class* $[x]$ is the set $\{y \in X : y \sim x\}$. By $X/\sim$ we mean the collection of all equivalence classes.

A.1.2. EXAMPLES.

(1) Equality is an equivalence relation on any set. Verify this.

(2) Consider the integers $\mathbb{Z}$. Say that $m \equiv n \pmod{12}$ if 12 divides $m - n$. Note that 12 divides $n - n = 0$ for any n, and thus $n \equiv n \pmod{12}$. So it is reflexive. Also if 12 divides $m - n$, then it divides $n - m = -(m - n)$. So $m \equiv n \pmod{12}$ implies that $n \equiv m \pmod{12}$ (i.e., symmetry). Finally, if $l \equiv m \pmod{12}$ and $m \equiv n \pmod{12}$, then we may write $l - m = 12a$ and $m - n = 12b$ for certain integers a, b. Thus $l - n = (l - m) + (m - n) = 12(a + b)$ is also a multiple of 12. Therefore, $l \equiv n \pmod{12}$, which is transitivity.

There are twelve equivalence classes $[r]$ for $0 \leqslant r < 12$ determined by the remainder r obtained when n is divided by 12. So $[r] = \{12a + r : a \in \mathbb{Z}\}$.

(3) Consider the set $\mathbb{R}$ with the relation $x \leqslant y$. This relation is reflexive ($x \leqslant x$) and transitive $x \leqslant y$ and $y \leqslant z$ implies $x \leqslant z$. However, it is *antisymmetric*: $x \leqslant y$ and $y \leqslant x$ both occur if and only if $x = y$. This is not an equivalence relation.

When dealing with functions defined on equivalence classes, we often define the function on an equivalence class in terms of a representative. In order for the function to be well defined, that is, for the definition of the function to make sense, we must check that we get same value regardless of which representative is used.

A.1.3. EXAMPLES.

(1) Consider the set of real numbers $\mathbb{R}$. Say that $x \equiv y \pmod{2\pi}$ if $x - y$ is an integer multiple of 2π. Verify that this is an equivalence relation. Define a function $f([x]) = (\cos x, \sin x)$. We are really defining a function $F(x) = (\cos x, \sin x)$ on $\mathbb{R}$ and asserting that $F(x) = F(y)$ when $x \equiv y \pmod{2\pi}$. Indeed, we then have $y = x + 2\pi n$ for some $n \in \mathbb{Z}$. As sin and cos are 2π-periodic, we have

$$\begin{aligned} F(y) &= (\cos y, \sin y) \\ &= (\cos(x + 2\pi n), \sin(x + 2\pi n)) \\ &= (\cos x, \sin x) = F(x). \end{aligned}$$

It follows that the function $f([x]) = F(x)$ yields the same answer for every $y \in [x]$. So f is well defined. One can imagine the function f as wrapping the real line around the circle infinitely often, matching up equivalent points.

(2) Consider $\mathbb{R}$ modulo 2π again, and look at $f([x]) = e^x$. Then $0 \equiv 2\pi \pmod{2\pi}$ but $e^0 = 1 \neq e^{2\pi}$. So f is not well defined on equivalence classes.

(3) Now consider Example A.1.2(2). We wish to define multiplication modulo 12 by $[n][m] = [nm]$. To check that this is well defined, consider two representatives $n_1, n_2 \in [n]$ and two representatives $m_1, m_2 \in [m]$. Then there are integers a and b so that $n_2 = n_1 + 12a$ and $m_2 = m_1 + 12b$. Then

$$\begin{aligned} n_2 m_2 &= (n_1 + 12a)(m_1 + 12b) \\ &= n_1 m_1 + 12(a m_1 + n_1 b + 12ab). \end{aligned}$$

Therefore, $n_2 m_2 \equiv n_1 m_1 \pmod{12}$. Consequently, multiplication modulo 12 is well defined.

A.2. A Construction of $\mathbb{R}$

Our description of the real numbers as the set of all infinite decimals modulo the issue with terminal 9's versus terminal 0's, (which is an equivalence relation!), was a bit problematic because the rules for addition and multiplication were very hard to formulate. Now that you have read Chapter 2, we can introduce a superior method for constructing the reals. The idea is to start with the rational numbers, $\mathbb{Q}$, and *complete* it.

Let $\mathcal{C}$ denote the set of all Cauchy sequences of rational numbers. Note that in the definitions of Cauchy sequence and limit, there is no harm in using only rational numbers for ε. Put an equivalence relation on $\mathcal{C}$ by $(x_n) \sim (y_n)$ if $\lim\limits_{n\to\infty} x_n - y_n = 0$. The three properties of an equivalence relation are very easy to check. Let $\mathcal{R} = \mathcal{C}/\sim$ be the set of equivalence classes. We define an imbedding J of $\mathbb{Q}$ into $\mathcal{R}$ by $J(r) = [(r, r, r, r, \dots)]$, the equivalence class of the constant sequence.

Next we define addition, multiplication and order.

- $[(x_n)] + [(y_n)] = [(x_n + y_n)]$.
- $[(x_n)] \cdot [(y_n)] = [(x_n y_n)]$.
- If $[(x_n)] \neq [(y_n)]$, say $[(x_n)] < [(y_n)]$ if there is some N so that $x_n < y_n$ for all $n \geqslant N$.

We need to verify that these notions are well defined. Suppose that $(x'_n) \sim (x_n)$ and $(y'_n) \sim (y_n)$. Then

$$\lim_{n\to\infty} (x'_n + y'_n) - (x_n + y_n) = \lim_{n\to\infty} x'_n - x_n + \lim_{n\to\infty} y'_n - y_n = 0.$$

Hence $(x'_n + y'_n) \sim (x_n + y_n)$, and thus addition is well defined. Similarly

$$\lim_{n\to\infty} (x'_n y'_n) - (x_n y_n) = \lim_{n\to\infty} (x'_n - x_n) y'_n + \lim_{n\to\infty} x_n (y'_n - y_n) = 0.$$

This uses that Cauchy sequences are bounded by Proposition 2.6.3.

The order is a bit more delicate. If $[(x_n)] \neq [(y_n)]$, then $x_n - y_n$ does not converge to 0. Thus there is some $\varepsilon > 0$ so that $|x_{n_i} - y_{n_i}| \geqslant \varepsilon$ for some subsequence. Use $\varepsilon/3$ in the definition of Cauchy sequence to obtain N so that for $N \leqslant m \leqslant n$,

$$|x_n - x_m| < \frac{\varepsilon}{3} \quad \text{and} \quad |y_n - y_m| < \frac{\varepsilon}{3}.$$

Then choose some $n_i \geqslant N$, and for definiteness, suppose that $x_{n_i} - y_{n_i} \geqslant \varepsilon$. (The other case is similar.) Then for $n \geqslant N$,

$$x_n - y_n = (x_{n_i} - y_{n_i}) - (x_{n_i} - x_n) - (y_n - y_{n_i} \geqslant \varepsilon - \frac{\varepsilon}{3} - \frac{\varepsilon}{3} = \frac{\varepsilon}{3}.$$

Hence $[(y_n)] < [(x_n)]$. Now take equivalent sequences (x'_n) and (y'_n). Then

$$\liminf x'_n - y'_n = \liminf (x'_n - x_n) + (x_n - y_n) + (y'_n - y_n) \geqslant 0 + \frac{\varepsilon}{3} + 0 = \frac{\varepsilon}{3}.$$

This shows that order is well defined. Moreover, we see that exactly one of

$$[(x_n)] < [(y_n)], \quad [(x_n)] = [(y_n)] \quad \text{or} \quad [(x_n)] > [(y_n)]$$

holds. You should check that $[(x_n)] < [(y_n)]$ and $[(y_n)] < [(z_n)]$ implies that $[(x_n)] < [(z_n)]$.

Next observe that the embedding $J : \mathbb{Q} \to \mathcal{R}$ preserves the ordered field properties:

$$J(r) + J(s) = J(r + s) \quad \text{and} \quad J(r) \cdot J(s) = J(rs) \quad \text{for} \quad r, s \in \mathbb{Q},$$

and $r < s$ implies that $J(r) < J(s)$. We let $\mathbf{0} = J(0)$ and $\mathbf{1} = J(1)$.

It is straightforward to verify all of the field operations. We give two examples. The distributive law follows from the corresponding property for $\mathbb{Q}$:

$$\begin{aligned}\big([(x_n)] + [(y_n)]\big) \cdot [(z_n)] &= [(x_n + y_n)] \cdot [(z_n)] = \big[\big((x_n + y_n) z_n\big)\big] \\ &= [(x_n z_n + y_n z_n)] = [(x_n z_n)] + [(y_n z_n)] \\ &= [(x_n)] \cdot [(z_n)] + [(y_n)] \cdot [(z_n)].\end{aligned}$$

Multiplicative inverses take a bit of work. If $[(x_n)] \neq \mathbf{0}$, then there is some ε_0 so that $|x_n| \geqslant \varepsilon_0$ for $n \geqslant N$. Let $y_n = 0$ for $n < N$ and $y_n = x_n^{-1}$ for $n \geqslant N$. We have to show that this is Cauchy. Choose $N_1 \geqslant N$ so that if $N_1 \leqslant m \leqslant n$, then $|x_n - x_m| < \varepsilon\varepsilon_0^2$. Then

$$|y_n - y_m| = \Big|\frac{1}{x_n} - \frac{1}{x_m}\Big| = \frac{|x_m - x_n|}{|x_n x_m|} < \frac{\varepsilon\varepsilon_0^2}{\varepsilon_0^2} = \varepsilon.$$

Thus $(y_n) \in \mathcal{R}$ and $[(x_n)] \cdot [(y_n)] = \mathbf{1}$.

We leave the straightforward verification of the other properties of an ordered field to the interested reader.

There is more work to be done. The ordered field $\mathcal{R}$ is *Archimedean*: if $[(x_n)] > \mathbf{0}$, then there is an $r \in \mathbb{Q}$ so that $\mathbf{0} < J(r) < [(x_n)]$. This follows because we showed that $\liminf x_n \geqslant \varepsilon > 0$ for some rational ε, and so any $0 < r < \varepsilon$ will work. If we like, we can take $r = \frac{1}{k}$ for $k \in \mathbb{N}$ sufficiently large. This also implies that that if $[(x_n)] \in \mathcal{R}$, then there is an integer $k \in \mathbb{N}$ so that $[(x_n)] < J(k)$. Indeed, if $[(x_n)] \leqslant \mathbf{0}$, then $k = 1$ suffices. If $\mathbf{x} = [(x_n)] > \mathbf{0}$, then $\mathbf{x}^{-1} > \mathbf{0}$. By the Archimedean property, $J(\frac{1}{k}) < \mathbf{x}^{-1}$ for some $k \in \mathbb{N}$, and thus $\mathbf{x} < J(k)$.

We need to verify the Least Upper Bound Property for $\mathcal{R}$. Let $\mathcal{S} \subset \mathcal{R}$ be a nonempty set which is bounded above by $z \in \mathcal{R}$, and let $s \in S$. Since R is Archimedean, we can find integers a, b so that $a < s \leqslant z < b$. Recursively define sequences x_n and y_n of *rational number*s as follows. Let $x_1 = a$ and $y_1 = b$. Suppose that x_i and y_i have been defined in $\mathbb{Q}$ for $1 \leqslant i < n$ so that $J(x_i)$ is not an upper bound for S and $J(y_i)$ is an upper bound for $\mathcal{S}$ and $y_i - x_i = 2^{1-i}(b-a)$. Let $c_n = \frac{1}{2}(x_{n-1} + y_{n-1})$. If c_n is an upper bound for S, then let $x_n = x_{n-1}$ and $y_n = c_n$; while if c_n is not an upper bound for S, then let $x_n = c_n$ and $y_n = y_{n-1}$. Let $\mathbf{x} = [(x_n)]$. Then $\mathbf{x} = [(y_n)]$ because $\lim_{n\to\infty} y_n - x_n = 0$. We claim that $\sup \mathcal{S} = \mathbf{x}$.

Let $\mathbf{s} = [(s_n)] \in \mathcal{S}$. If $\mathbf{s} > \mathbf{x}$, then by the Archimedean property, $\mathbf{s} > \mathbf{x} + J(\frac{1}{k})$ for some $k \in \mathbb{N}$. So there is an integer N so that $s_n > y_n + \frac{1}{2k}$ for all $n \geqslant N$. Choose $M \geqslant N$ so that $2^{1-M}(b-a) < \frac{1}{4k}$. Then for $n \geqslant M$

$$y_n = y_M + \sum_{i=M+1}^{m} (y_i - y_{i-1}) < y_M + \sum_{i=M+1}^{m} 2^{1-i}(b-a) < y_M + \frac{1}{4k}.$$

Therefore for $n \geqslant M$, we have $s_n > y_M + \frac{1}{4k}$; and hence $\mathbf{s} \geqslant J(y_M) + J(\frac{1}{4k})$. This contradicts the fact that $J(y_M)$ is an upper bound for $\mathcal{S}$. So no such $\mathbf{s}$ exists, and $\mathbf{x}$ is an upper bound for $\mathcal{S}$. A similar argument shows that if $\mathbf{z} < \mathbf{x}$, then $\mathbf{z}$ is not an upper bound.

It follows that $\mathcal{R}$ is an ordered field with the Least Upper Bound Property. This is exactly the property of $\mathbb{R}$ that we used to establish the various versions of completeness. We call this field the real numbers, $\mathbb{R}$. It is a subtle point that there is only one such field with these properties. This issue will not be addressed here.

A.3. Cardinality

Cardinality is the notion that measures the size of a set in the crudest of ways—by counting the numbers of elements. Obviously, the number of elements in a set could be 0, 1, 2, 3, 4, or some other finite number. Or a set can have infinitely many elements. Perhaps surprisingly, not all infinite sets have the same cardinality. For our purposes, infinite sets have two possible sizes: countable and uncountable (the uncountable ones are larger). The most important ideas to understand are what *countable* means and what distinguishes countable sets from those with larger cardinality.

A.3.1. DEFINITION. Two sets A and B have the same *cardinality* if there is a *bijection* f from A onto B. We write $|A| = |B|$ in this case. Similarly, we say that the cardinality of A is less than that of B ($|A| \leqslant |B|$) if there is an *injection* f from A into B.

The definition says simply that if all of the elements of A can be paired, one-to-one, with all of the elements of B, then A and B have the same size. If A fits inside B in a one-to-one manner, then A is smaller than B. One of the subtleties that we address later is whether $|A| \leqslant |B|$ and $|B| \leqslant |A|$ mean that $|A| = |B|$. The answer is yes, but this is not obvious for infinite sets.

A.3.2. EXAMPLES.

(1) The cardinality of any finite set is the number of elements, and this number belongs to $\mathbb{N}_0 = \{0, 1, 2, 3, 4, \dots\}$. Set theorists go to some trouble to define the natural numbers too. But we will take for granted that the reader is familiar with the notion of a finite set.

(2) Most sets encountered in analysis are infinite, meaning that they are not finite. The sets of natural numbers $\mathbb{N}$, integers $\mathbb{Z}$, rational numbers $\mathbb{Q}$, and real numbers $\mathbb{R}$ are all infinite. Moreover, we have the natural containments $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$. So $|\mathbb{N}| \leqslant |\mathbb{Z}| \leqslant |\mathbb{Q}| \leqslant |\mathbb{R}|$. Notice that the integers can be written as a list $0, 1, -1, 2, -2, 3, -3, \dots$. This amounts to defining a bijection $f : \mathbb{N} \to \mathbb{Z}$ by

$$f(n) = \begin{cases} (1-n)/2 & \text{if } n \text{ is odd} \\ n/2 & \text{if } n \text{ is even.} \end{cases}$$

Therefore, $|\mathbb{N}| = |\mathbb{Z}|$.

A.3.3. DEFINITION. A set A is a *countable set* is it is finite or if $|A| = |\mathbb{N}|$. The cardinal $|\mathbb{N}|$ is denoted by $\aleph_0$. This is the first letter of the Hebrew alphabet, aleph, with subscript zero. It is pronounced *aleph nought*.

An infinite set that is not countable is called an *uncountable set*.

Equivalently, A is countable and infinite if the elements of A may be listed as $a_1, a_2, a_3, \ldots$. Indeed, the list itself determines a bijection from $\mathbb{N}$ to A by $f(k) = a_k$. It is a basic fact that countable sets are the smallest infinite sets.

Notice that two uncountable sets could have different cardinalities.

A.3.4. LEMMA. *Every infinite subset of $\mathbb{N}$ is countable. Moreover, if A is an infinite set such that $|A| \leqslant |\mathbb{N}|$, then $|A| = |\mathbb{N}|$.*

PROOF. Any nonempty subset X of $\mathbb{N}$ has a smallest element. Indeed, as X is nonempty, it contains an integer n. Consider the elements of the finite set $\{1, 2, \ldots, n\}$ in order and pick the first one that belongs to X—that is, the smallest.

Let B be an infinite subset of $\mathbb{N}$. List the elements of B in increasing order as $b_1 < b_2 < b_3 < \ldots$. This is done by choosing the smallest element b_1, then the smallest of the remaining set $B\backslash\{b_1\}$, then the smallest of $B\backslash\{b_1, b_2\}$ and so on. The result is an infinite list of elements of B in increasing order. It must include every element $b \in B$ because $\{n \in B : n \leqslant b\}$ is finite, containing say k elements. Then $b_k = b$. As noted before the proof, this implies that $|B| = |\mathbb{N}|$.

Now consider a set A with $|A| \leqslant |\mathbb{N}|$. By definition, there is a injection f of A into $\mathbb{N}$. Let $B = f(A)$. Note that f is a bijection of A onto B. Then B is an infinite subset of $\mathbb{N}$. So $|A| = |B| = |\mathbb{N}|$. ■

A.3.5. PROPOSITION. *The countable union of countable sets is countable.*

PROOF. By the previous lemma, we may assume that there is a countably infinite collection of sets $A_1, A_2, A_3, \ldots$ that are each countably infinite. Write the elements of A_i as a list $a_{i,1}, a_{i,2}, a_{i,3}, \ldots$. Then we may write $A = \bigcup_{i \geqslant 1} A_i$ as a list as follows:

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

where the elements $a_{i,j}$ are written so that $i + j$ is monotone increasing, and within the set of pairs (i, j) with $i + j = n$, the terms are written with the i's in increasing order. See Figure 12.1. Thus A is countable. ■

A.3.6. COROLLARY. *The set $\mathbb{Q}$ of rational numbers is countable.*

PROOF. The set $\mathbb{Z} \times \mathbb{N} = \{(i, j) : i \in \mathbb{Z}, j \in \mathbb{N}\}$ is the disjoint union of the sets $A_i = \{(i, j) : j \in \mathbb{N}\}$ for $i \in \mathbb{Z}$. Each A_i is evidently countable. By Example A.3.2(2), $\mathbb{Z}$ is countable. Hence $\mathbb{Z} \times \mathbb{N}$ is the countable union of countable sets, and thus is countable by Proposition A.3.5.

Define a map from $\mathbb{Q}$ into $\mathbb{Z} \times \mathbb{N}$ by $f(r) = (a, b)$ if $r = a/b$, where a and b are integers with no common factor and $b > 0$. These conditions uniquely determine the pair (a, b) for each rational r, and so f is a function. Clearly, f is injective

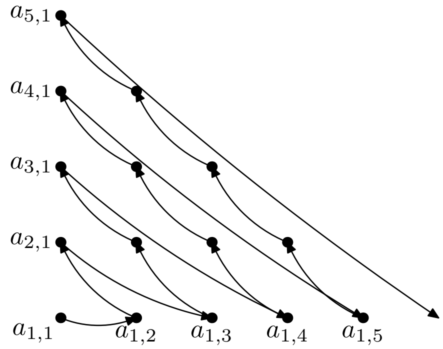


FIGURE 12.1. The set $\mathbb{N} \times \mathbb{N}$ is countable.

since r is recovered from (a, b) by division. Therefore, f is an injection of $\mathbb{Q}$ into a countable set. Hence $\mathbb{Q}$ is an infinite set with $|\mathbb{Q}| \leqslant |\mathbb{N}|$. So $\mathbb{Q}$ is countable by Lemma A.3.4. ■

A.3.7. COROLLARY. *If A and B are countable, then $A \times B$ is countable. Hence $\mathbb{Z}^n$ is countable for all $n \geqslant 1$.*

PROOF. First $A \times B = \bigcup_{b \in B} A \times \{b\}$ is a countable union of countable sets, and thus is countable. In particular, $\mathbb{Z}^2$ is countable. By induction, $\mathbb{Z}^n$ is countable for each $n \geqslant 1$. ■

There are infinite sets that are not countable.

A.3.8. THEOREM. *The set $\mathbb{R}$ of real numbers is uncountable.*

PROOF. The proof uses a *diagonalization* argument due to Cantor. Suppose to the contrary that $\mathbb{R}$ is countable. Then all real numbers may be written as a list $x_1, x_2, x_3, \ldots$. Express each x_i as an infinite decimal, which we write as $x_i = x_{i0}.x_{i1}x_{i2}x_{i3}\ldots$, where x_{i0} is any integer and x_{ik} is an integer from 0 to 9 for each $k \geqslant 1$. Our goal is to write down another real number that does not appear in this (supposedly exhaustive) list. Let $a_0 = 0$ and define $a_k = 7$ if $x_{ik} \in \{0, 1, 2, 3, 4\}$ and $a_k = 2$ if $x_{ik} \in \{5, 6, 7, 8, 9\}$. Define a real number $a = a_0.a_1a_2a_3\ldots$.

Since a is a real number, it must appear somewhere in this list, say $a = x_k$. However, the kth decimal place a_k of a and $x_{k,k}$ of x_k differ by at least 3. This cannot be accounted for by the fact that certain real numbers have two decimal expansions, one ending in zeros and the other ending in nines because this changes any digit by no more than 1 (counting 9 and 0 as being within 1). So $a \neq x_k$, and hence a does not occur in this list. It follows that there is no list containing all real numbers, and thus $\mathbb{R}$ is uncountable. ■

We conclude with the result promised in the start of this section.

A.3.9. Schroeder–Bernstein Theorem. *If A and B are sets with $|A| \leqslant |B|$ and $|B| \leqslant |A|$, then $|A| = |B|$.*

PROOF. The proof is surprisingly simple. Since $|A| \leqslant |B|$, there is an injection f mapping A into B. Likewise, as $|B| \leqslant |A|$, there is an injection g mapping B into A. Let $B_1 = B \backslash f(A)$. Recursively define $A_i = g(B_i)$ and $B_{i+1} = f(A_i)$ for $i \geqslant 1$. Define $A_0 = A \backslash \bigcup_{i \geqslant 1} A_i$ and $B_0 = B \backslash \bigcup_{i \geqslant 1} B_i$. We will show that the actions of f and g fit the scheme of Figure 12.2.

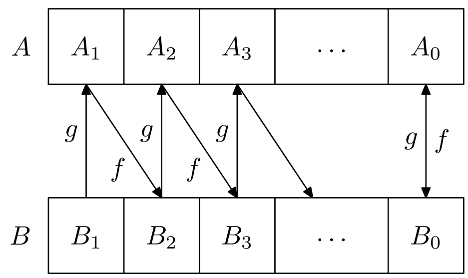


FIGURE 12.2. Schematic of action of f and g on A and B.

First we show that the B_i's are disjoint. Clearly each B_i for $i \geqslant 2$ is in the range of f and hence does not intersect B_1. Suppose that $1 < i < j$. Then $(fg)^{i-1}$ is an injection of B into itself that carries B_k onto B_{k+i-1} for every $k \geqslant 1$. In particular, B_1 is mapped onto B_i and B_{j-i+1} is mapped onto B_j. Since $B_1 \cap B_{j-i+1} = \varnothing$ and $(fg)^{i-1}$ is one-to-one, it follows that $B_i \cap B_j = \varnothing$.

By construction, g^{-1} is a bijection of each A_i onto B_i for $i \geqslant 1$. We claim that f maps A_0 onto B_0. Observe that f maps A_i onto B_{i+1} for each $i \geqslant 1$. Thus the remainder of A, namely A_0, is mapped onto the remainder of the image. Thus

$$f(A_0) = f(A) \backslash \bigcup_{i \geqslant 1} f(A_i) = (B \backslash B_1) \backslash \bigcup_{i \geqslant 1} B_{i+1} = B \backslash \bigcup_{i \geqslant 1} B_i = B_0.$$

This means that the function

$$h(a) = \begin{cases} g^{-1}(a) & \text{if} \quad a \in \bigcup_{i \geqslant 1} A_i \\ f(a) & \text{if} \quad a \in A_0 \end{cases}$$

is a bijection between A and B. Therefore $|A| = |B|$. ∎

A.4. e is transcendental

We first establish a relatively easy fact.

A.4.1. Theorem. *e is irrational.*

PROOF. We use the formula $e = \sum_{k=0}^{\infty} \frac{1}{k!}$. Suppose that $e = \frac{p}{q}$ where $p, q \in \mathbb{N}$. Then $q \geqslant 2$ since $2 < e < 3$. We have that

$$p(q-1)! = q!e = \sum_{k=0}^{q} \frac{q!}{k!} + \sum_{k=q+1}^{\infty} \frac{q!}{k!}$$

is an integer. Therefore we have an *integer*

$$\begin{aligned}\sum_{k=q+1}^{\infty} \frac{q!}{k!} &= \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \frac{1}{(q+1)(q+2)(q+3)} + \dots \\ &< \sum_{k=1}^{\infty} \frac{1}{(q+1)^k} = \frac{\frac{1}{q+1}}{1 - \frac{1}{q+1}} = \frac{1}{q} < 1.\end{aligned}$$

Hence this "integer" lies in $(0, 1)$, which is absurd. Therefore e is irrational. ■

A.4.2. DEFINITION. A real number α is *algebraic* if there is a polynomial $p(x)$ with integer coefficients with α as a root. A real number is *transcendental* if it is not algebraic.

A.4.3. PROPOSITION. *The set of algebraic numbers is countable.*

PROOF. First count the number of polynomials with integer coefficients of degree n. Such a polynomial has $n+1$ coefficients, which we can associate to a point in $\mathbb{Z}^{n+1}$, (or more precisely $\mathbb{Z}^n \times (\mathbb{Z}\backslash\{0\})$ since the leading coefficient $a_n \neq 0$,) which is countable by Corollary A.3.7. Each polynomial has n roots, so the set of roots is countable. Finally we take the union over all n, and the countable union of countable sets is countable. Thus the set of algebraic numbers is countable. ■

The following proof is tricky and pulls several functions out of thin air.

A.4.4. THEOREM. *e is transcendental.*

PROOF. Suppose that e is algebraic. Then there are integers $c_0, \dots, c_n$ so that

$$c_n e^n + c_{n-1} e^{n-1} + \dots + c_1 e + c_0 = 0.$$

We may suppose that $c_n \neq 0 \neq c_0$ because we can factor out an x if $c_0 = 0$. Let p be a prime which is much larger than $\max\{n, |c_0|\}$ to be specified more precisely later. Define a polynomial $f(x)$ of degree $r = (n+1)p - 1$ by

$$\begin{aligned}f(x) &= \frac{1}{(p-1)!} x^{p-1}(1-x)^p(2-x)^p \cdots (n-x)^p = \sum_{k=p-1}^{(n+1)p-1} \frac{a_k}{(p-1)!} x^k \\ &= \frac{(n!)^p}{(p-1)!} x^{p-1} + \frac{a_p}{(p-1)!} x^p + \frac{a_{p+1}}{(p-1)!} x^{p+1} + \dots + \frac{(-1)^n}{(p-1)!} x^{(n+1)p-1}.\end{aligned}$$

Note that each a_k is an integer; and $a_{p-1} = n!$ and $a_{(n+1)p-1} = (-1)^n$.

Claim: If $l \geqslant p$ and $j \in \mathbb{Z}$, then $f^{(l)}(j)$ is an integer multiple of p. Indeed, since f is a polynomial, and $l \geqslant p$, $\frac{d^l}{dx^l}\left(\frac{(n!)^p}{(p-1)!}x^{p-1}\right) = 0$ and

$$\begin{aligned}\frac{d^l}{dx^l}\Big(\frac{a_k}{(p-1)!}x^k\Big) &= pa_k\frac{k(k-1)\ldots(k+1-p)\ldots(k+1-l)}{p!}\\ &= pa_k\binom{k}{p}(k-p)\ldots(k+1-l).\end{aligned}$$

This is a product of integers including p; so is a multiple of p.

Claim: If $0 \leqslant l \leqslant p-1$ and $0 \leqslant j \leqslant n$, then $f^{(l)}(j) = 0$ except for $f^{(p-1)}(0) = (n!)^p$; and $(n!)^p$ is an integer but is not a multiple of p. For each $1 \leqslant j \leqslant n$, the factor $(j-x)^p$ has a zero of order p, and so $f^{(l)}(x)$ has a zero of order $p-l \geqslant 1$; and hence $f^{(l)}(j) = 0$. At $j = 0$, a similar argument shows that $f^{(l)}(0) = 0$ for $0 \leqslant l \leqslant p-2$. Finally, $f^{(p-1)}(x) = (n!)^p +$ higher order terms, so $f^{(p-1)}(0) = (n!)^p$. Since $n < p$, this has no factor of p.

Now we define two more functions.

$$F(x) = f(x) + f'(x) + f^{(2)}(x) + \cdots + f^{(r)}(x) = \sum_{k=0}^{\infty} f^{(k)}(x)$$

and $G(x) = e^{-x}F(x)$. Then

$$\begin{aligned}G'(x) &= e^{-x}(F'(x) - F(x))\\ &= e^{-x}\Big(\sum_{k=0}^{\infty} f^{(k+1)}(x) - \sum_{k=0}^{\infty} f^{(k)}(x)\Big) = -e^{-x}f(x).\end{aligned}$$

Apply the MVT on $[0,k]$ for $1 \leqslant k \leqslant n$ and find points $x_k \in (0,k)$ so that

$$\frac{G(k) - G(0)}{k} = G'(x_k) = -e^{-x_k}f(x_k).$$

Multiply by $c_k k e^k$ to get

$$c_k F(k) - c_k e^k F(0) = -kc_k e^{k-x_k} f(x_k)$$

and sum from 1 to n

$$\sum_{k=1}^{n} c_k F(k) - \big(\sum_{k=1}^{n} c_k e^k\big)F(0) = -\sum_{k=1}^{n} kc_k e^{k-x_k} f(x_k).$$

Now we observe that $\sum_{k=1}^{n} c_k e^k = -c_0$ by hypothesis, and by the two claims above, the LHS is a sum of integers and all but one is a multiple of p,

$$\sum_{k=0}^{n} c_k F(k) \equiv c_0(n!)^p \not\equiv 0 \pmod{p}.$$

In particular, this is a non-zero integer—and hence is at least 1 in absolute value. Now we make some estimates to show that the RHS is too small for p sufficiently large, reaching a contradiction.

Using the formula for $f(x)$, we get the crude estimate $\max_{0 \leqslant x \leqslant n} |f(x)| \leqslant \dfrac{(n^{n+1})^p}{(p-1)!}$. Therefore

$$|\text{RHS}| \leqslant \Big(\sum_{k=1}^{n} k|c_k|e^k \Big) \frac{(n^{n+1})^p}{(p-1)!}.$$

The bound has the form $\dfrac{AB^p}{(p-1)!}$ for constants A and B. However

$$\lim_{p\to\infty} \frac{AB^p}{(p-1)!} = 0.$$

Thus we can choose a prime p large enough that $|\text{RHS}| < 1$. Therefore e is not algebraic, and so is transcendental. ■

A.5. π is irrational

Here is another off-the-wall irrationality proof. It implies that π is irrational. The argument will need to use a trig function in order to encapsulate the value π somehow. It is much in the same spirit as the previous proof.

A.5.1. THEOREM. π^2 *is irrational.*

PROOF. Suppose that $\pi^2 = \frac{a}{b}$ where $a, b \in \mathbb{N}$. Since $\lim_{n\to\infty} \dfrac{\pi a^n}{n!} = 0$, we choose an n so that $0 < \dfrac{\pi a^n}{n!} < 1$. Let $f(x) = \dfrac{x^n(1-x)^n}{n!}$. Then $0 < f(x) < \frac{1}{n!}$ if $0 < x < 1$. Also

$$f(x) = \frac{1}{n!}x^n \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^k = \sum_{k=n}^{2n} \frac{c_k}{n!} x^k$$

where $c_k \in \mathbb{Z}$ are integers and $c_n = 1$. Compute

$$\begin{aligned} f^{(k)}(0) &= 0 && \text{if } 0 \leqslant k \leqslant n-1 \text{ and } k > 2n \\ f^{(n)}(0) &= n!\frac{c_n}{n!} = 1 && \text{if } k = n \\ f^{(k)}(0) &= k!\frac{c_k}{n!} = k(k-1)\cdots(n+1)c_k && \text{if } n+1 \leqslant k \leqslant 2n. \end{aligned}$$

In particular, $f^{(k)}(0)$ is always an integer. Notice that $f(1-x) = f(x)$, and hence $f^{(k)}(1-x) = (-1)^k f^{(k)}(x)$, Therefore, $f^{(k)}(1)$ is also always an integer.

Define another polynomial

$$F(x) = b^n\big(\pi^{2n} f(x) - \pi^{2n-2} f^{(2)}(x) + \cdots + (-1)^n f^{(2n)}(x)\big)$$
$$= b^n \sum_{k=0}^{n} (-1)^k \pi^{2n-2k} f^{(2k)}(x).$$

Since $\pi^2 = \frac{a}{b}$ by assumption, $b^n\pi^{2n-2k}$ are integers. Also by the previous paragraph, $F(0)$ and $F(1)$ are integers. Compute

$$\pi^2 F(x)+F''(x)=b^n \sum_{k=0}^{n}(-1)^k\pi^{2n+2-2k} f^{(2k)}(x)+b^n \sum_{k=0}^{n}(-1)^k\pi^{2n-2k} f^{(2k+2)}(x)$$
$$= b^n \sum_{k=0}^{n}(-1)^k\pi^{2n+2-2k} f^{(2k)}(x)-b^n \sum_{l=1}^{n+1}(-1)^l\pi^{2n+2-2l} f^{(2l)}(x)$$
$$= b^n\pi^{2n+2} f(x) = a^n\pi^2 f(x).$$

Now we introduce some trig functions. Define

$$G(x) = F'(x)\sin\pi x - \pi F(x)\cos\pi x.$$

Then

$$G'(x) = F''(x)\sin\pi x + \pi F'(x)\cos\pi x - \pi F'(x)\cos\pi x + \pi^2 F(x)\sin\pi x$$
$$= (F''(x) + \pi^2 F(x))\sin\pi x = a^n\pi^2 f(x)\sin\pi x.$$

Now $G(0) = -\pi F(0)$ and $G(1) = \pi F(1)$. Therefore

$$\frac{G(1) - G(0)}{\pi} = -F(0) - F(1) \in \mathbb{Z}.$$

The MVT provides an $x_0 \in (0,1)$ so that $G(1) - G(0) = G'(x_0)$. Hence

$$|-F(0) - F(1)| = \frac{|G(1) - G(0)|}{\pi} = \frac{|G'(x_0)|}{\pi} = a^n\pi|f(x_0)\sin\pi x_0| > 0$$

and $a^n\pi|f(x_0)\sin\pi x_0| < \frac{a^n\pi}{n!} < 1$. This is an integer in $(0,1)$, which is a contradiction. Therefore π^2 is irrational. ∎

A.6. Stirling's Formula

In this section, we derive an asymptotic formula for $n!$ known as Stirling's formula. It is based on approximating the integral of $f(x) = \ln x$, $A_n = \int_1^n \ln x\, dx$, for $n \geqslant 2$. Since the second derivative $f''(x) = -\frac{1}{x^2} < 0$, the curve $f(x)$ is curving downwards or concave. This means that the (red) line segments from $(k-1, \ln k)$ and $(k, \ln k)$ lies below the curve for $2 \leqslant k \leqslant n$. Hence the sum of the areas of these trapezoids provides a lower bound for the integral. On the other hand, the tangent line through $(k-\frac{1}{2}, \ln(k-\frac{1}{2}))$ lies above the curve. We use them to bound the area from above.

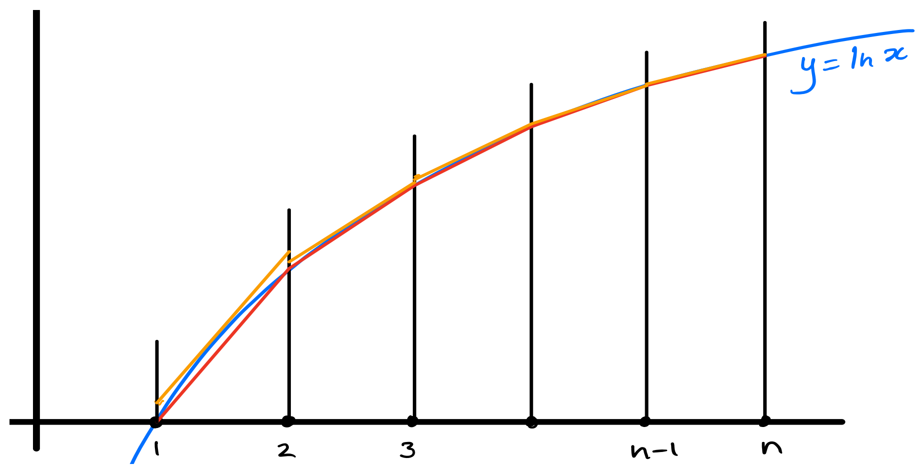


FIGURE 12.3. Approximate $\int_1^n \ln x\,dx$

The lower bound is

$$A_n > \sum_{k=2}^{n} \frac{\ln k - 1 + \ln k}{2} = \sum_{k=2}^{n-1} \ln k + \frac{1}{2}\ln n = \ln n! - \frac{1}{2}\ln n =: B_n$$

Now

$$A_n = \int_1^n \ln x\,dx = x\ln x - x\Big|_1^n = n\ln n - (n-1).$$

Define the error to be

$$\begin{aligned} E_n &= A_n - B_n = n\ln n - (n-1) - \ln n! + \tfrac{1}{2}\ln n \\ &= (n+\tfrac{1}{2})\ln n - (n-1) - \ln n!. \end{aligned}$$

Therefore

$$\ln n! = (n+\tfrac{1}{2})\ln n - n + (1-E_n).$$

Exponentiating yields

$$n! = e^{1-E_n} n^n \sqrt{n} e^{-n}.$$

Now we consider the upper bound. The tangent line through $(k-\frac{1}{2}, \ln(k-\frac{1}{2}))$ from $k-1$ to k has average height $\ln(k-\frac{1}{2})$. Hence the upper bound is

$$A_n < \sum_{k=2}^{n} \ln(k-\tfrac{1}{2}) =: C_n.$$

Thus using $\ln(1+x) < x$, as $f(x)$ lies below the tangent line through $(1,0)$,

$$\begin{aligned}
E_n < C_n - B_n &= \frac{1}{2}\sum_{k=2}^{n} 2\ln(k-\tfrac{1}{2}) - (\ln k - 1 + \ln k) \\
&= \frac{1}{2}\sum_{k=2}^{n} \ln\frac{k^2-k+\frac{1}{4}}{k^2-k} = \frac{1}{2}\sum_{k=2}^{n}\ln\Big(1+\frac{1}{4(k-1)k}\Big) \\
&< \frac{1}{2}\sum_{k=2}^{n}\frac{1}{4(k-1)k} = \frac{1}{8}\sum_{k=2}^{n}\frac{1}{k-1}-\frac{1}{k} \\
&= \frac{1}{8}\big(1-\frac{1}{n}\big) < \frac{1}{8}.
\end{aligned}$$

This yields the *Stirling inequality*

$$e^{7/8}\Big(\frac{n}{e}\Big)^n\sqrt{n} < n! < e\Big(\frac{n}{e}\Big)^n\sqrt{n}.$$

However we can do better by computing $E := \lim\limits_{n\to\infty} E_n$. Note that E_n is monotone increasing and bounded above, so that the limit exists by the Monotone Convergence Theorem. We know that

$$e^{1-E} = \lim_{n\to\infty}\frac{n!e^n}{n^n\sqrt{n}}.$$

Set $b_n = \dfrac{n!e^n}{n^n\sqrt{n}}$. Compute

$$\begin{aligned}
\frac{b_n^2}{b_{2n}} &= \frac{(n!)^2e^{2n}}{n^{2n+1}}\,\frac{(2n)^{2n+\frac{1}{2}}}{(2n)!e^{2n}} = \sqrt{\frac{2}{n}}\,\frac{(2^n n!)^2}{(2n)!} \\
&= \sqrt{\frac{2}{n}}\,\frac{2^2\cdot 4^2\cdot 6^2\cdots(2n)^2}{1\cdot 2\cdot 3\cdot 4\cdots(2n-1)(2n)} \\
&= \sqrt{\frac{2}{n}}\,\frac{2\cdot 4\cdot 6\cdots(2n)}{1\cdot 3\cdot 5\cdots(2n-1)} \\
&= \sqrt{\frac{2(2n+1)}{n}}\sqrt{\frac{2^2\cdot 4^2\cdot 6^2\cdots(2n)^2}{1\cdot 3^2\cdot 5^2\cdots(2n-1)^2(2n+1)}}
\end{aligned}$$

By the Wallis product formula, we obtain

$$e^{1-E} = \lim_{n\to\infty}\frac{b_n^2}{b_{2n}} = 2\sqrt{\frac{\pi}{2}} = \sqrt{2\pi}.$$

Therefore we have established:

A.6.1. STIRLING'S FORMULA. $\displaystyle\lim_{n\to\infty}\frac{n!}{n^n e^{-n}\sqrt{2\pi n}} = 1.$

A.7. $\int_0^\infty \frac{\sin x}{x}\,dx$

There are many different ways to compute the integral in the title. None of them are easy, and many would not be considered elementary from the point of view of first year calculus. This one draws inspiration from the theory of Fourier seres, but we hide this by proving the necessary results without their real motivation. That makes the proof seem artificial. In fact, every method for this integral has some sort of trick. The first lemma is an easy result about the *Dirichlet kernel*.

A.7.1. LEMMA. $\displaystyle\int_0^\pi \frac{\sin(n+\frac12)x}{\sin\frac{x}{2}}\,dx = \pi.$

PROOF. Let $f(x) = \begin{cases} 2n+1 & \text{if } x=0 \\ \dfrac{\sin(n+\frac12)x}{\sin\frac{x}{2}} & \text{if } 0 < x \leqslant \pi \end{cases}$. This function is continuous because

$$\lim_{x\to 0} \frac{\sin(n+\frac12)x}{\sin\frac{x}{2}} = 2n+1.$$

We use the identity $2\sin x\cos y = \sin(x+y) - \sin(x-y)$ in the following calculation.

$$\begin{aligned} \sin\frac{x}{2}\Big(1 + 2\sum_{k=1}^n \cos ku\Big) &= \sin\frac{x}{2} + \sum_{k=1}^n 2\sin\frac{x}{2}\cos ku \\ &= \sin\frac{x}{2} + \sum_{k=1}^n \sin(k+\tfrac12)x - \sin(k-\tfrac12)x \\ &= \sin(n+\tfrac12)x. \end{aligned}$$

Therefore $f(x) = 1 + 2\sum_{k=1}^n \cos ku$ for $0 < x \leqslant \pi$ and also at $x = 0$.

Now we can integrate

$$\int_0^\pi \frac{\sin(n+\frac12)x}{\sin\frac{x}{2}}\,dx = \int_0^\pi 1 + 2\sum_{k=1}^n \cos ku\,dx = \pi. \qquad \blacksquare$$

A.7.2. LEMMA. *Define* $g(x) = \dfrac{1}{x} - \dfrac{1}{2\sin\frac{x}{2}}$ *for* $0 < x \leqslant \pi$. *Then* $g(x)$ *and* $g'(x)$ *extend to be continuous at* $x = 0$.

PROOF. The argument using the Taylor expansion $\sin x = x - \frac{1}{6}x^3 + O(x^5)$ near $x = 0$ is straightforward.

$$\begin{aligned}\lim_{x\to 0} g(x) &= \lim_{x\to 0} \frac{1}{x} - \frac{1}{2\sin\frac{x}{2}} = \lim_{x\to 0} \frac{2\sin\frac{x}{2} - x}{2x\sin\frac{x}{2}} \\ &= \lim_{x\to 0} \frac{2(\frac{1}{2}x - \frac{1}{48}x^3 + O(x^5)) - x}{2x(\frac{1}{2}x + O(x^3)} \\ &= \lim_{x\to 0} \frac{-\frac{1}{24}x^3 + O(x^5)}{x^2 + O(x^4)} = \lim_{x\to 0} \frac{-\frac{1}{24}x + O(x^3)}{1 + O(x^2)} = 0.\end{aligned}$$

So g extends to be continuous at $x = 0$. We have

$$g'(x) = -\frac{1}{x^2} + \frac{\cos\frac{x}{2}}{4\sin^2\frac{x}{2}} = \frac{x^2\cos\frac{x}{2} - 4\sin^2\frac{x}{2}}{4x^2\sin^2\frac{x}{2}}.$$

Therefore, since $\cos x = 1 - \frac{1}{2}x^2 + O(x^4)$ near $x = 0$,

$$\begin{aligned}\lim_{x\to 0} g'(x) &= \lim_{x\to 0} \frac{x^2\cos\frac{x}{2} - 4\sin^2\frac{x}{2}}{4x^2\sin^2\frac{x}{2}} \\ &= \lim_{x\to 0} \frac{x^2(1 - \frac{1}{8}x^2 + O(x^4)) - 4(\frac{1}{2}x - \frac{1}{48}x^3 + O(x^5))^2}{4x^2(\frac{1}{2}x - \frac{1}{48}x^3 + O(x^5))^2} \\ &= \lim_{x\to 0} \frac{x^2 - \frac{1}{8}x^4 + O(x^6) - 4(\frac{1}{4}x^2 - \frac{1}{48}x^4 + O(x^6))}{4x^2(\frac{1}{4}x^2 - \frac{1}{48}x^4 + O(x^6))} \\ &= \lim_{x\to 0} \frac{-\frac{1}{24}x^4 + O(x^6)}{x^4 - \frac{1}{48}x^6 + O(x^8))} = \lim_{x\to 0} \frac{-\frac{1}{24} + O(x^2)}{1 + O(x^2)} = -\frac{1}{24}.\end{aligned}$$

Thus g' extends to be continuous at $x = 0$ by setting $g'(0) = -\frac{1}{24}$. ■

The following is a special case of the *Riemann-Lebesgue Lemma*.

A.7.3. LEMMA. *If $g(x)$ is C^1 on $[0, \pi]$, then* $\displaystyle\lim_{t\to\infty} \int_0^\pi g(x)\sin(tx)\,dx = 0$.

PROOF. Integrating by parts yields

$$\begin{aligned}\Big|\int_0^\pi g(x)\sin(tx)\,dx\Big| &= \Big| - g(x)\frac{1}{t}\cos(tx) + \frac{1}{t}\int_0^\pi g'(x)\cos(tx)\,dx\Big| \\ &\leqslant \frac{1}{t}(2\|g\|_\infty + \|g'\|_\infty).\end{aligned}$$

Clearly this goes to 0 as $t \to \infty$. ■

A.7.4. REMARK. It is easy to show that every continuous function can be approximated by a C^1 function uniformly with ε. So an easy argument upgrades this lemma to continuous functions. That would mean that we did not have to show that $g(x)$ was differentiable in the previous lemma. Also one can easily replace the interval $[0,\pi]$ with $[a,b]$ without change.

A.7.5. Theorem. $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$

Proof. In Example 8.1.7, we showed that the improper integral

$$\int_0^\infty \frac{\sin x}{x}\,dx = \lim_{b\to\infty}\int_0^b \frac{\sin x}{x}\,dx$$

exists. Using $g(x) = \dfrac{1}{x} - \dfrac{1}{2\sin\frac{x}{2}}$, we see that

$$\begin{aligned}\int_0^\pi g(x)\sin(n+\tfrac12)x\,dx &= \int_0^\pi \frac{\sin(n+\frac12)x}{x}\,dx - \int_0^\pi \frac{\sin(n+\frac12)x}{2\sin\frac{x}{2}}\,dx \\ &= \int_0^{(n+\frac12)\pi} \frac{\sin x}{x}\,dx - \frac{\pi}{2}.\end{aligned}$$

We used a change of variables for the first integral and Lemma A.7.1 for the second. By Lemma A.7.3, the LHS tends to 0. Since the improper integral exists,

$$\begin{aligned}\int_0^\infty \frac{\sin x}{x}\,dx &= \lim_{n\to\infty}\int_0^{(n+\frac12)\pi} \frac{\sin x}{x}\,dx \\ &= \frac{\pi}{2} + \lim_{n\to\infty}\int_0^\pi g(x)\sin(n+\tfrac12)x\,dx = \frac{\pi}{2}.\end{aligned}$$

■

A.8. Isoperimetric inequality

A famous problem from the time of the ancient Greeks is to determine the closed curve of given perimeter which encloses the largest area. The Greeks believed that the answer was a circle, but could not prove it. There are now a variety of proofs and generalizations to higher dimensions. Here we give a proof based on calculus.

First notice that any closed curve that has a chance to encircle the largest area must be convex. Thus if we arrange the figure so that the x-coordinates run from a to b and back, in the counterclockwise direction, then the first part of the curve must be a convex function, and the second part is a concave function. We showed in Part I, Theorem 5.6.5 and Corollary 5.6.6, that a convex function has a left derivative everywhere and it is monotone increasing. Similarly the concave part has a monotone decreasing left derivative. This isn't quite enough to be piecewise C^1, however monotone functions are Riemann integrable. So with a bit of care, Green's

Theorem 8.5.3 is still valid. Indeed, convex curves fall under the special case for the first part of that proof. So our proof of the isoperimetric inequality is actually valid in general.

A.8.1. Isoperimetric Inequality. *Let γ be a piecewise C^1 closed curve of arc length 2π. Then the area enclosed by γ is at most π, and this occurs only when γ is a circle.*

Proof. We parameterize the curve by arc length s. Then $\gamma(s) = (x(s), y(s))$ for $0 \leqslant s \leqslant 2\pi$ and

$$\left|\frac{d\gamma}{ds}\right|^2 = x'(s)^2 + y'(s)^2 = 1.$$

We can reposition the curve by translation and rotation (which does not affect arc length or area) so that $\gamma(0)$ and $\gamma(\pi)$ lie on the x-axis with $x(\pi) < x(0)$. So $y(0) = y(\pi) = 0$. The discussion above shows that $\gamma([0, \pi])$ lies above the axis, and $\gamma([\pi, 2\pi])$ lies below. Note that if $0 \leqslant s \leqslant \pi$, we have $x'(s) \leqslant 0$ and $y(s) \geqslant 0$; while for $\pi \leqslant x \leqslant 2\pi$, we have $x'(s) \geqslant 0$ and $y(s) \leqslant 0$. So $-y(s)x'(s) \geqslant 0$ for all s. By Green's Theorem 8.5.3, the area enclosed by γ is

$$A = \int_0^{2\pi} -y(s)x'(s)\,ds = \int_0^{2\pi} |y(s)x'(s)|\,ds.$$

Recall that $ab \leqslant \frac{1}{2}(a^2 + b^2)$, a special case of the AMGM inequality. Thus

$$A = \int_0^{2\pi} |y(s)x'(s)|\,ds \leqslant \int_0^{2\pi} \tfrac{1}{2}(y(s)^2 + x'(s)^2)\,ds = \frac{1}{2}\int_0^{2\pi} y(s)^2 + 1 - y'(s)^2\,ds.$$

Define $u(s) = \dfrac{y(s)}{\sin s}$. Observe that

$$\lim_{s\to 0} \frac{y(s)}{\sin s} = \lim_{s\to 0} \frac{y(s)}{s}\,\frac{s}{\sin s} = y'(0)$$

and

$$\lim_{s\to \pi} \frac{y(s)}{\sin s} = \lim_{h\to 0} \frac{y(\pi + h)}{h}\,\frac{h}{\sin \pi + h} = -y'(\pi).$$

Thus u is continuous on $[0, 2\pi]$, and differentiable except possibly at $0, \pi$ and 2π. Then $y(s) = u(s)\sin s$ and $y'(s) = u(s)\cos s + u'(s)\sin s$. Therefore

$$\begin{aligned} A &\leqslant \frac{1}{2}\int_0^{2\pi} u(s)^2 \sin^2 s + 1 - \big(u(s)\cos s + u'(s)\sin s\big)^2\,ds \\ &= \frac{1}{2}\int_0^{2\pi} \Big(u(s)^2(\sin^2 s - \cos^2 s) - 2u(s)u'(s)\sin s \cos s\Big) + 1 - u'(s)^2 \sin^2 s\,ds \\ &= -\frac{1}{2}u(s)^2 \sin s \cos s\Big|_0^{2\pi} + \frac{1}{2}\int_0^{2\pi} 1 - u'(s)^2 \sin^2 s\,ds \\ &\leqslant 0 + \pi = \pi. \end{aligned}$$

Equality holds in the last inequality only if $u'(s) = 0$ everywhere, so that u is constant, which yields $y(s) = c\sin s$. The AMGM inequality is an equality only if $|x'(s)| = |y(s)|$, so that $x'(s) = -c\sin s$ because $y(s)x'(s) \leqslant 0$. Hence $x(s) = c\cos s + d$. Our initial choices show that $c > 0$. Also

$$1 = x'(s)^2 + y'(s)^2 = c^2\sin^2 s + c^2\cos^2 s = c^2.$$

So $c = 1$. That is, $\gamma(s) = (d, 0) + (\cos s, \sin s)$ for $0 \leqslant s \leqslant 2\pi$. This is a circle of radius 1 and area π. ■

A.9. Euler's sum

We will establish the famous formula of Euler: $\sum\limits_{n\geqslant 1} \dfrac{1}{n^2} = \dfrac{\pi^2}{6}$ and some related formulae. The arguments use knowledge of the complex numbers including *de Moivre's Theorem*:

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta \quad \text{for} \quad \theta \in \mathbb{R} \text{ and } n \in \mathbb{N}.$$

A.9.1. Lemma. $\sum\limits_{k=1}^{n} \cot^2\left(\frac{k\pi}{2n+1}\right) = \dfrac{n(2n-1)}{3}.$

Proof. By de Moivre's Theorem,

$$\begin{aligned}\cos(2n+1)\theta + i\sin(2n+1)\theta &= (\cos\theta + i\sin\theta)^{2n+1} \\ &= \sin^{2n+1}\theta(\cot\theta + i)^{2n+1} \\ &= \sin^{2n+1}\theta \sum_{k=0}^{2n+1} \binom{2n+1}{k} i^k \cot^{2n+1-k}.\end{aligned}$$

Taking imaginary parts and dividing by $\sin^{2n+1}\theta$ yields

$$\begin{aligned}\frac{\sin(2n+1)\theta}{\sin^{2n+1}\theta} &= \binom{2n+1}{1}\cot^{2n}\theta - \binom{2n+1}{3}\cot^{2n-2}\theta + \dots + (-1)^n \\ &= \sum_{m=0}^{n} (-1)^m \binom{2n+1}{2m+1} \cot^{2n-2m}\theta = P_n(\cot^2\theta)\end{aligned}$$

where we define the polynomial

$$P_n(x) = \sum_{m=0}^{n} (-1)^m \tbinom{2n+1}{2m+1} x^{n-m} = \tbinom{2n+1}{1}x^n - \tbinom{2n+1}{3}x^{n-1} + \tbinom{2n+1}{5}x^{n-2} + \dots .$$

Then

$$P_n\big(\cot^2 \tfrac{k\pi}{2n+1}\big) = \frac{\sin k\pi}{\sin^{2n+1}\frac{k\pi}{2n+1}} = 0 \quad \text{for} \quad 1 \leqslant k \leqslant n.$$

Since deg $P_n = n$, this is a complete list of the roots of P_n. Because $P_n(x) = \binom{2n+1}{1}x^n - \binom{2n+1}{3}x^{n-1} + \dots$, the sum of the roots is

$$\sum_{k=1}^{m} \cot^2\left(\tfrac{k\pi}{2n+1}\right) = \frac{\binom{2n+1}{3}}{\binom{2n+1}{1}} = \frac{(2n+1)(2n)(2n-1)}{6(2n+1)} = \frac{n(2n-1)}{3}. \qquad \blacksquare$$

A.9.2. EULER'S THEOREM. $\displaystyle\sum_{n\geqslant 1} \frac{1}{n^2} = \frac{\pi^2}{6}$.

PROOF. For $0 < x < \frac{\pi}{2}$, we have $\sin x < x < \tan x$. Therefore

$$\cot^2 x < \frac{1}{x^2} < \csc^2 x = 1 + \cot^2 x.$$

Therefore

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

Applying Lemma A.9.1, we obtain

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

Simplifying we get

$$\frac{\pi^2}{6} \frac{2n(2n-1)}{(2n+1)^2} < \sum_{k=1}^{n} \frac{1}{k^2} < \frac{\pi^2}{6} \frac{4n(n+2)}{(2n+1)^2}.$$

Now let $n \to \infty$ and apply the squeeze theorem to get $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$. $\blacksquare$

Now we will push this a bit harder to compute $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^4}$. We first need a formula for the sum of the squares of the roots of a polynomial.

A.9.3. LEMMA. *If* $p(x) = a_x x^n + a_{n-1}x^{n-1} + \ldots + a_0 = a_n \prod_{k=1}^{n}(x - r_k)$, *then*

$$\sum_{k=1}^{n} r_k^2 = \left(\frac{a_{n-1}}{a_n}\right)^2 - 2\left(\frac{a_{n-2}}{a_n}\right).$$

PROOF. Expanding the first few terms of the product, we get

$$\prod_{k=1}^{n}(x - r_k) = x^n - \left(\sum_{k=1}^{n} r_k\right)x^{n-1} + \left(\sum_{1\leqslant j<k\leqslant n} r_j r_k\right)x^{n-2} + \cdots + (-1)^n \prod_{k=1}^{n} r_k.$$

Hence

$$\sum_{k=1}^{n} r_k = \frac{-a_{n-1}}{a_n} \quad \text{and} \quad \sum_{1\leqslant j<k\leqslant n} r_j r_k = \frac{a_{n-2}}{a_n}.$$

Therefore

$$\sum_{k=1}^{n} r_k^2 = \Big(\sum_{k=1}^{n} r_k\Big)^2 - 2\sum_{1\leqslant j<k\leqslant n} r_j r_k = \Big(\frac{a_{n-1}}{a_n}\Big)^2 - 2\Big(\frac{a_{n-2}}{a_n}\Big).$$

■

A.9.4. LEMMA. $\sum_{k=1}^{n} \cot^4\left(\frac{k\pi}{2n+1}\right) = \dfrac{8}{45}n^4 + O(n^3).$

PROOF. We use the polynomial $P_n(x)$ from Lemma A.9.1 and apply the preceding Lemma.

$$\begin{aligned}
\sum_{k=1}^{n} \cot^4\left(\frac{k\pi}{2n+1}\right) &= \left(\frac{\binom{2n+1}{3}}{\binom{2n+1}{1}}\right)^2 - 2\frac{\binom{2n+1}{5}}{\binom{2n+1}{1}} \\
&= \frac{(n^2(2n-1)^2}{9} - \frac{2(2n+1)(2n)(2n-1)(2n-2)(2n-3)}{120(2n+1)} \\
&= \frac{(n(2n-1)}{45}(5(2n^2-n)-3(2n^2-5n+3)) \\
&= \frac{(n(2n-1)}{45}(4n^2+10n-9) = \frac{8}{45}n^4 + \text{ lower order terms.}
\end{aligned}$$

The lower order terms are bounded by a multiple of n^3 for large n, so the sum equals $\frac{8}{45}n^4 + O(n^3)$. ■

A.9.5. THEOREM. $\sum_{n\geqslant 1} \dfrac{1}{n^4} = \dfrac{\pi^4}{90}.$

PROOF. Again we use the inequality $\cot^2 x < \dfrac{1}{x^2} < 1 + \cot^2 x$. Thus

$$\cot^4 x < \frac{1}{x^4} < (1+\cot^2 x)^2.$$

Therefore

$$\begin{aligned}
\sum_{k=1}^{n} \cot^4\left(\tfrac{k\pi}{2n+1}\right) < \sum_{k=1}^{n} \left(\frac{2n+1}{k\pi}\right)^4 &< \sum_{k=1}^{n} \Big(1 + \cot^2\left(\tfrac{k\pi}{2n+1}\right)\Big)^2 \\
&= \sum_{k=1}^{n} \cot^4\left(\tfrac{k\pi}{2n+1}\right) + 2\sum_{k=1}^{n} \cot^2\left(\tfrac{k\pi}{2n+1}\right) + n.
\end{aligned}$$

By Lemma A.9.4 and Lemma A.9.1,

$$\frac{\pi^4}{(2n+1)^4}\Big(\frac{8}{45}n^4 + O(n^3)\Big) < \sum_{k=1}^{n}\frac{1}{k^4} < \frac{\pi^4}{(2n+1)^4}\Big(\frac{8}{45}n^4 + O(n^3)\Big).$$

Simplifying, we get

$$\sum_{k=1}^{n}\frac{1}{k^4} = \frac{\pi^4}{90} + O(\tfrac{1}{n}).$$

Letting n go to infinity, we obtain $\sum\limits_{k=1}^{\infty}\frac{1}{k^4} = \frac{\pi^4}{90}$. ■

There are other ways to get these identities. Once you learn about Fourier series, easier proofs will be available.

There is a general formula $\sum\limits_{k=1}^{\infty}\frac{1}{k^{2n}} = \frac{(2\pi)^{2n}|B_{2n}|}{2(2n)!}$ where the B_n are the *Bernoulli numbers*. They are given by $B_0 = 1$ and $B_n = -\sum\limits_{k=0}^{n-1}\binom{n}{k}\frac{B_k}{n+1-k}$. Thus $B_1 = -\frac{1}{2}$, $B_{2n+1} = 0$ for $n \geqslant 1$, $B_2 = \frac{1}{6}$, $B_4 = \frac{-1}{30}$, $B_6 = \frac{1}{42}$, $B_8 = \frac{-1}{30}, \dots$. Also

$$1 = \Big(B_0 + \frac{B_1}{1!}x + \frac{B_2}{2!}x^2 + \frac{B_3}{3!}x^3 + \dots\Big)\Big(1 + \frac{1}{2!}x + \frac{1}{3!}x^2 + \frac{1}{4!}x^3 + \dots\Big).$$

A.10. The Gamma function

The *gamma function* is a continuous (in fact, C^∞) function which behaves like the factorial function. It arises naturally in various places in classical analysis.

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt \quad \text{for} \quad x > 0.$$

The integand is positive, so this improper integral is defined as long as it is bounded above. We establish this in two steps. Fix $x > 0$.

$$\int_0^1 t^{x-1}e^{-t}\,dt \leqslant \int_0^1 t^{x-1}\,dt = \frac{t^x}{x}\Big|_{t=0}^{t=1} = \frac{1}{x} < \infty.$$

Since $\lim\limits_{t\to\infty}\frac{t^{x-1}}{e^{t/2}} = 0$, $M_x = \sup_{t\geqslant 1}\frac{t^{x-1}}{e^{t/2}} < \infty$ by the Extreme Value Theorem. Therefore

$$\int_1^\infty t^{x-1}e^{-t}\,dt \leqslant \int_1^\infty M_x e^{-t/2}\,dt = -2M_x e^{-t/2}\Big|_{t=1}^{t=\infty} = \frac{2M_x}{\sqrt{e}} < \infty.$$

Therefore $\Gamma(x)$ is defined for all $x > 0$.

CLAIM: $\Gamma(x+1) = x\Gamma(x)$. Let $\varepsilon > 0$ and $R > 0$ large. Integrate by parts:

$$\int_\varepsilon^R t^x e^{-t}\,dt = -t^x e^{-t}\Big|_\varepsilon^R + \int_\varepsilon^R x t^{x-1} e^{-t}\,dt$$
$$= \varepsilon e^{-\varepsilon} - R^x e^{-R} + x\int_\varepsilon^R t^{x-1}e^{-t}\,dt.$$

Now let $\varepsilon \to 0^+$ and $R \to +\infty$ and we get

$$\Gamma(x+1) = x\Gamma(x) + \lim_{\varepsilon\to 0^+} \varepsilon e^{-\varepsilon} - \lim_{R\to\infty} R^x e^{-R} = x\Gamma(x).$$

CLAIM: $\Gamma(n+1) = n!$ for $n \in \mathbb{N}_0$.

$$\Gamma(1) = \int_0^\infty e^{-t}\,dt = -e^{-t}\Big|_0^\infty = 1.$$

The claim is now established by induction. It is true for $n = 0$. If it is true for $n-1$, then $\Gamma(n) = (n-1)!$. Thus $\Gamma(n+1) = n\Gamma(n) = n!$. This establishes the induction step.

CLAIM: $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Substitute $t = u^2$ and $dt = 2u\,du$ to get

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt = \int_0^\infty u^{2x-2}e^{-u^2}2u\,du = 2\int_0^\infty u^{2x-1}e^{-u^2}\,du.$$

In particular, $\Gamma(\frac{1}{2}) = 2\displaystyle\int_0^\infty e^{-u^2}\,du$. The trick for this integral is to square it and consider this as an integral over the first quadrant Q. Then use polar coordinates and the fact (compare with area in polar coordinates) that $du\,dv = r\,dr d\theta$.

$$\Gamma(\tfrac{1}{2})^2 = 4\int_0^\infty e^{-u^2}\,du \int_0^\infty e^{-v^2}\,dv = 4\iint_Q e^{-u^2-v^2}\,du\,dv$$
$$= 4\int_0^{\pi/2}\int_0^\infty e^{-r^2} r\,dr\,d\theta = 4\frac{\pi}{2}\big(-\tfrac{1}{2}e^{-r^2}\big)\Big|_0^\infty = \pi.$$

By symmetry, we also have $\displaystyle\int_{-\infty}^\infty e^{-u^2}\,du = 2\int_0^\infty e^{-u^2}\,du = \sqrt{\pi}$.

VOLUMES OF SPHERES. Let the unit ball in $\mathbb{R}^n$ be B_n. The surface of the ball B_n is called S_{n-1}, or the $n-1$-sphere. The notation reflects the fact that S_{n-1} is an $n-1$-dimensional object. Observe that

$$\pi^{n/2} = \Big(\int_{-\infty}^\infty e^{-u^2}\,du\Big)^n = \int\cdots\int_{\mathbb{R}^n} e^{-(u_1^2+u_2^2+\cdots+u_n^2)}\,du_1\,du_2\,\ldots du_n.$$

We employ the spherical shell method to compute this integral. Observe that on the sphere rS_{n-1} of radius r, the function $e^{-(u_1^2+u_2^2+\cdots+u_n^2)}$ takes the constant value e^{-r^2}. Thus integrating over rS_{n-1} will yield e^{-r^2} times the $n-1$-dimensional volume of rS_{n-1}. This must be r^{n-1} times the $n-1$-dimensional volume $|S_{n-1}|$ of S_{n-1}. Now integrate with respect to r from 0 to ∞ to obtain the integral over $\mathbb{R}^n$.

$$\pi^{n/2} = \int_0^\infty r^{n-1} e^{-r^2} |S_{n-1}|\, dr = \tfrac{1}{2}\Gamma(\tfrac{n}{2})|S_{n-1}|.$$

Therefore the sphere S_{n-1} has $n-1$-dimensional volume $\dfrac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}$.

So the circumference of a circle of radius r is $r\dfrac{2\pi^{2/2}}{\Gamma(\frac{2}{2})} = 2\pi r$. The sphere of radius r has area

$$r^2 \frac{2\pi^{3/2}}{\Gamma(\frac{n}{3}2)} = \frac{2\pi^{3/2}}{\frac{1}{2}\Gamma(\frac{1}{2})} r^2 = 4\pi r^2.$$

The 3-sphere S_3 has volume and the 4-sphere has 4-dimensional volume

$$|S_3| = \frac{2\pi^{4/2}}{\Gamma(\frac{4}{2})} = \frac{2\pi^2}{1!} = 2\pi^2 \quad \text{and} \quad |S_4| = \frac{2\pi^{5/2}}{\Gamma(\frac{5}{2})} = \frac{2\pi^{5/2}}{\frac{3}{2}\frac{1}{2}\Gamma(\frac{1}{2})} = \frac{8}{3}\pi^2.$$

We compute the n-volume of $B_n(R)$ of radius R by the spherical shell technique:

$$|B_n(R)| = \int_0^R r^{n-1} |S_{n-1}|\, dr = \frac{2\pi^{n/2}}{n\Gamma(\frac{n}{2})} R^n.$$

INDEX