Calculus 2: Advanced

MAST10021

Volker Schlue
University of Melbourne

Semester 2, 2023

(version: November 21, 2023)

Contents

Preface

I have written these lecture notes for the subject **Calculus 2: Advanced** which was introduced at the University of Melbourne in 2020. Its ambition is to teach all the topics covered in the “standard” Calculus 2 subject, but with more emphasis on concepts, and proofs. Even though the name would suggest it, there is no **Calculus 1: Advanced** subject. For most students, this is the first enounter with any proofs in Calculus. Thus to get from the basics (say the definition of a limit) to the more advanced topics (say discussions of solutions to differential equations), the pace is quite high, and a selection of topics had to be made: I am mostly following (Spivak, *Calculus*), but I have tried to make a division into “core material” (these are the *Notes*, that I expect every student to study), and “additional topics” (that I usually do not have time to cover in class, but give an opportunity for further study, along with some references to the literature). The result is a class that can be taught over 12 weeks, accompanied by tutorials (which are not contained in these lecture notes). I would like to thank the students who took this subject in the years 2020-2023 for their feedback and I hope that taking this class made them curious about mathematics. – VS, Melbourne, November 2023.

Literature. The following text books were useful when preparing these lecture notes.

(Spivak, *Calculus*) This is a beautiful introduction to Calculus. About 2/3 of topics covered in this subject can be found in this text, especially foundations.

(Folland, *Advanced Calculus*) This is a good text, but with a more advanced starting point. About 1/4 of the topics covered in this course can be found this text, especially related to multivariable calculus.

(Osserman, *Two-Dimensional Calculus*) I have found this text to be a delightful treatment of Calculus in two and three variables. Unfortunately, we barely have time to cover these topics.

(Apostel, *Calculus I*) This is a quiet comprehensive treatment of Calculus, which in the end I have not found very useful. We consult it for topics related to differential equations.

(Walter, *Gewöhnliche Differentialgleichungen*) This is a very nice introduction to differential equations, but it requires some familiarity with the German language.

Module I.

Numbers, Functions, and Graphs

Note 1.

Numbers

1.1. Natural numbers and induction

The simplest numbers are the **natural numbers** $1, 2, 3, \ldots$, which are collectively referred to as $\mathbb{N}$.

The natural numbers are very important in particular in relation to the principle of **mathematical induction**: Suppose $P(x)$ means that the property P holds for the number x. Then the principle of mathematical induction states that $P(x)$ is true for all natural numbers x provided that

- $P(1)$ is true,
- If $P(k)$ is true, then $P(k+1)$ is true.

Example 1.1. Suppose we want to prove that for any natural number n,

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}\,. \tag{1.1}$$

Then it sufficies to demonstrate that this is valid in the case $n = 1$, and to show that its validity for $n = k$, implies that the formula holds for $n = k + 1$. Now assuming that the formula holds for $n = k$, we have that

$$1 + \ldots + k + (k+1) = \frac{k(k+1)}{2} + (k+1) = \frac{k^2 + 3k + 2}{2} = \frac{(k+2)(k+1)}{2} \tag{1.2}$$

which shows that the formula holds for $n = k + 1$.

Closely related to proofs by induction are *recursive defininitions*.

Example 1.2. The number $n!$ ("n factorial") is defined as follows:

- $1! = 1$
- $(n+1)! = (n+1)n!$

Example 1.3. We write $\sum_{i=1}^{n} a_i$ for the sum $a_1 + \ldots + a_n$. For example, the formula above could be expressed as

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}\,. \tag{1.3}$$

A careful definition of this symbol would be a recursive one:

- $\sum_{i=1}^{1} a_i = a_1$
- $\sum_{i=1}^{n+1} a_i = \sum_{i=1}^{n} a_i + a_{n+1}$.

1.2. Rational and real numbers

We could extend the system of natural numbers $\mathbb{N}$ to the **integers**, which is the collection of numbers $\mathbb{Z}$, given by

$$\ldots, -2, -1, 0, 1, 2, 3, \ldots .$$

A still larger system of numbers are the **rational numbers** $\mathbb{Q}$ obtained by taking quotients m/n of integers with $n \neq 0$. These numbers satisfy all the "basic properties" of numbers that we are familiar with (and which we will spell out in the additional material below). It is tempting to think that these are all the numbers we would ever need.

However, it is really *Calculus* that will lead us to yet a larger collection of numbers, the **real numbers** $\mathbb{R}$. The additional properties of these numbers are quite profound, and very different from basic arithmetic properties. We will return to these in Lecture 5.

Instead of discussing these additional properties now it might be instructive to convince ourselves that there *are* irrational numbers (real numbers which are not rational).

Example 1.4. The number $\sqrt{2}$ is *not* rational. The proof relies on the following observation: The square of an **even** natural number, namely a number of the form $n = 2k$, is again even; $(2k)^2 = 2(2k^2)$. The same is true for **odd** natural numbers, namely numbers of the form $n = 2k + 1$:

$$(2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

This means that if n^2 is even, then n is even; and if n^2 is odd, then n is odd. Now suppose $\sqrt{2}$ were rational, and equals p/q for some integers p and $q \neq 0$, and we can assume that p and q have no common divisor. Then

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

which shows that p^2, and hence p is even. So for $p = 2k$, we get

$$2k^2 = q^2$$

which now shows that also q is even. But if both p and q are even, this contradicts the fact that p and q have no common divisor.

The argument in this example shows that there is no rational number x such that $x^2 = 2$. We have *not* shown that there exists a number $\sqrt{2}$ whose square is 2.

1.3. Inequalities

Although inequalities are often not discussed in elementary mathematics, they play a prominent role in Calculus.

We write $a < b$ to say "a is less than b", and $a > b$ to say "a is greater than b". $a < b$ means the same as $b > a$ and are merely two ways of writing the same assertion. The main properties of numbers pertaining to inequalities are:

(Trichotomy) For any real numbers a and b, one, and only one, of the following holds:

1. $a = b$
2. $a < b$
3. $b < a$

(Ordering) For any numbers a, b, and c, if $a < b$ and $b < c$, then $a < c$.

(Closure under addition) For any numbers a, b, and c, if $a < b$, then $a + c < b + c$.

(Closure under multiplication) For any numbers a, b, and c, if $a < b$, and $0 < c$, then $ac < bc$.

The numbers satisfying $a > 0$ are called **positive**, while the numbers $a < 0$ are **negative**. We also write $a \leq b$ to mean $a < b$ or $a = b$, and $a \geq b$ to mean $a > b$ or $a = b$.

All familiar facts about inequalities, howevery elementary they may seem, can be derived from these basic properties.

Example 1.5. If $a < 0$ is negative, then $-a > 0$ is positive. Since $a < 0$ means the same as $0 > a$ it follows (from closure under addition) that $0 - a > a - a = 0$, or simply $-a > 0$.

Example 1.6. More generally, if $a < 0$ and $b < 0$ are negative numbers, then $ab > 0$ is positive: We have already shown that $-a > 0$ and $-b > 0$ are positive numbers, hence (using closure under multiplication) it follows that $(-a)(-b) > 0$.

Exercise 1.1. If $a > 0$ and $b > 0$ are positive, then $ab > 0$ is positive.

The fact that $ab > 0$ if $a > 0$, $b > 0$ and also if $a < 0$, $b < 0$ has one special consequence: $a^2 > 0$ whenever $a \neq 0$. In particular, we have proven that $1 = 1 \cdot 1 > 0$.

We will encounter many more examples of inequalities in Problem 7 below.

Example 1.7. If $a < b$ and $c < 0$ then $ac > bc$. Since $-c > 0$ we have $a(-c) = -ac < b(-c) = -bc$ which is the same as $bc < ac$.

1.4. Absolute value

For any number a, we define the **absolute value** of a as:

$$|a| = \begin{cases} a & a \geq 0 \\ -a & a \leq 0\,. \end{cases}$$

An important inequality is the **triangle inequality**, which is the statement that for all numbers a and b, we have

$$|a + b| \leq |a| + |b|\,. \tag{1.4}$$

Proof. Since the absolute value is defined by cases, the proof amounts to verifying this inequality in the following four cases:

$$
\begin{aligned}
&a \geq 0 \qquad b \geq 0 &&\text{(A)}\\
&a \geq 0 \qquad b \leq 0 &&\text{(B)}\\
&a \leq 0 \qquad b \geq 0 &&\text{(C)}\\
&a \leq 0 \qquad b \leq 0 &&\text{(D)}
\end{aligned}
$$

In the case (A), the stated inequality is certainly true, because $|a+b| = a+b = |a|+|b|$. Similarly in the case (D), we have $|a + b| = -(a + b) = -a - b = |a| + |b|$.

Now let us look at the case (B). Here we need to prove that

$$|a + b| \leq a - b\,. \tag{1.5}$$

From the assumption we do not know if $a + b$ is positive, negative, or zero, but we can consider the case $a + b \geq 0$ first. Then we need to show that $a + b \leq a - b$, which is the same as $b \leq -b$, which is certainly true since $b \leq 0$. In the other case, when $a + b \leq 0$, we have to show that $-(a + b) \leq a - b$ which is the same as $-a \leq a$, which is certainly true for $a \geq 0$.

Exercise 1.2. Verify the inequality in the case (C). □

Another proof of the triangle inequality is based on fact that

$$|a| = \sqrt{a^2}\,. \tag{1.6}$$

Therefore for all numbers a and b,

$$|a + b|^2 = (a + b)^2 = a^2 + 2ab + b^2 \leq |a|^2 + 2|a||b| + |b|^2 = (|a| + |b|)^2\,. \tag{1.7}$$

Exercise 1.3. Why does this imply (1.4)? See Problem 7.

1.5. The real number line and euclidean space

Many properties of the real numbers are intuitively understood by picturing them geometrically as points on a line. This is the **real (number) line** $\mathbb{R}$. To associate to each point on the line a number pick arbitrarily a point which we label 0, and one to the right which we label 1, which fixes the *scale*, and we arrive at all natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$ by marking equal distances to the right, and at all integers $\mathbb{Z} = \{\ldots, -2, 1, 0, 1, 2, \ldots\}$ by proceeding similarly to the left. By subdividing the interval from 0 to 1, it is clear how to make sense of the rational numbers $\mathbb{Q} = \{p/q : p \in \mathbb{Z}, q \in \mathbb{N}\}$, and we take it for granted for now that also the *irrational* numbers fit into this scheme. The relation $a < b$ can be interpreted to mean that a lies to the left of b, and the number $|a - b|$ is the distance between a and b.

We frequently encounter the set of numbers x which satisfy $|x-a|<\varepsilon$. This is the collection of points whose distance from a is less than $\varepsilon>0$. This is an example of an **open interval**

$$(a-\varepsilon, a+\varepsilon)=\{x: a-\varepsilon<x<a+\varepsilon\} \tag{1.8}$$

There are also **closed intervals** denoted by

$$[a, b]=\{x: a \leq x \leq b\} \tag{1.9}$$

where it is always understood that $a \leq b$. Moreover we write

$$(a, \infty)=\{x: x>a\} \qquad [a, \infty)=\{x: x \geq a\} \tag{1.10}$$

but we emphasize that the symbol ∞ is purely suggestive, because there is no number "∞" such that $a<\infty$ for all $a \in \mathbb{R}$. Nonetheless with this notation the real line can be viewed as an interval:

$$\mathbb{R}=(-\infty, \infty) \tag{1.11}$$

We can intersect two real number lines at a right angle to form the plane $\mathbb{R}^2$ consisting of *pairs* of numbers (a, b), and we also refer to a and b as the **coordinates** of the point (a, b). Alternatively, we can view $\vec{v}=(a, b)$ as a **vector**, namely an arrow from the origin to the point (a, b). This allows us add two vectors $\vec{v}=(a, b)$, $\vec{w}=(c, d)$, and interpret $v+w=(a+c, b+d)$ geometrically. Similarly we can define $\lambda \vec{v}=(\lambda a, \lambda b)$ for any real number λ, and interprete the operation of *scalar multiplicaion* geometrically is scaling. The **distance** between two points in the plane (a, b), and (c, d) is

$$\sqrt{(a-c)^2+(b-d)^2}\,. \tag{1.12}$$

More generally, we can view an ordered n-tuple of real numbers $(a_1, a_2, \ldots, a_n)$ as the coordinates of a point in the n-dimensional space $\mathbb{R}^n$. Alternatively, we can view $\vec{a}=(a_1, a_2, \ldots, a_n)$ as a **vector**. The geometry of **euclidean space** is tied to the **dot product**

$$\vec{a} \cdot \vec{b}=\sum_{i=1}^n a_i b_i\,, \tag{1.13}$$

which allows us to speak of vectors which are **orthogonal**, and to define the **length** of a vector. We say two vectors $\vec{a}$ and $\vec{b}$ are orthogonal if $\vec{a} \cdot \vec{b}=0$. Moreover, we define the length (or **norm**) of a vector $\vec{a}$ by

$$|\vec{a}|=\sqrt{\vec{a} \cdot \vec{a}}=\sqrt{a_1^2+a_2^2+\ldots+a_n^2}\,. \tag{1.14}$$

With this notion of length the **triangle inequality** that we have seen above for numbers extends to vectors:

$$|\vec{a}+\vec{b}| \leq|\vec{a}|+|\vec{b}|\,. \tag{1.15}$$

Remark 1.1. Indeed, it follows from the definition that

$$|\vec{a}+\vec{b}|^2 = |\vec{a}|^2 + 2\vec{a}\cdot\vec{b} + |\vec{b}|^2 \tag{1.16}$$

and so the stated inequality follows, provided we can show that $|\vec{a}\cdot\vec{b}| \leq |\vec{a}||\vec{b}|$. This is **Cauchy's inequality** ; (see the **additional notes** to Module I).

The **distance** between two points $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ in $\mathbb{R}^n$ is defined by $|\vec{a}-\vec{b}|$; note that this agrees with (1.12) in $\mathbb{R}^2$. The reason (1.15) is called the triangle inequality is because it implies that for any vectors $\vec{a}, \vec{b}, \vec{c}$,

$$|\vec{a}-\vec{c}| \leq |\vec{a}-\vec{b}| + |\vec{b}-\vec{c}|\,. \tag{1.17}$$

In other words, the distance from $\vec{a}$ to $\vec{c}$ is at most the sum of the distances from $\vec{a}$ to $\vec{b}$, and from $\vec{b}$ to $\vec{c}$, for any intermediate point $\vec{c}$.

Problems

1. Prove the following formula by induction:

$$1^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

2. Find a formula for

$$\sum_{i=1}^{n}(2i-1) = 1+3+5+\ldots+(2n-1)$$

3. Find all the numbers x for which
   a) $4-x < 3-2x$
   b) $5-x^2 < -2$
   c) $(x-1)(x-3) > 0$
   d) $x^2-2x+2 > 0$
   e) $x^2-x+10 > 16$
   f) $\frac{1}{x} + \frac{1}{1-x} > 0$
   g) $|x-3| = 8$
   h) $|x-3| < 8$
   i) $|x+4| < 2$
   j) $|x-1| + |x-2| > 1$

4. a) Prove by induction on n that

$$1+r+r^2+\ldots+r^n = \frac{1-r^{n+1}}{1-r}$$

b) Derive this result by setting $S = 1 + r + r^2 + \ldots + r^n$, multiplying this equation by r, and solving the two equations for S.

5. Prove that $\sqrt{3}$ is irrational.

   Hint: Every interger is of the form $3n$, or $3n + 1$, or $3n + 2$.

6. The following is a recursive definition of a^n: $a^1 = a$, and $a^{n+1} = a^n \cdot a$. Prove, by induction, that
$$a^{n+m} = a^n \cdot a^m \qquad (a^n)^m = a^{nm}\,.$$

7. Use the basic properties of Section 1.3 to prove the following:
   a) If $a < b$ and $c < d$, then $a + c < b + d$.
   b) If $a < b$, then $-b < -a$.
   c) If $a < b$ and $c > d$, then $a - c < b - d$.
   d) If $a < b$ and $c < 0$ then $ac > bc$.
   e) If $a > 1$ then $a^2 > a$.
   f) If $0 < a < 1$ then $a^2 < a$.
   g) If $0 \le a < b$ and $0 \le c < d$ then $ac < bd$.
   h) If $0 \le a < b$, then $a^2 < b^2$.
   i) If $a, b \ge 0$ and $a^2 < b^2$, then $a < b$.

8. Let $\vec{v} = (v_1, v_2)$, and $\vec{w} = (w_1, w_2)$ be vectors in the plane.
   a) Given v, find a vector w such that $v \cdot w = 0$.
   b) Show that $v \cdot (w + z) = v \cdot w + v \cdot z$
   c) Show that $v \cdot w = \frac{1}{4}(|v + w|^2 - |v - w|^2)$

9. Let $\vec{a} = (3, -1, 2)$ and $\vec{b} = (2, 1, 0)$. Compute the norms of $\vec{a}$ and $\vec{b}$.

10. Given vectors $\vec{x}$, and $\vec{y}$ in $\mathbb{R}^n$ show that
    a) $|\vec{x} + \vec{y}|^2 = |\vec{x}|^2 + 2\vec{x} \cdot \vec{y} + |\vec{y}|^2$
    b) $|\vec{x} + \vec{y}|^2 + |\vec{x} - \vec{y}|^2 = 2(|\vec{x}|^2 + |\vec{y}|^2)$

11. Show that $||\vec{a}| - |\vec{b}|| \le |\vec{a} - \vec{b}|$ for every $\vec{a}$, $\vec{b}$ in $\mathbb{R}^n$.

Note 2.

Functions and Graphs

2.1. Functions

A **function** f is a rule which assigns to each real number x (or only certain real numbers x) another real number $f(x)$, the *value of f at x*.

Example 2.1. The function which assigns to each number the square of that number,

$$f(x) = x^2 \tag{2.1}$$

Example 2.2. The rule which assigns to each number $c \neq 1, -1$ the number $c^3/(c^2-1)$,

$$f(c) = \frac{c^3}{c^2-1} \qquad (c \neq 1, -1) \tag{2.2}$$

Example 2.3. The rule that assigns to each number t the number $t^3 + x$. This rule obviously depends on x, and defines a *family of functions* f_x,

$$f_x(t) = t^3 + x \tag{2.3}$$

Remark 2.1. A function need not be expressed by an algebraic formula, it can be *any* rule that assigns numbers to certain other numbers.

Example 2.1 is a special example of an extremely important class of functions, namely the **polynomial functions**. If $a_0, a_1, \ldots, a_n$ are real numbers, and $a_n \neq 0$, then we say

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \tag{2.4}$$

is a polynomial of **degree** n.

Example 2.2 is an example of a **rational function**, namely a function of the form p/q where p, and q are polynomials.

Given two functions f, g, they can be combined to form a new function in various ways: $f + g$ is called the **sum** of f and g, $f \cdot g$ is the **product**, and f/g the **quotient** of f, and g. All these are defined in the obvious ways, but we already see that some thought has to be given to the **domain**: For example $(f+g)(x) = f(x) + g(x)$ only makes sense for numbers x for which *both* f and g are defined. The *domain* of a function is the set of numbers to which the rule can be applied. So in the above example if A is the domain of f, and B the domain of g, then $f + g$ is only defined on the *intersection* of A, and B, denoted by $A \cap B$. Similarly for $f \cdot g$.

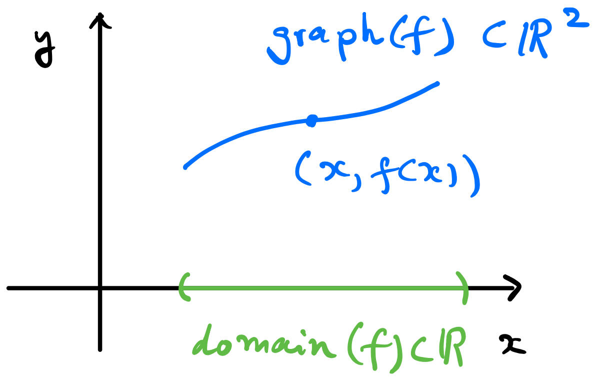

Figure 2.1.: Graph of a function.

In order to indicate the domain A of a real valued function f, we write

$$f : A \longrightarrow \mathbb{R} . \tag{2.5}$$

This is often also indicated implicitly, when we give a formula for $f(x)$ followed by

$$(x \in A) .$$

Exercise 2.1. What is the domain of f/g?

If f and g are any two functions, then the **composition** $f \circ g$ of f and g is defined by

$$(f \circ g)(x) = f(g(x)) \tag{2.6}$$

on the domain $\{x : x \text{ is in the domain of } g, \text{and } g(x) \text{ is in the domain of } f\}$.

Example 2.4. The function $x \mapsto \sin(x^2)$ is the composition of the functions $f(x) = \sin(x)$ and $g(x) = x^2$.

2.2. Graphs

We call the **graph** of a function f the set of points

$$(x, f(x)) \in \mathbb{R}^2 \tag{2.7}$$

where x is in the domain of f; see Fig. 2.2.

Remark 2.2. In fact, one could go one step further and *identify* a function with its graph. That leads to the formal definition of a function mentioned below (Additional).

The graph of a **linear function** $f(x) = cx + d$ is a straight line with **slope** c through the point $(0, d)$.

Exercise 2.2. Given two points (a, b) and (c, d) find the linear function f whose graph goes through both points.

Example 2.5. The graph of the function $f(x) = x^2$ is a **parabola**.

Exercise 2.3. Sketch the graphs of

1. $f(x) = x^n$ for $n = 2, 3, 4, \ldots$
2. $f(x) = x^2 + x$
3. $f(x) = x^3 - 3x$
4. $f(x) = 1/x \quad (x \neq 0)$
5. $f(x) = \frac{1}{1+x^2}$

Example 2.6. Consider the function $f(x) = \sin(1/x)$ on its domain $\mathbb{R} \setminus \{0\}$. To draw the graph it helps to observe that

$$f(x) = 0 \quad \text{for } x = 1/\pi, 1/2\pi, 1/3\pi, \ldots \tag{2.8}$$

$$f(x) = 1 \quad \text{for } x = \frac{1}{\pi/2}, \frac{1}{\pi/2 + 2\pi}, \frac{1}{\pi/2 + 4\pi}, \ldots \tag{2.9}$$

$$f(x) = -1 \quad \text{for } x = \frac{1}{3\pi/2}, \frac{1}{3\pi/2 + 2\pi}, \frac{1}{3\pi/2 + 4\pi}, \ldots \tag{2.10}$$

Moreover, when x is large, $1/x$ is small, so also $f(x)$ is small; similarly when $|x|$ is large, for $x < 0$. We take away from this that while the graph of f approaches the horizontal axis as $|x| \to \infty$ (from above on the right, and from below on the left), it *oscillates* infinitely many times between -1, and 1 near 0.

Exercise 2.4. Draw the graph of $f(x) = x \sin(1/x)$.

2.3. Functions of two variables

A function of two variables is a rule f that assigns to each point in the plane (x, y) a real number, the value of f at that point, denoted by $f(x, y)$.

Example 2.7. $f(x, y) = 1 - 2x - y$

Example 2.8. $f(x, y) = x^2 + y^2$

Example 2.9. $f(x, y) = 2xy/(x^2 + y^2)$, $(x, y) \neq (0, 0)$.

The first two are examples of **polynomials in two variables**, namely the sum of terms of the form $ax^m y^n$, where a is a constant and m, and n are nonnegative integers. The number $m + n$ is called the **degree** of this term, and the degree of the polynomial is the highest degree of the terms it contains.

As we have seen for functions of one variable, one can gain a great deal of insight by associating with a function $f(x)$ its **graph**, consisting of those points in the xy plane for

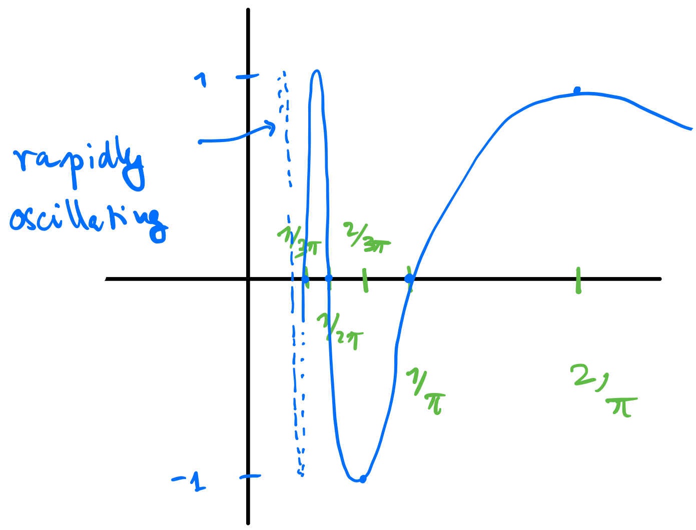

Figure 2.2.: The graph of the function $f(x) = \sin(1/x)$ for $x > 0$.

which $y = f(x)$. Similary when visualizing a function $f(x, y)$ of two variables, it is useful to think about the graph of f as a *surface* in $\mathbb{R}^3$ consisting of all points

$$(x, y, z = f(x, y)) \in \mathbb{R}^3 . \tag{2.11}$$

Example 2.10. The graph of the function in Example 2.7 is a plane. It intersects the z-axis at $z = 1$, because $f(0, 0) = 1$. Similarly we obtain the intersections with the x and y axes, and these three points, $(0, 0, 1)$, $(1/2, 0, 0)$, and $(0, 1, 0)$ determine uniquely the plane.

Example 2.11. The graph of the function in Example 2.8 is a *surface of revolution*, since the value of f depends only on the distance $|(x, y)|$ from the z-axis. The surface is a *paraboloid*.

Example 2.12. An important difference between Example 2.9 and the previous ones, is that here f is not defined for all values of x, and y. More generally the **quotient** of two polynomials

$$f(x, y) = \frac{P(x, y)}{Q(x, y)} \tag{2.12}$$

is not defined at points where the denominator vanishes. In Example 2.9 this a single point, the origin. However, since the numerator also vanishes, it is not immediately clear how this function behaves near the origin.

Exercise 2.5. Introduce polar coordinates in the plane to study the behaviour of the function $f(x, y) = 2xy/(x^2 + y^2)$ near the origin. Set $x = r\cos(\theta)$, and $y = r\sin(\theta)$, to express $f(x, y)$ as a function of r, and θ. Sketch the graph of f.

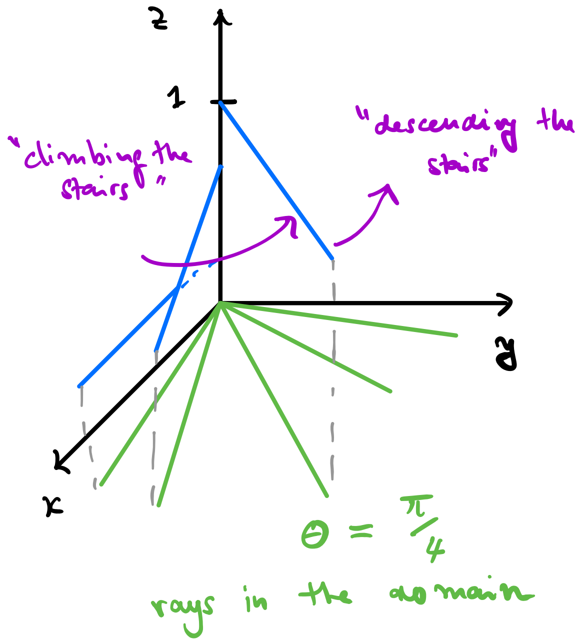

Figure 2.3.: The graph of the function from Example 2.9 is generated by rotating a ray from center, while moving it up and down.

Solution 2.6. We have

$$f(r\cos(\theta), r\sin(\theta)) = 2\cos(\theta)\sin(\theta) = \sin(2\theta)\,. \tag{2.13}$$

In other words, the value only depends on θ, not on r. Therefore the graph of f is constant over the rays $(r\cos\theta, r\sin(\theta))$, $r > 0$, for constant θ, and the graph $z = f(x, y)$ lies entirely between the planes $z = 1$, and $z = -1$. We can visualize the surface as a kind of "spiral ramp", see Fig. 2.3.

A common way to sketch a function of two variables is to consider the intersections of the graph of f in $\mathbb{R}^3$ with a plane $ax + by = 0$. In other words, by restricting ourselves to points in the xy plane with $ax + by = 0$ for some fixed a, and b, the values of f are a function of one variable, and its graph is visualized as a curve in the plane.

Example 2.13.

$$f(x, y) = \cos(xe^y) \tag{2.14}$$

Note that the graph of f lies between the horizontal planes $z = 1$ and $z = -1$. On each straight line $y = c$ in the xy plane, the function $f(x, c) = \cos(kx)$, with $k = e^c$, is an oscillating function in x, which is oscillating more rapidly the larger the value of c. The graph of $x \to f(x, c)$ is the cross section of the graph of f, as a surface in $\mathbb{R}^3$, with the planes $y = c$.

Exercise 2.7. Sketch the graph of the function of the previous example.

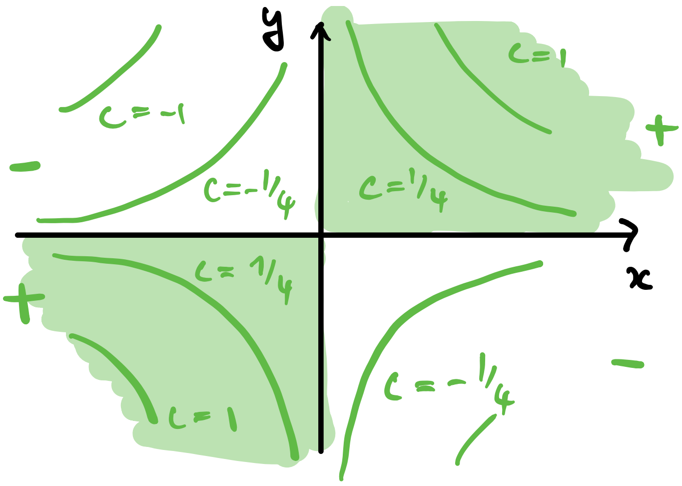

Figure 2.4.: Level sets of the function $f(x, y) = xy$.

Another approach is to intersect the graph of a function of two variables with the planes $z = c$. Here we are interested in the set of points (x, y) in the plane for which f takes a given value c. The set of points (x, y) in the xy plane for which $f(x, y) = c$, where c is a given constant, is a **level curve** of f. By choosing various values of c, and constructing the corresponding level curves, we can often obtain a picture of the graph of f.

Exercise 2.8. Sketch the level curves of the function

$$f(x, y) = xy\,. \tag{2.15}$$

Convince yourself that the graph of f is a "saddle-shaped" surface.

Solution 2.9. The level curves $xy = c$ are **hyperbolas** if $c \neq 0$. For $c > 0$, they lie in the first and third quadrants, while for $c < 0$, they lie in the second and forth quadrants of the xy plane; see Fig. 2.4. The level curve $xy = 0$ consists of the set of points on the x and y axes.

Exercise 2.10. Ask yourself what is the relevance of level curves to reading a map of a mountainous terrain?

Problems

1. Let

$$f(x) = \frac{1}{1 + x} \tag{2.16}$$

   What is

a) $f(1/x)$

b) $f(x+y)$

c) $f(x)+f(y)$

2. Find the domain of the functions defined by the following formulas

 a) $f(x)=\sqrt{1-x^2}$

 b) $f(x)=\frac{1}{x-1}+\frac{1}{x-2}$

3. A function is **even** if $f(x)=f(-x)$ and **odd** if $f(x)=-f(-x)$. For example, the functions $f(x)=x^2$, and $f(x)=|x|$ are *even*, while the function $f(x)=x$, or $f(x)=\sin(x)$ are *odd*.

 a) Determine whether $f+g$ is even, odd, or not necessarily either, in the four cases obtained by choosing f even or odd, and g even or odd.

 b) Do the same for $f\cdot g$, and $f\circ g$.

 c) Prove that every even function f can be written as $f(x)=g(|x|)$, with a function g that is *not* uniquely determined.

4. Prove or give a counterexample for each of the following assertions:

 a) $f\circ(g+h)=f\circ g+f\circ h$

 b) $(g+h)\circ f=g\circ f+h\circ f$

 c) $\frac{1}{f\circ g}=f\circ(1/g)$

5. Indicate on the real line the set of x satisfying the following relations, and write these sets using the notation of intervals.

 a) $|x-3|\leq 1$

 b) $\frac{1}{1+x^2}\leq a$

 c) $|x^2-1|<\frac{1}{2}$

6. Draw the set of all points (x,y) satisfying the following conditions.

 a) $x>y$

 b) $|x-y|<1$

 c) $1/(x+y)$ is an integer.

 d) $x=y^2$

 e) $x=|y|$

7. Sketch the graphs of the following functions, by plotting enough points to get a good idea of its shape.

 a) $f(x)=x-1/x$

 b) $f(x)=x^2+1/x^2$

8. Describe the general features of the graph of f if
   a) f is even
   b) f is odd
   c) f is **periodic**, namely $f(x) = f(x + a)$ for all x with some **period** $a > 0$.
9. For each of the following functions of two variables,
   a) $f(x, y) = x^2 - y^2$
   b) $f(x, y) = xy^2$
   c) $f(x, y) = \sin(x + y)$

   draw the intersection of their graphs with the planes,
   a) $y = c$
   b) $x = c$
   c) $z = c$ (level curves)

   for various values of c, and use these to sketch the graph of each function.
10. For each of the following functions of two variables,
    a) $f(x, y) = x + y^2$
    b) $f(x, y) = x^2 + 2xy + y^2$
    c) $f(x, y) = x^3 - 3xy^2$

    draw the level curve $z = 0$, and shade the region where $f(x, y) > 0$.
11. Let $h(t)$ be a strictly increasing function of t, and let $g(x, y) = h(f(x, y))$.
    a) How are the level curves of f and g related?
    b) How are the graphs of f and g related?

Additional: Vectors, functions, and graphs

Further Reading

(Folland, *Advanced Calculus*, Chapter 1.1)

Vector Calculus: Advanced (MAST20032)

Euclidean Spaces and Vectors

Similarly to numbers we have the basic operations of **addition** and **scalar multiplication** for vectors $\vec{x} = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$:

$$\vec{x} + \vec{y} = (x_1 + y_1, \ldots, x_n + y_n) \tag{2.1}$$

$$\lambda\vec{x} = (\lambda x_1, \ldots, \lambda x_n) \qquad \lambda \in \mathbb{R} \tag{2.2}$$

We have already introduced the dot product between two vectors and the norm of a vector. The fundamental inequality relating the two is **Cauchy's inequality**:

Proposition 2.1. *For any $\vec{x}, \vec{y} \in \mathbb{R}^n$, we have*

$$|\vec{x} \cdot \vec{y}| \leq |\vec{x}||\vec{y}| \,. \tag{2.3}$$

Exercise 2.1. This inequality has a geometrically instructive proof. Draw pictures while you are reading it!

Proof. Let us assume $\vec{b} \neq 0$. Consider the function

$$f(t) = |\vec{a} - t\vec{b}|^2 = |\vec{a}|^2 - 2t\vec{a} \cdot \vec{b} + t^2|\vec{b}|^2 \,. \tag{2.4}$$

We know, on one hand, that $f(t) \geq 0$ is non-negative. On the other hand, this is a quadratic in t, and its minimum occurs at $t = \vec{a} \cdot \vec{b}/|\vec{b}|^2$, where f takes the value

$$f(\vec{a} \cdot \vec{b}/|\vec{b}|^2) = |\vec{a}|^2 - \frac{(\vec{a} \cdot \vec{b})^2}{|\vec{b}|^2} \,. \tag{2.5}$$

Since $f \geq 0$, in particular at the minimum, we obtain the inequality after multiplying through by $|\vec{b}|^2$. □

A consequence is the **triangle inequality**:

Proposition 2.2. *For any $\vec{x}, \vec{y} \in \mathbb{R}^n$, we have*

$$|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}| \,. \tag{2.6}$$

Proof. We have $|\vec{x} + \vec{y}|^2 = |\vec{x}|^2 + 2\vec{x} \cdot \vec{y} + |\vec{y}|^2$, and by Cauchy's inequality $|\vec{x} \cdot \vec{y}| \leq |\vec{x}||\vec{y}|$. Therefore,

$$|\vec{x} + \vec{y}|^2 \leq (|\vec{x}| + |\vec{y}|)^2 \,, \tag{2.7}$$

so the inequality follows by taking square roots. □

Further Reading

(Spivak, *Calculus*, Chapter 3, 4)

A mathematical definition of functions and domains

We have introduced functions as "rules" but we will now illustrate how to turn this idea into a proper mathematical definition, which in particular avoids the confusion that while for example

$$f(x) = x^2 \tag{2.8}$$

and

$$f(x) = x^2 + 3x + 3 - 3(x + 1) \tag{2.9}$$

are *different* rules, they certainly define the *same* function.

Definition 2.1. A **function** is a collection of pairs of numbers with the following property: if (a, b) and (a, c) are both in the collection, then $b = c$.

Definition 2.2. The **domain** of f is the set of numbers of all a for which there is some b such that (a, b) is in f.

If a is in the domain of f, it follows from the first definition that there is a *unique* number b such that (a, b) is in f. This unique number is denoted by $f(a)$.

Mappings

More generally, a **map** (or **mapping**) is a rule f that assigns to each element of some set A an element of some other set B (possibly equal to A). We write $f : A \to B$. If $x \in A$, the element in B assigned to x by f is called the value $f(x)$. Thus functions are maps, but the term "function" is typically reserved for mappings whose values are real numbers (or complex numbers).

Given $f : A \to B$ we refer to A as the domain of f. If S is a subset of A, we denote by

$$f(S) = \{f(x) : x \in A\}\,, \tag{2.10}$$

and the set $f(A)$ (a subset of B) is called the **range** of f.

A mapping $f : A \to B$ is said to be **invertible** if there is another mapping $g : B \to A$ such that $g(f(x)) = x$ for all x in A and $f(g(y)) = y$ for all y in B. If this mapping exists g is called the inverse of f.

Figures in plane geometry

It is important to realise however that some of the most important figures of *plane geometry* are *not* graphs of functions. Let us first introduce the notion of distance, as captured by the *Pythagorean theorem.*

Definition 2.3. The **distance** d between two points (a, b) and (c, d) in the plane $\mathbb{R}^2$ is defined by

$$d = \sqrt{(a-c)^2 + (b-d)^2} \tag{2.11}$$

Example 2.1. The **circle** with centre (a, b) of radius r is the set of points (x, y) whose distance from (a, b) is equal to r. Since for example both $(a, b + r)$, and $(a, b - r)$ are in this collection of points, it is *not* a graph of a function.

Example 2.2. Given two points in the plane, an **ellipse** is the set of points, for which the sum of the distances to the two **focal** points is constant. If we take for simplicity the two focal points to be $(-c, 0)$, and $(c, 0)$ on the horizontal axis, and the distance to be $2a > 0$, then these are all points (x, y) for which

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a \tag{2.12}$$

After some calculation one finds this implies

$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1\,, \tag{2.13}$$

where we take $a > c$.

Example 2.3. The hyperbola is defined analogously, except that we require the *difference* of the two distances to be a constant. These leads to the same equation, but we now take $c > a$. The hyperbola has two **branches** because we can take the difference in two different ways. Note that while the hyperbola is also *not* a graph, we can write for example in the case $a^2 = 2$, and $a^2 - c^2 = -2$, that

$$(x + y)(x - y) = 2 \tag{2.14}$$

and so hyperbola coincides with the graph of the function $f(x) = 1/x$ *after a rotation of the axes* by an angle of $\pi/4$.

Additional problems

1. a) If f is any function, define a new function $|f|$ by $|f|(x) = |f(x)|$. If f and g are functions, define two new functions $\max(f, g)$, $\min(f, g)$ by

$$\max(f, g)(x) = \max(f(x), g(x)) \tag{2.15}$$
$$\min(f, g)(x) = \min(f(x), g(x)) \tag{2.16}$$

   Find an expression for $\max(f, g)$, and $\min(f, g)$ in terms of $|\cdot|$.

   b) Show that $f = f_+ + f_-$ where $f_+ = \max(f, 0)$ is the **positive part**, and $f_- = \min(f, 0)$ is the **negative part** of f.

   c) A function f is called **nonnegative** if $f(x) \geq 0$ for all x. Prove that any function f can be written as $f = g - h$, where g and h are nonnegative, (and *not* uniquely determined).

2. Prove that the graphs of the linear functions

$$f(x) = mx + b \qquad f(x) = nx + c \tag{2.17}$$

   are perpendicular if $mn = -1$.

   Hint: Consider the triangle with vertices at $(0, 0)$, $(1, m)$, and $(1, n)$, use the Pythagorean theorem.

3. a) If $x_1, \ldots, x_n$ are distinct numbers, find a polynomial function f_i of degree $n-1$ which is 1 at x_i and 0 at x_j for $j \neq i$.

   b) Find a polynomial function f of degree $n - 1$ such that $f(x_i) = a_i$, where $a_1, \ldots, a_n$ are given numbers.

4. For which numbers a, b, c, d does the function

$$f(x) = \frac{ax + b}{cx + d} \tag{2.18}$$

   satisfy $f(f(x)) = x$ (for all x in the domain of $f \circ f$)?

5. Convince yourself that the set of points (x, y) satisfying

$$ax^2 + bx + cy^2 + dy + e = 0 \tag{2.19}$$

   is either a parabola, an ellipse, a hyperbola, or in a degenerate cases two lines, one line, a point, or the empty set.

Module II.

Limits, and Continuity

Note 3.
Limits

In this lecture we shall make precise one of the most important notions in Calculus, namely the *limit* of a function. We would like to say that "a function f approaches the limit l near a, if we can make $f(x)$ as close as we like to l by requiring that x be sufficiently close to, but not equal to, a."

Here it is *irrelevant* how or even if f is defined at the point a. For example the functions

$$f(x) = x^2\,, \qquad g(x) = x^2 \quad (x \neq a)\,, \qquad h(x) = \begin{cases} x^2 & x \neq a \\ b & x = a \end{cases} \tag{3.1}$$

should all have the same limit $l = a^2$ at a.

A way to picture what we mean by "we can make $f(x)$ as close as we like to l", is to draw the graph of f, and first choose an interval B around l, which determines two horizontal lines in the plane. Then "by requiring that x be sufficiently close to a" means that we can find an interval A around a, so that the graph of the function f above A lies between the two horizontal lines, except perhaps at a. The idea is that if "f approaches the limit l near a", this should be possible no matter how small we choose the interval B.

Example 3.1. Consider the function $f(x) = 3x$ with $a = 5$. The limit should be $l = 15$. Suppose we want to make that $f(x)$ is within $1/10$ of 15. This means we want

$$15 - \frac{1}{10} < 3x < 15 + \frac{1}{10} \tag{3.2}$$

which we can also write as

$$-\frac{1}{30} < x - 5 < \frac{1}{30} \tag{3.3}$$

or simply $|x - 5| < 1/30$. This means that as long as we take x to within $1/30$ of distance from a, $f(x)$ will be within a distance of $1/10$ from l.

Exercise 3.1. Convince yourself that the function $f(x) = x^2$ approaches $l = 9$ near $a = 3$ in this sense. Suppose we would like to make $f(x)$ within distance 1 from $l = 9$. How close does x have to be to $a = 3$?

Example 3.2. Consider the function $f(x) = 1/x$ for $x \neq 0$, and let us try to show that $f(x)$ approaches $l = 1/3$ near $a = 3$. Let us convince ourselves that $f(x)$ will be *arbitrarily* close to l, if x is *sufficiently* close $a = 0$. This means we would like to show that for *any* chosen distance $\varepsilon > 0$ the inequality

$$|\frac{1}{x} - \frac{1}{3}| < \varepsilon \tag{3.4}$$

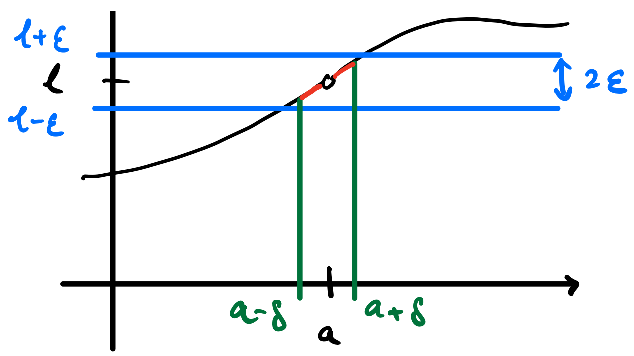

Figure 3.1.: Geometric interpretation of the definition of a limit.

can be satisifed *provided* x is in close range of a, namely $|x-3|$ is *small*. We begin with

$$|\frac{1}{x}-\frac{1}{3}| = |\frac{3-x}{3x}| = \frac{1}{3}\frac{1}{|x|}|x-3| \tag{3.5}$$

We have to make sure that the factor $1/|x|$ is not too large in the range we consider, so let us first require that $|x-3|<1$, which ensures that $2<x<4$, so that

$$\frac{1}{4} < \frac{1}{x} < \frac{1}{2} \tag{3.6}$$

and so $1/|x| < 1/2$ because $0<x=|x|$. Therefore

$$|\frac{1}{x}-\frac{1}{3}| < \frac{1}{6}|x-3| \qquad (|x-3|<1) \tag{3.7}$$

which shows that $|1/x-1/3|<\varepsilon$ provided $|x-3|<6\varepsilon$, which is also <1 as long as we started out with $\varepsilon<1/6$. The emphasis is here on being able to choose ε *small*, any upper bound on ε is irrelevant for the notion of a limit.

Definition 3.1 (Limit)**.** A function f **approaches the limit** l **near** a if for every $\varepsilon>0$, there is some $\delta>0$ such that, for all x, if $0<|x-a|<\delta$, then $|f(x)-l|<\varepsilon$.

This is a very important definition and you need to know it by heart!

Let us also make sure to get the logical negation of this statement right, namely to understand what it means for a function *not* to approach a limit l at a:

> A function does *not* approach the limit l at a, if there is *some* $\varepsilon>0$ such that for *every* $\delta>0$ there is *some* x which satisfies $0<|x-a|<\delta$, but *not* $|f(x)-l|<\varepsilon$.

Example 3.3. The function $f(x)=\sin(1/x)$ does not approach 0 near 0, because for $\varepsilon=1/2$ and any $\delta>0$, there is some x with $0<|x|<\delta$ such that $\sin(1/x)\geq 1/2$. Indeed, we only need to choose $x=1/(\pi/2+2\pi n)$ for some $n\in\mathbb{N}$, which becomes arbitrarily small for n large.

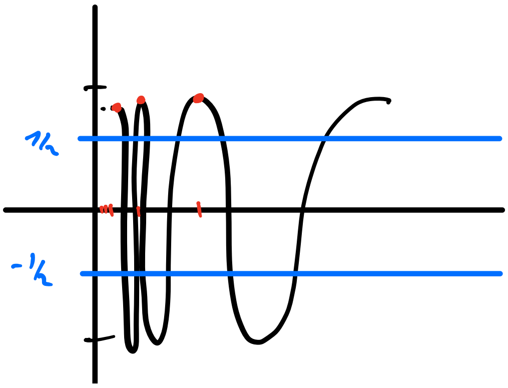

Figure 3.2.: The function $\sin(1/x)$ does not have a limit at 0.

Exercise 3.2. In fact, more is true: The function $f(x) = \sin(1/x)$ does not approach *any* limit l near 0.

Example 3.4. The function $f(x) = x\sin(1/x)$ approaches the limit 0 near 0. Since for all $x \neq 0$,

$$|x\sin(1/x)| \leq |x| \tag{3.8}$$

we can make $|f(x)| < \varepsilon$ simply by requiring that $0 < |x| < \delta$ with $\delta = \varepsilon$.

Exercise 3.3. Show that the function $f(x) = x^2\sin(1/x)$ approaches 0 near 0. What about $f(x) = \sqrt{x}\sin(1/x)$?

Since a function f cannot approach two different limits, we can talk about *the* limit l that f approaches near a, which is denoted by

$$\lim_{x\to a} f(x)\,. \tag{3.9}$$

The statement $\lim_{x\to a} f(x) = l$ has exactly the same meaning as the phrase "f approaches l near a." The possibility remains that f does not approach l near a for any l, and that is expressed by saying "$\lim_{x\to a} f(x)$ does not exist."

While the examples at the beginning of the lecture may give the impression that every function in question has to be dealt with separately, the idea is of course to establish general theorems which will make it easy to find limits.

Theorem 3.1 (Limit laws)**.** *If* $\lim_{x\to a} f(x) = l$ *and* $\lim_{x\to a} g(x) = m$*, then*

$$\lim_{x\to a}(f+g)(x) = l + m \tag{3.10}$$

$$\lim_{x\to a}(f\cdot g)(x) = l\cdot m \tag{3.11}$$

Moreover, if $m \neq 0$*, then*

$$\lim_{x\to a}\left(\frac{f}{g}\right) = \frac{l}{m} \tag{3.12}$$

Example 3.5. Using the Theorem we can prove, trivially, such statements as

$$\lim_{x\to a} \frac{x^3 + 7x^5}{x^2 + 1} = \frac{a^3 + 7a^5}{a^2 + 1} \tag{3.13}$$

without going through the laborious process of finding a δ, for a given ε.

We only give a proof of the first "limit law".

Proof. Let $\varepsilon > 0$. By assumption we know that there are $\delta_1, \delta_2 > 0$ such that, for all x,

$$\begin{aligned} 0 < |x - a| < \delta_1 &\Longrightarrow |f(x) - l| < \varepsilon/2 \\ 0 < |x - a| < \delta_2 &\Longrightarrow |g(x) - m| < \varepsilon/2. \end{aligned}$$

So if we choose $\delta = \min(\delta_1, \delta_2)$ to be the smallest of the two, then both statements are true for all $0 < |x - a| < \delta$, and moreover

$$|(f + g)(x) - (l + m)| = |f(x) - l + g(x) - m| \leq |f(x) - l| + |g(x) - m| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

whenever $0 < |x - a| < \delta$.

The proofs of the other two statements are very similar but the choices of δ are more intricate. For example, to achieve that

$$|f(x)g(x) - lm| < \varepsilon$$

we would write

$$|f(x)g(x) - lm| = |(f(x) - l)g(x) + l(g(x) - m)| \leq |f(x) - l||g(x)| + |l||g(x) - m|$$

and so a good choice of δ_1 ensures that

$$\text{if } 0 < |x - a| < \delta_1, \text{ then } |f(x) - l| < \frac{\varepsilon}{2(|m| + 1)}$$

which together with a choice of δ_2 such that

$$\text{if } 0 < |x - a| < \delta_2, \text{ then } |g(x) - m| < \min(1, \frac{\varepsilon}{2(|l| + 1)})$$

yields the desired inequality for all $0 < |x - a| < \delta$, where again $\delta = \min(\delta_1, \delta_2)$.

□

Problems

1. Find the following limits

   a) $\lim_{x\to 1} \frac{1-\sqrt{x}}{1-x}$

   b) $\lim_{x\to 0} \frac{1-\sqrt{1-x^2}}{x}$

2. In each of the following cases, determine the limit l for the given a, and prove that it is the limit by showing how to find a δ such that $|f(x) - l| < \varepsilon$ for all x satisfying $0 < |x - a| < \delta$.
   a) $f(x) = x^2 + 5x - 2$, $a = 2$
   b) $f(x) = x^4$, for any $a > 0$.
   c) $f(x) = \sqrt{|x|}$, $a = 0$.

3. Give an example of a function f for which the following assertion is *false*: If $|f(x) - l| < \varepsilon$ when $0 < |x - a| < \delta$, then $|f(x) - l| < \varepsilon/2$ when $0 < |x - a| < \delta/2$.

4. a) If $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ do not exist, can $\lim_{x\to a}(f(x) + g(x))$ exist? What about $\lim_{x\to a} f(x)g(x)$?
   b) If $\lim_{x\to a} f(x)$ exists and $\lim_{x\to a}(f(x) + g(x))$ exists, must $\lim_{x\to a} g(x)$ exist?
   c) If $\lim_{x\to a} f(x)$ exists and $\lim_{x\to a} g(x)$ does not exist, can $\lim_{x\to a}(f(x) + g(x))$ exist?

5. a) Prove that $\lim_{x\to a} f(x) = l$ if and only if $\lim_{x\to a}(f(x) - l) = 0$.
   b) Prove that $\lim_{x\to 0} f(x) = \lim_{x\to a} f(x - a)$.
   c) Prove that $\lim_{x\to 0} f(x) = \lim_{x\to 0} f(x^3)$.

6. Prove that if $\lim_{x\to 0} g(x) = 0$ and $|h(x)| \le M$ for all x, then $\lim_{x\to 0} g(x)h(x) = 0$.

Note 4.
Continuous functions

4.1. Definition of continuity

Intuitively, a function f is *continuous* if the graph contains no breaks, jumps, or wild oscillations.

Definition 4.1 (Continuity)**.** A function is **continuous at** a if

$$\lim_{x \to a} f(x) = f(a) \tag{4.1}$$

Remark 4.1. There are several ways this can fail. For example, f might not be defined at a, or the limit may not exist, in which cases this identity makes no sense. It could also be that f *is* defined at a *and* the limit of $f(x)$ at a exists, but these two numbers are not the same.

Example 4.1. The function $f(x) = \sin(1/x)$ is not continuous at 0, because it is not even defined at 0.

Example 4.2. The function $f(x) = x \sin(1/x)$ is not defined at 0 either, but the limit $\lim_{x\to 0} x \sin(1/x)$ exists and is 0, so while f is not continuous at 0, we can define an *extension* of f, namely the function

$$F(x) = \begin{cases} f(x) & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{4.2}$$

which is continuous at 0.

Example 4.3. Any monomial $f(x) = x^n$ is obviously continuous at any a because $\lim_{x\to a} x^n = a^n = f(a)$.

Theorem 4.1 (Continuity laws)**.** *If f and g are continuous at a, then $f + g$ and $f \cdot g$ are continuous at a, and moreover if $g(x) \neq 0$, then $1/g$ is continuous at a.*

Proof. This follows directly from Theorem 3.1. Indeed, if $\lim_{x\to a} f(x) = f(a)$ and $\lim_{x\to a} g(x) = g(a)$, then

$$\lim_{x\to a} (f+g)(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = f(a) + g(a) = (f+g)(a)\,. \tag{4.3}$$

Similarly for $f \cdot g$, and $1/g$. □

The theorem allows us to infer that *rational* functions are continuous at every point in their domain. We defer the proof that the *trigonometric* functions are continuous, but even if we know that we are still unable to prove the continuity of functions like $f(x) = \sin(x^2)$, before making a statement about *compositions*:

Theorem 4.2 (Composition of continuous functions)**.** *If g is continuous at a, and f is continuous at $g(a)$, then $f \circ g$ is continuous at a.*

Proof. Let $\varepsilon > 0$. Since f is continuous at $g(a)$ we can find a $\delta > 0$ such that, if $|y - g(a)| < \delta$, then $|f(y) - f(g(a))| < \varepsilon$. So now choose $\eta > 0$, so that for all x, if $|x - a| < \eta$, we have $|g(x) - g(a)| < \delta$, which is possible because g is continuous at a. □

Example 4.4. With this theorem we can now infer that $F(x)$ from (4.2) is continuous at *every* point. Similarly for functions like $f(x) = \sin(x^2 + \sin(x))$, etc.

Exercise 4.1. Give another proof of the quotient case in the limit laws using Theorem 4.2. In other words, use the theorem about compositions of continuous functions to show that if g is continuous at a, and $g(a) \neq 0$, then

$$\lim_{x \to a} \frac{1}{g(x)} = \frac{1}{g(a)} \,. \tag{4.4}$$

So far we have talked about continuity *at a point.* The consequences of continuity are more powerful when it refers to *continuity on an interval*: We say f is **continuous on** (a, b) if $f(x)$ is continuous at x for all $x \in (a, b)$. This is a special case of a function being **continuous on** $\mathbb{R} = (-\infty, \infty)$.

4.2. Consequences of continuity

A continuous function is sometimes described, intuitively, as one whose graph can be drawn without lifting your pencil off the paper. While Example 4.2 shows this description is a little too optimistic, it is true that all the following theorems about continuity are clear in this picture:

Theorem 4.3. *Suppose f is continuous at a, and $f(a) > 0$. Then $f(x) > 0$ for all x in some interval containing a.*

Proof. Since f is continuous at a, we know that for $\varepsilon = f(a) > 0$ there is a $\delta > 0$ so that for all x, if $|x - a| < \delta$, then $|f(x) - f(a)| < \varepsilon$, or $-\varepsilon < f(x) - f(a) < \varepsilon$, which means

$$0 < f(x) < 2f(a) \tag{4.5}$$

for all x in the interval $(a - \delta, a + \delta)$. □

Exercise 4.2. Formulate and prove an analogous statement under the assumption $f(a) < 0$.

The next theorem says, geometrically, that the graph of a continuous function which starts below the horizontal axis and ends above the horizontal axis must cross this axis at some point.

Theorem 4.4. *If f is continuous on an interval (c, d), $c < a < b < d$, and*

$$f(a) < 0 < f(b)\,,$$

then there is some x in (a, b) such that $f(x) = 0$.

This theorem requires continuity on the *whole interval.*

Exercise 4.3. Give an example of a function for which continuity fails to hold at *a single point*, and the conclusion is false.

There are simple generalisations of this theorem: Firstly there is nothing special about the number 0. In the situation that $f(a) < c < f(b)$, there is some $x \in (a, b)$ such that $f(x) = c$. Indeed, simply apply the theorem to the function $g = f - c$. Moreover if $f(a) > c > f(b)$, then we can apply the theorem to $g = -f + c$. This can be summarized in the **Intermediate Value Theorem**: *If a continuous function on an interval takes on two values, then it takes on every value in between.*

The proof of Theorem 4.4 relies on a deep property of the real numbers that we will not discuss in this subject.However, we can guess that the proof of the intermediate value theorem must rely on a property of the real numbers, as opposed to the rationals, because Theorem 4.4 implies immediately *the existence of square roots*:

Theorem 4.5. *Every positive number has a square root.*

Proof. Consider the function $f(x) = x^2$. We want to show that if $\alpha > 0$, there exists a number x such that $f(x) = x^2 = \alpha$. There is obviously a number b such that $f(b) > \alpha$. (In fact, if $\alpha > 1$, take for example $b = \alpha$, and if $\alpha < 1$, take for example $b = 1$.) Since f is continuous, and $f(0) = 0 < \alpha < f(b)$, there exists a number x in the interval $[0, b]$ such that $f(x) = \alpha$, so $x^2 = \alpha$.

□

Problems

1. For which of the following functions f is there a continuous extension F of f? In other words, for which of the following functions is there a continuous function F on the real line such that $F(x) = f(x)$ for all x in the domain of f.

   a) $f(x) = \frac{x^2-4}{x-2}$

   b) $f(x) = \frac{|x|}{x}$

   c) $f(x) = x^2 \sin(1/x^2)$

2. a) Suppose that f is a function satisfying $|f(x)| \leq |x|$ for all x. Show that f is continuous at 0.

   b) Suppose that g is continuous at 0 and $g(0) = 0$, and $|f(x)| \leq |g(x)|$. Prove that f is continuous at 0.

3. Prove that if f is continuous at a then for any $\varepsilon > 0$ there is a $\delta > 0$ so that whenever $|x - a| < \delta$ and $|y - a| < \delta$ we have $|f(x) - f(y)| < \varepsilon$.

4. Find an integer n such that $f(x) = 0$ for some x between n and $n + 1$, where

$$f(x) = x^3 - x + 3 \tag{4.6}$$

5. Suppose f is continuous, and that $f(x) = 0$ only for $x = a$.
   a) Suppose $f(x) > 0$ for some $x > a$ as well as for some $x < a$. What can be said about $f(x)$ for all $x \neq a$?
   b) Suppose $f(x) > 0$ for some $x > a$ and $f(x) < 0$ for some $x < a$. What can be said about $f(x)$ for $x \neq a$?

6. Suppose that f is a continuous function on $[0, 1]$ and that $f(x)$ is in $[0, 1]$ for each x. (Draw a picture!) Prove that $f(x) = x$ for some number x in the unit interval.

Note 5.

Theorems about continuity

5.1. Global properties of continuous functions

We have already seen one important theorem about continuity, namely the **Intermediate value theorem** which states that

Theorem 5.1 (Intermediate value theorem). *If f is continuous on $[a, b]$ and*

$$f(a) < t < f(b),$$

then there is some x in (a, b) such that $f(x) = t$.

Let us now state two more theorems about continuity and explore some of their consequences:

Theorem 5.2 (Bounded value theorem). *If f is continuous on $[a, b]$, then f is bounded above on $[a, b]$, that is there is some number N such that $f(x) \leq N$ for all x in $[a, b]$.*

Geometrically, this means that the graph of f lies below *some* horizontal line.

The third theorem states that a continuous function on a closed interval always achieves a maximum:

Theorem 5.3 (Extreme value Theorem). *If f is continuous on $[a, b]$, then there is some number y in $[a, b]$ such that $f(y) \geq f(x)$ for all x in $[a, b]$.*

These theorems all rely on the continuity of f on the *interval* $[a, b]$. Indeed, the conclusions are false if continuity fails at a single point

Example 5.1. For Theorem 5.2, take the function

$$f(x) = \begin{cases} 1/x & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{5.1}$$

which is continuous at every point except 0, but f is not bounded above.

Example 5.2. For Theorem 5.3, consider the function

$$f(x) = \begin{cases} x^2 & x < 1 \\ 0 & x \geq 1 \end{cases} \tag{5.2}$$

On the interval $[0, 1]$ the function is bounded above, but there is no y in $[0, 1]$ such that $f(y) \geq f(x)$ for all x in the interval.

These important theorems are stated in the simplest setting and are easily generalised. For example, a continuous function on a closed interval always achieves a minimum, too:

Exercise 5.1. Use Theorem 5.3 to show that if f is continuous on $[a, b]$ then there is some y in $[a, b]$ such that $f(y) \leq f(x)$ for all x in $[a, b]$.

5.2. Roots of polynomial equations

We illustrate the power of these theorems with an application to polynomials. We have already seen that Theorem 5.1 implies the existence of square roots. More generally, we have

Theorem 5.4. *If n is odd, then for any numbers $a_1, \dots, a_n$, the equation*

$$x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0 \tag{5.3}$$

has a root.

This is true because the function $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ is positive for large positive x, and negative for large negative x, so $f(x) = 0$ for some x. More precisely, for n odd,

$$\lim_{x \to \pm\infty} f(x) = \pm\infty\,. \tag{5.4}$$

Remark 5.1. The statement that $\lim_{x\to\infty} f(x) = \infty$ does not mean that f has a limit as x becomes large, but that f is unbounded, in the sense that, for every $N > 0$, there exists $M > 0$, so that,

$$x > M \Longrightarrow f(x) > N\,. \tag{5.5}$$

Similary for statements such as $\lim_{x\to a} f(x) = \infty$; (see the **additional notes** to Module II).

Exercise 5.2. Prove (5.4). If this is difficult, read on and return to it after the proof of the Theorem below.

Examples of equations like $x^2 + 1 = 0$ show that we cannot hope to show the same result for n even. However in this case the equation $x^2 + 1 = c$ has a solution for $c \geq 1$. While we are changing the problem — in effect we are playing with the constant a_0 — a result of this type can be proven more generally with the help of Theorem 5.3.

Theorem 5.5. *Suppose n is even. Then there is a number m such that*

$$x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = c \tag{5.6}$$

has a solution for $c \geq m$, and no solution for $c < m$.

The idea is here to show first that $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ has an absolute minimum, namely there is a number y such that $f(y) \leq f(x)$ for all x. Given such a y, set $m = f(y)$, then $f(x) = c$ obviously has no solution for $c < m$, but for $c = m$ we have that $x = y$ is a solution. In the case $c > m$ we can find a number $b > y$ with $f(b) > c$, so $f(y) = m < c < f(b)$ and we can apply Theorem 5.1 to infer the existence of a solution $f(x) = c$.

Exercise 5.3. Why can we find a number $b > y$ with $f(b) > c$?

It remains to show:

Lemma 5.6. *If n is even and $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$, then there is a number y such that $f(y) \leq f(x)$ for all x.*

Proof. First note that
$$\lim_{x\to\pm\infty} f(x) = \infty\,. \tag{5.7}$$
Hence we can find an interval $[-b, b]$ so large, that for any points x outside that interval,
$$|x| > b \Longrightarrow f(x) > |f(0)|\,. \tag{5.8}$$
Now we can use Theorem 5.3, to infer that f has a minimum on the closed interval $[-b, b]$:
$$f(x) \geq f(y) \tag{5.9}$$
for some $y \in (-b, b)$. The point y is a minimum not just on the interval $[-b, b]$, because for $|x| > b$ we have
$$f(x) > |f(0)| \geq f(0) \geq f(y)\,. \tag{5.10}$$
□

Solution 5.4. In the proofs of the theorems above we have used that for any polynomial of degree $n \in \mathbb{N}$,
$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0 \tag{5.11}$$
we have
$$\lim_{x\to\infty} f(x) = \infty\,. \tag{5.12}$$
We can see this as follows:
$$f(x) = x^n\Big(1 + \frac{a_{n-1}}{x} + \ldots + \frac{a_0}{x^n}\Big) \qquad (x > 0) \tag{5.13}$$
Now by choosing $M = \max(1, 2n|a_{n-1}|, \ldots 2n|a_0|)$ we get that for $|x| \geq M$,
$$\Big|\frac{a_{n-1}}{x} + \ldots + \frac{a_0}{x^n}\Big| \leq \Big(|a_{n-1}| + \ldots + |a_0|\Big)\frac{1}{M} \leq \frac{1}{2}\,. \tag{5.14}$$
Therefore, for all x, with $|x| \geq M$,
$$\frac{1}{2}x^n \leq x^n\Big(1 + \frac{a_{n-1}}{x} + \ldots + \frac{a_0}{x^n}\Big) = f(x) \tag{5.15}$$

This shows that for any $N > 0$, we can find $M > 0$, so that, if $x > M$, then $f(x) > N$. Indeed, for a given $N > 1$, let us choose $M = \max\{1, 2n|a_{n-1}|, \ldots, 2n|a_0|, 2N\}$, then, whenever $x > M$, we have
$$f(x) > \frac{1}{2}x^n > 2^{n-1}N^n > N\,. \tag{5.16}$$

5.3. What about the proof?

We have deferred the proofs of the theorems about continuity because they rely on a property of the real numbers that we have not discussed yet. Let us see what this property might be by trying to give a proof of Theorem 4.4, namely:

> *If f is continuous on $[a, b]$ and $f(a) < 0 < f(b)$, then there is some x in $[a, b]$ such that $f(x) = 0$.*

In the situation that $f(a) < 0$ a promising idea seems to be to locate the first point where $f(x) = 0$, namely *the smallest point* in $[a, b]$ such that $f(x) = 0$. To find this point consider the set of numbers A which contains all points x in $[a, b]$ such that f is negative on $[a, x]$. In view of Theorem 4.3 the set A contains some points close to a, while all points sufficiently close to b are *not* in A.

Now it is clear that we can find a number α that is greater than all the numbers in A (for example $\alpha = b$ is such a choice), but it is not so clear if we can find the *smallest* number α which is greater than all the numbers in A. It is precisely this property of the real numbers, that ensures the existence of α which allows us to proceed. A thorough discussion of "least upper bounds" is beyond the scope of this subject.

We claim that $f(\alpha) = 0$, and prove this by eliminating the possibilities that $f(\alpha) < 0$ or $f(\alpha) > 0$.

Suppose first that $f(\alpha) < 0$. Then by Theorem 4.3 $f(x) < 0$ for all x in a small interval containing α, in particular for some numbers greater than α, which contradicts that α is an upper bound for A.

On the other hand, if $f(\alpha) > 0$ then by Theorem 4.3 $f(x) > 0$ for all x in a small interval containing α, in particular for some numbers smaller that α, which means that these numbers are *not* in A, and we could have chosen a smaller upper bound for A, which contradicts that α is the smallest upper bound for A.

Hence $f(\alpha) = 0$.

Problems

1. For each of the following functions, decide which are bounded above and below on the indicated interval, and which take on their maximum or minimum value.
   a) $f(x) = x^2$ on $(-1, 1)$
   b) $f(x) = x^3$ on $(-1, 1)$
   c) $f(x) = x^2$ on $\mathbb{R}$
   d) $f(x) = x^2$ on $[0, \infty)$
   e) $f(x) = \begin{cases} x^2, & x \leq a \\ a + 2, & x > a \end{cases}$ on $(-a - 1, a + 1)$. (Assume here that $a > -1$.)
   f) $f(x) = \begin{cases} x^2, & x < a \\ a + 2, & x \geq a \end{cases}$ on $[-a - 1, a + 1]$. (Again assume here that $a > -1$.)

2. Suppose f and g are continuous on $[a, b]$ and that $f(a) < g(a)$, but $f(b) > g(b)$. Prove that $f(x) = g(x)$ for some x in $[a, b]$.

3. Suppose that f is a continuous function with $f(x) > 0$ for all x, and

$$\lim_{x \to -\infty} f(x) = 0 = \lim_{x \to \infty} f(x)\,. \tag{5.17}$$

   Prove that there is some number y such that $f(y) \geq f(x)$ for all x.

Additional: Limts and Continuity

Recommended Reading

(Spivak, *Calculus*, Chapter 5)

Real Analysis: Advanced (MAST20033)

In the lecture we have mentioned without proof that:

Theorem 5.1. *A function f cannot approach two different limits near a.*

Proof. We want to show that if f approaches l near a, and f approaches m near a, then $l = m$.

By definition, for a given $\varepsilon > 0$, there exist $\delta_1 > 0$, and $\delta_2 > 0$ such that

$$0 < |x - a| < \delta_1 \Longrightarrow |f(x) - l| < \varepsilon \tag{5.1}$$

$$0 < |x - a| < \delta_2 \Longrightarrow |f(x) - m| < \varepsilon \tag{5.2}$$

and in particular both conclusions are true provided $|x - a| < \delta$, if we choose $\delta = \min(\delta_1, \delta_2)$. If it were true that $l \neq m$, then we could choose $\varepsilon = |l - m|/2 > 0$, and it follows that for all x, if $0 < |x - a| < \delta$,

$$|l - m| = |l - f(x) + f(x) - m| \leq |l - f(x)| + |f(x) - m| < 2\varepsilon = |l - m| \tag{5.3}$$

which is a contradiction. □

Exercise 5.1. First interpret separately, precisely in the sense of Definition 3.1, and then prove equality of the expressions

$$\lim_{x \to a} f(x) \quad \text{and} \quad \lim_{h \to 0} f(a + h)\,. \tag{5.4}$$

There are times we would like to talk about the limit of f as x approaches a "from above", or "from below". These are situations when $f(x)$ may not be defined for all x in $|x - a| < \delta$, but only say for $x > a$, or $x < a$, but the "one-sided" limits still exist.

Definition 5.1. A function f has **a limit l as x approaches a from above**, if for every $\varepsilon > 0$ there is a $\delta > 0$, such that for x, if $0 < x - a < \delta$, then $|f(x) - l| < \varepsilon$. We then write

$$\lim_{x \to a^+} f(x) = l\,. \tag{5.5}$$

As an exercise write down the definition of a limit **from below**.

Example 5.1. Consider the function

$$f(x) = \begin{cases} -1, & x < 0 \\ 1, & x > 0 \end{cases} \tag{5.6}$$

This function does not approach any number near 0, but the limits from above and below do exist:

$$\lim_{x\to 0^+} f(x) = 1 \qquad \lim_{x\to 0^-} f(x) = -1\,. \tag{5.7}$$

Additional Problems

1. Prove that $\lim_{x\to a} f(x)$ exists if $\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x)$.

2. The function $f(x) = 1/x^2$ does not approach a limit near 0. Nonetheless it is common to write $\lim_{x\to 0} f(x) = \infty$. In general we define "$\lim_{x\to a} f(x) = \infty$" to mean that for all N there is a $\delta > 0$ such that, for all x, if $0 < |x-a| < \delta$, then $f(x) > N$.

   a) Show that $\lim_{x\to 3} 1/(x-3)^2 = \infty$.

   b) Prove that if $f(x) > \epsilon > 0$ for all x, and $\lim_{x\to a} g(x) = 0$, then

   $$\lim_{x\to a} \frac{f(x)}{|g(x)|} = \infty\,.$$

Recommended Reading

(Spivak, *Calculus*, Chapter 6, 7)

Real Analysis: Advanced (MAST20033)

Exercise 5.2. You might have noticed that in the proof of Theorem 4.2 we have taken $\lim_{x\to a} f(x) = f(a)$ to mean that for every $\varepsilon > 0$, there exists a $\delta > 0$, so that for all x, if $|x-a| < \delta$, then $|f(x) - f(a)| < \varepsilon$. So for a function f which is *continuous* at a, we have dropped the condition that $0 < |x-a|$. Why?

For the theorems about global properties of continuous functions, the notion of continuity on a *closed* interval is important. We say a function is **continuous on** $[a, b]$ if f is continuous at all x in (a, b), and

$$\lim_{x\to a^+} f(x) = f(a) \qquad \lim_{x\to b^-} f(x) = f(b)\,. \tag{5.8}$$

Recall here from Defintion 5.1 (Additional: Limits) what it means for a function to approach a limit *from above* or *from below.*

Another variation of the "one-sided" limit occurs when we talk about the limit of a function $f(x)$ "as x approaches ∞."

Example 5.2. We have

$$\lim_{x\to\infty} \sin(1/x) = 0\,. \tag{5.9}$$

Definition 5.2. A function $f(x)$ approaches a limit l "as x goes to infinity,"

$$\lim_{x\to\infty} f(x) = l$$

if for every $\varepsilon > 0$, there is a number $N \in \mathbb{N}$ such that, for all x,

$$x > N \Longrightarrow |f(x) - l| < \varepsilon\,.$$

Additional Problems

1. a) Prove the following version of Theorem 4.3 for "right-hand continuity": Suppose that $\lim_{x\to a^+} f(x) = f(a)$, and $f(a) > 0$. Then there is a number $\delta > 0$ such that $f(x) > 0$ for all x satisfying $0 \le x - a < \delta$. Similarly if $f(a) < 0$, then there is a number $\delta > 0$ such that $f(x) < 0$ for all x satisfying $0 \le x - a < \delta$.

   b) Prove a version of Theorem 4.3 when $\lim_{x\to b^-} f(x) = f(b)$.

2. a) Prove that if f is continuous on $[a, b]$, then there is a function g which is continuous on $\mathbb{R}$ which satisifies $g(x) = f(x)$ for all $x \in [a, b]$.

   b) Give an example to show that this assertion is false if $[a, b]$ is replaced by (a, b).

3. Prove that $\lim_{x\to 0^+} f(1/x) = \lim_{x\to\infty} f(x)$.

Note 6.

Functions of two variables: Limits and Continuity

In Note 3 we have arrived at a definition of what it means for a function of one variable to approach a limit at a point a. In this note we want to extend this notion to functions $f(x, y)$ of two variables.

Definition 6.1 (Limit)**.** A function $f(x, y)$ of two variables **has the limit** l **as** (x, y) **approaches** (a, b) if for every $\varepsilon > 0$, there is some $\delta > 0$ so that

$$|f(x,y) - l| < \varepsilon \text{ whenever } 0 < |(x,y) - (a,b)| < \delta\,.$$

In other words, the definition is *conceptually* exactly the same as for one variable just that $|\cdot|$ now refers to the distance in $\mathbb{R}^2$: The set of points (x, y) satisfying

$$0 < |(x,y) - (a,b)| = \sqrt{(x-a)^2 + (y-b)^2} < \delta \tag{6.1}$$

is a (punctured, and open) **disc** of radius δ centered at the point (a, b).

Remark 6.1. It is often more convenient to use the following *equivalent* formulation:

A function $f(x, y)$ has a limit l as (x, y) approaches (a, b) if for every $\varepsilon > 0$ there is a $\delta > 0$ so that $|f(x, y) - l| < \varepsilon$ whenever $(x, y) \neq (a, b)$ and

$$\max\{|x-a|, |y-b|\} < \delta\,.$$

Indeed, if $|f(x, y) - l| < \varepsilon$ on a punctured disc of radius δ, then this will also be the case on a puctured square with side length $2\delta'$, provided $\sqrt{2}\delta' \leq \delta$. Conversely, if $|f(x, y) - l| < \varepsilon$ on a punctured square of side-length 2δ, then this will be true for all points (x, y) on a punctured disc that fits into this square, namely a disc of radius δ.

The definition absolves us from the following dilemma. For a function of one variable $f(x)$ it is clear what we mean by "x approaches a": we can approach the point a either from the left, or the right. However, for functions of two variables there are infinitely many ways to approach the point (a, b) in the plane, because any curve in the plane through (a, b) gives a way to approach (a, b). Moreover, the value that a function approaches at a point may depend on the direction in which this point is approached.

Example 6.1. Let

$$g(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0)\,. \end{cases} \tag{6.2}$$

First observe that $g(x,0) = 0$ and $g(0,y)$. In particular, *along the coordinate axes* g goes to 0. Now let $c \neq 0$, and consider g evaluated on the points (x, cx):

$$g(x,cx) = \frac{cx^3}{x^4 + c^2x^2} = \frac{cx}{c^2 + x^2} \to 0 \qquad (x \to 0)\,. \tag{6.3}$$

This shows that *along any straight line through the origin* the function g tends to 0. However, if we approach the origin on an *parabola* $y = cx^2$, for any $c \neq 0$, then

$$g(x,cx^2) = \frac{c}{1+c^2} \neq 0\,. \tag{6.4}$$

Our definition does not make reference to the "way" in which (a, b) is approached. In fact, according to our definition the function $g(x, y)$ of the previous example does not have a limit at $(0, 0)$ at all!

Exercise 6.1. Why does our working of Example 6.1 show that $g(x, y)$ does not have a limit at the origin?

Exercise 6.2. Let

$$f(x,y) = \frac{xy}{x^2+y^2} \qquad (x,y) \neq (0,0) \tag{6.5}$$

and let $f(0, 0) = 0$. Show that f does *not* have a limit as (x, y) approaches $(0, 0)$.

Solution 6.3. First note that $f(0, y) = 0$, and $f(x, 0) = 0$. Now consider the values of f on the straight lines through the origin: For any $c \neq 0$,

$$f(x,cx) = \frac{cx^2}{x^2 + c^2x^2} = \frac{c}{1+c^2} \qquad (x \neq 0)\,. \tag{6.6}$$

We need to show that for any l, we can find $\varepsilon > 0$, so that for all $\delta > 0$, there exists a point (x, y) in the punctured disc of radius δ, where $|f(x, y) - l| > \varepsilon$.

For $l = 0$, we can always choose a point (x, x) on the straight line $y = x$ arbitrarily close to the origin, where $f(x, x) = 1/2$, so we can arrange this for any $\varepsilon < 1/2$.

For $l \neq 0$, simply choose $\epsilon = l/2$, then for any point on the axis $|g(x, 0) - l| = l > \varepsilon$.

Notice that the definition of a limit, which if it does exist for a function $f(x, y)$ at the point (a, b) we denote by

$$\lim_{(x,y)\to(a,b)} f(x,y)\,, \tag{6.7}$$

does not involve the value $f(a, b)$ at all; only the values of the function *near* (a, b) are relevant here. Indeed the function need not even be defined at (a, b). However, if f is defined at (a, b), and its value at the point agrees with its limit as we approach (a, b), then the function is said to be continuous at (a, b):

Definition 6.2 (Continuity)**.** A function $f(x, y)$ is **continuous** at (a, b) if it has a limit at (a, b), and

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b)\,. \tag{6.8}$$

While the examples above may have given the impression that functions of two variables, in general, do not have a limit, this is rather the exception than the rule:
The functions

$$f(x, y) = x + y \tag{6.9}$$
$$f(x, y) = xy \tag{6.10}$$

are continuous everywhere in the plane. Moreover

$$f(x, y) = x - y \tag{6.11}$$
$$f(x, y) = x/y \qquad (y \neq 0) \tag{6.12}$$

are continuous in their domain. Furthermore; (see the **additional notes** to Module II):

(Continuity laws) the sum, and product of two continuous functions is continuous. Moreover, **the quotient of two continuous functions is continuous** (on the set where the denominator is not zero).

Since the "elementary functions" of one variable, in particular polynomials and trigonometric functions, are all continuous (on their domains), it is almost immediate that all the "elementary functions" of two variables, namely those built up of these functions of one variable, by arithmetic operations and compositions, are also continuous, where they are defined.

Example 6.2. The function

$$f(x, y) = \frac{\sin(3x + 2y)}{x^2 - y} \tag{6.13}$$

is continuous everywhere, except along the parabola $y = x^2$.

Exercise 6.4. Let f be defined by

$$f(x, y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0)\,. \end{cases} \tag{6.14}$$

Decide if h is continuous at the origin.

Solution 6.5. Experimentation with lines does not give any indication that the limit of f near $(0, 0)$ does not exist, or is anything other than 0. So let us try to show that

$$\lim_{(x,y)\to(0,0)} f(x, y) = 0 \tag{6.15}$$

which, if true, then shows that f is continuous at the origin. Since $|x^2 - y^2| \leq x^2 + y^2$, we have

$$|f(x, y)| \leq |xy| \tag{6.16}$$

for all $(x, y) \neq (0, 0)$. Since the function $h(x, y) = xy$ is continuous at the origin, it follows that f is continuous at the origin. Indeed, let $\varepsilon > 0$. Then for all $(x, y) \neq (0, 0)$, with $|(x, y)| < \delta$, we have

$$|f(x, y)| \leq |xy| \leq \frac{1}{2}(x^2 + y^2) < \varepsilon\,, \tag{6.17}$$

provided δ is chosen so that $\delta^2 < 2\varepsilon$, say $\delta = \sqrt{\varepsilon}$.

Problems

1. Show that the following functions do *not* have a limit at the origin:
   a) $f(x, y) = \frac{x^2+y}{\sqrt{x^2+y^2}} \qquad (x, y) \neq (0, 0)$
   b) $f(x, y) = \frac{x}{x^4+y^4} \qquad (x, y) \neq (0, 0)$
2. Show that the following functions *do* have a limit at the origin:
   a) $f(x, y) = \frac{x^2y^2}{x^2+y^2}$
   b) $f(x, y) = \frac{3x^5-xy^4}{x^4+y^4}$
3. Let
   $$f(x, y) = \frac{1}{x}\sin(xy) \qquad (x \neq 0) \tag{6.18}$$
   How should we define $f(0, y)$ for any number y so as to make f a continuous function on the plane?
4. Consider again the function
   $$f(x, y) = \frac{xy}{x^2 + y^2}\,. \tag{6.19}$$
   In Exercise 6.2 we have shown that f is not continuous at $(0, 0)$. Nonetheless, prove that $f(x, b)$, and $f(a, y)$ are continous functions of x, and y, respectively, for any numbers a, and b (including $a = 0$, $b = 0$).

   We say f is **separately continuous**.

Additional: Limits and Continuity

Further Reading

(Folland, *Advanced Calculus*, Chapter 1.3)

Vector Calculus: Advanced (MAST20032)

Functions of several variables

Definition 6.1 extends in an obvious way to functions of several variables: If f is a function on $\mathbb{R}^n$, then we say f has a limit l as $\vec{x} = (x^1, x^2, \ldots, x^n)$ approaches $\vec{a} = (a^1, \ldots, a^n)$ if for every $\varepsilon > 0$, we can find a $\delta > 0$ so that, if $0 < |\vec{x} - \vec{a}| < \delta$, then $|f(\vec{x}) - f(\vec{a})| < \varepsilon$.

Continuity of elementary functions of two variables

In our discussion of continuous functions of two variables we have stated without proof that:

Proposition 6.1. *The functions $f_1(x, y) = x + y$ and $f_2(x, y) = xy$ are continuous functions on the plane.*

This can be shown in an elementary way using the definition of continuity.
We have also used without proof that compositions preserve continuity.

Proposition 6.2. *Suppose $f : A \to \mathbb{R}^2$, and $g : B \to \mathbb{R}$ are continuous functions on their domains $A, B \subset \mathbb{R}^2$, respectively. If $f(A) \subset B$, then $g \circ f : A \to \mathbb{R}$ is a continuous function.*

This is used to obtain results about the continuity of elementary functions.

Proposition 6.3. *The functions $f_3(x, y) = x - y$ is continuous on the plane, and the function $f_4(x, y) = x/y$ is continuous on $\{(x, y) : y \neq 0\}$.*

Proof. We have $f_3(x, y) = f_1(x, f_2(-1, y))$, hence a composition of continuous functions. Moreover, $f_4(x, y) = f_2(x, g(y))$, where $g(y) = 1/y$ is continuous away from 0, hence f_2 is a composition of continuous functions on the set where it is defined. $\square$

Module III.

Differentiation

Note 7.

Differentiation in one variable

7.1. Differentiability in one variable

Following our intuition that continuous functions are those "whose graphs can be drawn without lifting the pen off the paper" they are still allowed to have "sharp corners". A differentiable function does not have graphs like that and admits a well-defined "tangent line" at each point.

Definition 7.1. A function is **differentiable at** a if

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \tag{7.1}$$

exists. In this case the limit is denoted by $f'(a)$ and is called the **derivative of** f **at** a**.**

Remark 7.1. Note that the *difference quotient* $(f(x+h)-f(x))/h$ is the slope of the line through the points $(x, f(x))$ and $(x+h, f(x+h))$. Therefore define that **tangent line** to the graph of f at $(a, f(a))$ to be the line through the point $(a, f(a))$ with slope $f'(a)$.

We say f is differentiable if f is differentiable at every point on its domain. More generally, we say f is differentiable on say an interval $A = (a, b)$ (or some set of points A) if f is differentiable at every point $a \in A$, and we call the function f' the **derivative** of f on the domain A.

Example 7.1. The constant function $f(x) = c$ is differentiable and $f'(x) = 0$.

Exercise 7.1. The **linear functions** $f(x) = cx + d$ are differentiable, and $f'(x) = c$.

Example 7.2. Let us compute the derivative of the function $f(x) = x^2$ at $x = a$:

$$f'(a) = \lim_{h\to 0} \frac{(a+h)^2 - a^2}{h} = \lim_{h\to 0}(2a+h) = 2a \tag{7.2}$$

Exercise 7.2. Show that the function $f(x) = x^3$ is differentiable and $f'(a) = 3a^2$.

Example 7.3. The function

$$f(x) = |x| \tag{7.3}$$

is *not* differentiable at 0. Indeed, the difference quotient at 0 is simply $|h|/h$, which is 1 for $h > 0$, and -1 for $h < 0$, so the limit as h approaches 0 does not exist.

Exercise 7.3. Show that the function $f(x) = |x|$ is differentiable at every point $a \neq 0$.

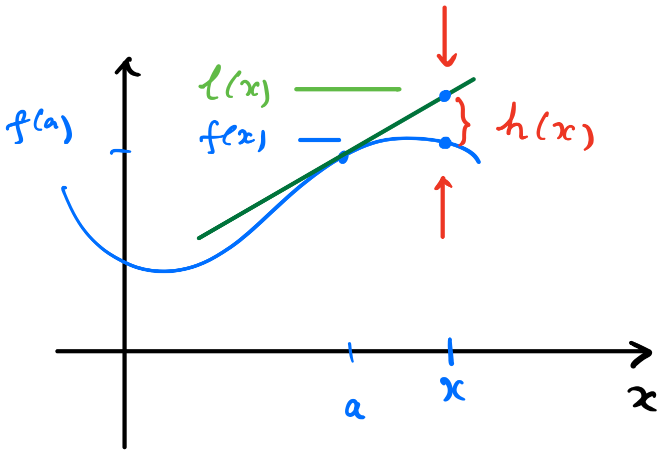

Figure 7.1.: Linear approximation of a function.

Example 7.4. The function $f(x) = \sqrt{|x|}$ is also not differentiable at 0. In fact, the slopes of the tangent lines at $(x, f(x))$ become infinite as we approach 0 from the right, and negative infinite as we approach from the left.

These are example of functions which are continuous, but not differentiable. Conversely, we have:

Theorem 7.1. *If f is differentiable at a, then f is continuous at a.*

Proof.

$$\lim_{h\to 0} f(a+h) - f(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \cdot h = f'(a) \lim_{h\to 0} h = 0\,. \tag{7.4}$$

□

7.1.1. Geometric interpretation of the derivative

Let us return to the geometric interpretation of the derivative given in Remark 7.1, cf. Figure 7.1.

Suppose f is differentiable at a point a. This means that the graph of the linear function

$$l(x) = f'(a)(x-a) + f(a) \tag{7.5}$$

approximates the graph of the function f near a. More precisely, we know that the difference in height

$$h(x) = f(x) - l(x) \tag{7.6}$$

tends to zero *at a faster rate than* $x - a$, as x approaches a:

$$\lim_{x\to a} \frac{h(x)}{x-a} = \lim_{x\to a} \frac{f(x) - l(x)}{x-a} = \lim_{x\to a} \frac{f(x) - f(a)}{x-a} - f'(a) = 0\,. \tag{7.7}$$

In conclusion, we have

$$f(x) = l(x) + h(x) = f(a) + f'(a)(x - a) + h(x)\,, \tag{7.8}$$

and we can view this as a *linear approximation of the function* f near a, and h as an *error* which goes to zero *faster* than the distance to a.

Remark 7.2. The statement that

$$\lim_{x\to a} h(x) = 0 \tag{7.9}$$

is the statment that $f(x)$ is *continuous* at a. The fact that (7.9) is implied by (7.7) is another proof of Theorem 7.1.

7.1.2. Higher derivatives

We have seen in Note 4 that the function

$$f(x) = \begin{cases} x\sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{7.10}$$

is continuous. However, once again this function is not differentiable at 0: For any $h \neq 0$,

$$\frac{f(h) - f(0)}{h} = \sin(1/h) \tag{7.11}$$

and this function does not have a limit as $h \to 0$. A very similar function, which *is* differentiable at 0, is

$$f(x) = \begin{cases} x^2\sin(1/x) & x \neq 0 \\ 0 & x = 0\,. \end{cases} \tag{7.12}$$

However, we will see that for this function the *second derivative* fails to exist at 0.

For any function f, we obtain by taking the derivative another function f' (whose domain may be smaller than the domain of f). Clearly, now starting with the function f', we obtain another function $(f')'$ whose domain we take to be all points where f' is differentiable. This is the **second derivative** f'' of f. In general, we also write

$$f^{(0)} = f \tag{7.13}$$

$$f^{(1)} = f' \tag{7.14}$$

$$f^{(2)} = f'' \tag{7.15}$$

$$f^{(k+1)} = (f^{(k)})'\,, \tag{7.16}$$

and we also call $f^{(k)}$, for $k \geq 2$, the **higher order derivatives** of f. The idea is that the more derivatives of a function exist, the more *regular* it is.

Example 7.5. Consider the function

$$f(x) = \begin{cases} x^2 & x \geq 0 \\ -x^2 & x \leq 0\,. \end{cases} \tag{7.17}$$

We know that $f'(a) = 2a$ for $a > 0$, and $f'(a) = -2a$ for $a < 0$. Moreover,

$$\Big|\frac{f(h) - f(0)}{h}\Big| = |h| \tag{7.18}$$

and so f is differentiable at 0, and $f'(0) = 0$. We can summarize this conveniently by

$$f'(x) = 2|x|\,. \tag{7.19}$$

As we have seen this function is continuous but not differentiable, hence $f''(0)$ does not exist.

7.2. Differentiation

The aim is now to prove a few theorems that will allow us to *differentiate* a large number of functions without invoking the definition, and investigating a limit, every time.

Theorem 7.2 (Sum and product rule)**.** *If f and g are differentiable at a, then $f + g$ and $f \cdot g$ are also differentiable at a, and*

$$(f + g)'(a) = f'(a) + g'(a) \tag{7.20}$$

$$(f \cdot g)'(a) = f'(a)g(a) + f(a)g'(a) \tag{7.21}$$

The second formula is also called the **product rule**.

Proof of (7.21)*.*

$$\begin{aligned}(fg)'(a) &= \lim_{h\to 0} \frac{(fg)(a+h) - (fg)(a)}{h} \\ &= \lim_{h\to 0} \frac{(f(a+h) - f(a))g(a+h) + f(a)(g(a+h) - g(a))}{h} \\ &= f'(a)g(a) + f(a)g'(a)\end{aligned} \tag{7.22}$$

where in the last step we have used Theorem 3.1. □

Another proof of (7.21)*.* We can also use the ideas from Section 7.1.1 to give another proof of the product rule: If f and g are differentiable at a, then they have a linear approximation near a,

$$f(x) = f(a) + f'(a)(x - a) + h(x)\,, \qquad g(x) = g(a) + g'(a)(x - a) + k(x)\,, \tag{7.23}$$

where $h(x)$, and $k(x)$ are functions that tend to zero faster than $x - a$. Therefore

$$f(x)g(x) = f(a)g(a) + \Big[f(a)g'(a) + f'(a)g(a)\Big](x - a) + E(x) \tag{7.24}$$

where $E(x)$ is a function that satisfies

$$\lim_{x\to a} \frac{E(x)}{x - a} = 0\,. \tag{7.25}$$

Therefore the graph of the function fg is approximated by a linear function with slope $f(a)g'(a) + f'(a)g(a)$. □

Example 7.6. If $f(x) = x^n$ for some natural number $n \in \mathbb{N}$, then we can now prove by induction that

$$f'(a) = na^{n-1}\,. \tag{7.26}$$

Exercise 7.4. The above allows us to compute easily the derivatives of polynomials:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0 \tag{7.27}$$

Compute a few derivatives, $p'(x)$, $p^{(2)}$, and note in particular that

$$f^{(n)}(x) = n!a_n\,, \qquad f^{(k)}(x) = 0 \qquad (k > n)\,. \tag{7.28}$$

Note that the product rule can also be used to differentiate any product of functions. For example to compute $(f \cdot g \cdot h)'$, we could either write $f \cdot g \cdot h = (fg) \cdot h$, and apply the product rule to the functions fg, and h, or we could write $f \cdot g \cdot h = f \cdot (gh)$, and apply the product rule to f, and gh; both has the same result:

$$(fgh)' = (f')gh + f(g')h + fg(h') \tag{7.29}$$

Next we turn to the **quotient rule**.

Theorem 7.3 (Quotient rule)**.** *If f and g are differentiable at a and $g(a) \neq 0$, then f/g is differentiable at a, and*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{(g(a))^2} \tag{7.30}$$

Proof. Let us start with the special case

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{(g(a))^2}\,. \tag{7.31}$$

We have

$$\begin{aligned}\left(\frac{1}{g}\right)'(a) &= \lim_{h\to 0}\frac{1}{h}\frac{g(a) - g(a+h)}{g(a+h)g(a)}\\ &= -\lim_{h\to 0}\frac{g(a+h) - g(a)}{h}\lim_{h\to 0}\frac{1}{g(a+h)g(a)} = -\frac{g'(a)}{(g(a))^2}\,.\end{aligned} \tag{7.32}$$

Note that for the difference quotient to make sense we needed $g(a+h) \neq 0$, at least for $|h|$ sufficiently small. However, we know that g is differentiable, hence continuous at a, and since $g(a) \neq 0$ it follows that $g(a+h) \neq 0$, as long as $|h| < \delta$, for some $\delta > 0$; cf. Theorem 7.1 and Theorem 4.3.

Since

$$\frac{f}{g} = f \cdot \frac{1}{g}\,, \tag{7.33}$$

the formula for the derivative of the quotient then follows from the product rule. □

Example 7.7.

$$f(x) = \frac{x^2-1}{x^2+1} \qquad f'(x) = \frac{4x}{(x^2+1)^2} \tag{7.34}$$

Example 7.8. If $f(x) = x^{-n}$ for some natural number $n \in \mathbb{N}$, then

$$f(x) = \frac{1}{x^n}\,, \qquad f'(x) = \frac{-nx^{n-1}}{x^{2n}} = (-n)x^{-n-1}\,. \tag{7.35}$$

To investigate our favourite functions from above, say

$$f(x) = x^2 \sin(1/x) \qquad (x \neq 0) \tag{7.36}$$

we need to understand differentiation of compositions, and we need to know the derivatives of the trigonometric functions which we take for now without proof:

$$\sin'(x) = \cos(x) \qquad \cos'(x) = -\sin'(x)\,. \tag{7.37}$$

Finally the **chain rule**, whose proof can defer to Additional.

Theorem 7.4 (Chain rule)**.** *If g is differentiable at a, and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at a, and*

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a)\,. \tag{7.38}$$

Example 7.9. Consider again the function $f(x)$ from (7.36), or more precisely the extension

$$f(x) = \begin{cases} x^2 \sin(1/x) & x \neq 0 \\ 0 & x = 0\,. \end{cases} \tag{7.39}$$

We have already seen, directly from the defintion, that $f'(0) = 0$. With the chain rule, we can also compute that for $x \neq 0$,

$$f'(x) = 2x\sin(1/x) + x^2\cos(1/x) \cdot \Big(-\frac{1}{x^2}\Big) = 2x\sin(1/x) - \cos(1/x) \tag{7.40}$$

In particular, $f'(x)$ is *not* continuous at 0.

We illustrate how the chain rule is applied in practice with some more examples.

Example 7.10.

$$f(x) = \sin(x^2) \tag{7.41}$$

When we apply the chain rule we view this as a composition

$$f = \sin \circ S \tag{7.42}$$

where $S(x) = x^2$ (say S for “taking the square”). Then it is clear that

$$f'(x) = \cos(S(x)) \cdot S'(x) = \cos(x^2) \cdot 2x\,, \qquad S'(x) = 2x\,. \tag{7.43}$$

Example 7.11.

$$f(x) = \sin^2(x^2) \tag{7.44}$$

We could view f as the composition

$$f = S \circ (\sin \circ S) \tag{7.45}$$

and the chain rule gives

$$f' = S'(\sin \circ S) \cdot (\sin \circ S)' \tag{7.46}$$

$$f'(x) = 2(\sin \circ S)(x) \cdot (\cos \circ S)(x) \cdot 2x = 2\sin(x^2) \cdot \cos(x^2) \cdot 2x \tag{7.47}$$

The above notation is useful to clarify the compositions that make up a function, but in practice one does not usually introduce additional notation.

Example 7.12.

$$f(x) = \sin(\sin(x^2)) \tag{7.48}$$

We compute directly:

$$f'(x) = \cos(\sin(x^2)) \cdot \cos(x^2) \cdot 2x \tag{7.49}$$

Exercise 7.5. Compute the derivatives of

$$f(x) = \sin\Big((\sin(x))^2\Big) \tag{7.50}$$

$$f(x) = \sin^2(x \sin x) \tag{7.51}$$

Problems

1. Prove directly using the defintion, that if $f(x) = \sqrt{x}$, then $f'(a) = \frac{1}{2\sqrt{a}}$ for any $a > 0$.

2. Prove that if $g(x) = f(x) + c$, then $g'(x) = f'(x)$. Also show that if $g(x) = cf(x)$ then $g'(x) = cf'(x)$.

3. Let f be a function such that $|f(x)| \leq x^2$ for all x. Prove that f is differentiable at 0.

4. Show that if f is differentiable at a, then

   a)
   $$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \tag{7.52}$$

   b)
   $$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a-h)}{2h} \tag{7.53}$$

5. a) Prove that if f is even, then f' is odd:

$$f'(x) = -f'(-x)\,. \tag{7.54}$$

   Draw a picture!

   b) Prove that if f is odd, then f' is even: $f'(x) = f'(-x)$.

6. Find the derivative f' for the following functions f.
   a) $f(x) = \sin(\cos(x))$
   b) $f(x) = \sin(x + \sin(x))$

7. Find f' for each of the following functions f.
   a) $f(x) = \sin^3(x^2 + \sin(x))$
   b) $f(x) = \sin^2((x + \sin(x))^2)$
   c) $f(x) = (x + \sin^5(x))^6$
   d) $f(x) = \frac{\sin(x^2)\sin^2(x)}{1+\sin(x)} \qquad (x \neq -\pi/2 + k2\pi, k \in \mathbb{Z})$

8. Find the derivative f' for the following functions f.
   a) $f(x) = \sin(x + x^2)$
   b) $f(x) = \sin\left(\frac{\cos(x)}{x}\right) \qquad (x \neq 0)$

9. Find f' for each of the following functions f.
   a) $f(x) = \sin^3(x^2 + \sin(x))$
   b) $f(x) = \sin^2((x + \sin(x))^2)$
   c) $f(x) = (x + \sin^5(x))^6$
   d) $f(x) = \frac{\sin(x^2)\sin^2(x)}{1+\sin(x)} \qquad (x \neq -\pi/2 + k2\pi, k \in \mathbb{Z})$

Note 8.

Differentiation in two variables

8.1. Partial derivatives

The simplest notion of a derivative of a function $f(x, y)$ of two variables is the *partial derivative*, which is the derivative of the function with respect to *one* of the variables, while keeping the other fixed.

Definition 8.1 (Partial derivatives)**.** For a given function $f(x, y)$, the limit of the following difference quotients, if they exists, are called the *partial derivative* $\partial_x f$, and $\partial_y f$ of f at (a, b):

$$\frac{\partial f}{\partial x}(a,b) = \lim_{h\to 0} \frac{f(a+h,b)-f(a,b)}{h} \tag{8.1}$$

$$\frac{\partial f}{\partial y}(a,b) = \lim_{h\to 0} \frac{f(a,b+h)-f(a,b)}{h}\,. \tag{8.2}$$

Example 8.1. The partial derivatives of a function

$$f(x,y) = \frac{e^{2x}\sin(y)}{1+y^2} \tag{8.3}$$

are

$$\partial_x f(x,y) = \frac{2e^{2x}\sin(y)}{1+y^2} = 2f(x,y) \tag{8.4}$$

$$\partial_y f(x,y) = \frac{e^{2x}\cos(y)}{1+y^2} - \frac{e^{2x}2y\sin(y)}{(1+y^2)^2} = \left(\cos(y) + \frac{2y\sin(y)}{1+y^2}\right)\frac{e^{2x}}{1+y^2}\,. \tag{8.5}$$

The partial derivatives of a function tell us how the values of a function change along the coordinate axes. However, even if they exist at a point (a, b), they do not necessarily give us information about the behaviour of the function near (a, b).

Example 8.2. Let us take another look at the example

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases} \tag{8.6}$$

Here both partial derivatives exist at the origin:

$$\partial_x f(0,0) = \lim_{h\to 0} h^{-1}(f(h,0) - f(0,0)) = 0 \tag{8.7}$$

$$\partial_y f(0,0) = \lim_{h\to 0} h^{-1}(f(0,h) - f(0,0)) = 0 \tag{8.8}$$

However, this does not tell us anything about the behaviour of the function $f(x,y)$ near (a,b). The reason is, as we shall see, that this function is not differentiable at $(0,0)$.

8.2. Differentiable functions

The idea to characterize differentiable functions of two variables is the same as for functions of one variable: Functions which are differentiable at a point should be well approximated by *linear* functions near that point. In other words, the graph of a function which is differentiable at a point should be approximately a plane through that point.

More precisely, a function $f(\vec{x})$ of two variables $\vec{x} = (x_1, x_2)$ is differentiable at a point $\vec{a} = (a_1, a_2)$, if there is a linear function $l(\vec{x})$ such that $l(\vec{a}) = f(\vec{a})$, and the difference $h(\vec{x}) = f(\vec{x}) - l(\vec{x})$ tends to zero *faster* than $|\vec{x} - \vec{a}|$, as $\vec{x}$ approaches $\vec{a}$. Now, a *linear* function of two variables takes the form

$$l(\vec{x}) = b + \vec{c}\cdot\vec{x}\,, \tag{8.9}$$

for some $b \in \mathbb{R}$, $\vec{c} = (c_1, c_2) \in \mathbb{R}^2$, and the condition $l(\vec{a}) = f(\vec{a})$ implies that $b = f(\vec{a}) - \vec{c}\cdot\vec{a}$, so

$$l(\vec{x}) = f(\vec{a}) + \vec{c}\cdot(\vec{x} - \vec{a})\,. \tag{8.10}$$

Definition 8.2 (Differentiability)**.** A function $f(x_1, x_2)$ is *differentiable* at a point $\vec{a} = (a_1, a_2)$, if there is a vector $\vec{c} \in \mathbb{R}^2$ such that

$$f(\vec{a} + \vec{h}) = f(\vec{a}) + \vec{c}\cdot\vec{h} + E(\vec{h}) \tag{8.11}$$

where the function $E(\vec{h})$ satisfies

$$\lim_{\vec{h}\to 0} \frac{E(\vec{h})}{|\vec{h}|} = 0\,. \tag{8.12}$$

As for functions of one variable, all differentiable functions are continuous.

Theorem 8.1. *If f is differentiable at $\vec{a}$, then f is continuous at $\vec{a}$.*

The proof follows immediately from the linear approximation of a differentiable function. Indeed if f is differentiable at $\vec{a}$, then

$$f(\vec{a} + \vec{h}) = f(\vec{a}) + \nabla f(\vec{a})\cdot\vec{h} + E(\vec{h})\,. \tag{8.13}$$

where $E(\vec{h})/|\vec{h}| \to 0$, hence in particular

$$\lim_{\vec{h}\to 0} f(\vec{a} + \vec{h}) = f(\vec{a})\,. \tag{8.14}$$

8.3. Gradient of a function

We have seen in Definition 8.2 that if a function $f(\vec{x})$ is differentibale at a point $\vec{a}$, then there exists a vector $\vec{c}$, so that

$$f(\vec{x}) = f(\vec{a}) + \vec{c} \cdot (\vec{x} - \vec{a}) + E(\vec{x}) \tag{8.15}$$

where $E(\vec{x})/|\vec{x} - \vec{a}| \to 0$ as $\vec{x} \to \vec{a}$. Let us now determine the components of this vector $\vec{c} = (c_1, c_2)$. In (8.15) simply choose $\vec{x} = (a_1 + h, a_2)$, then

$$f(a_1 + h, a_2) - f(a_1, a_2) = \vec{c} \cdot (h, 0) + E(a_1 + h, a_2) = c_1 h + E(a_1 + h, a_2) \tag{8.16}$$

Hence, dividing by $h \neq 0$, and taking the limit $h \to 0$, shows that

$$\frac{\partial f}{\partial x}(\vec{a}) = c_1 \,. \tag{8.17}$$

Similarly for the other component of $\vec{c}$, and we conclude

$$\nabla f(\vec{a}) = \begin{pmatrix} \dfrac{\partial f}{\partial x}(\vec{a}) \\[2ex] \dfrac{\partial f}{\partial y}(\vec{a}) \end{pmatrix} . \tag{8.18}$$

We call this vector the **gradient** of f at (a, b).

In particular, we have proven that for a differentiable function the partial derivatives always exist.

Theorem 8.2. *If $f(x, y)$ is differentiable at (a, b), then the partial derivatives $\partial_x f$, and $\partial_y f$ of f exist at $\vec{a}$ and are the components of the gradient vector $\nabla f(\vec{a})$.*

Conversely, however, we have seen an example of a function whose partial derivatives exist at a point, but which is nonetheless *not* differentiable at that point:

Example 8.3. The function

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases} \tag{8.19}$$

is not continuous at the origin, hence in particular not differentiable at $(0, 0)$, by Theorem 8.1. Nonetheless the partial derivatives exist:

$$\partial_x f(0, 0) = \partial_y f(0, 0) = 0 \,. \tag{8.20}$$

In summary, for a function $f(x, y)$ to be differentiable at (a, b) it is *necessary* for the partial derivatives $\partial_x f(a, b)$, $\partial_y f(a, b)$ to exist, *but not sufficient.*

A sufficient criterion for differentiability is the following, whose proof we will defer:

Theorem 8.3. *Suppose the partial derivatives $\partial_x f$, and $\partial_y f$ of a function $f(x, y)$ exist at every point, and are themselves as functions $\partial_x f(x, y)$, and $\partial_y f(x, y)$* continuous *at (a, b). Then $f(x, y)$ is differentiable at (a, b).*

Example 8.4. For the example above we compute away from the origin,

$$\partial_x f(x, y) = \frac{y^3 - x^2 y}{(x^2 + y^2)^2} \qquad \partial_y f(x, y) = \frac{x^3 - xy^2}{(x^2 + y^2)^2} \qquad (x, y) \neq (0, 0)\,, \tag{8.21}$$

and these functions are continuous everywhere, *except* at the origin.

8.4. Directional derivatives

Let $\vec{a} = (a_1, a_2)$ be a point, and $\vec{u} = (u_1, u_2)$ a direction, namely a vector of unit length, $|u| = 1$. Then for each number t,

$$\vec{x}(t) = \vec{a} + t\vec{u} \tag{8.22}$$

is a point on the straight line through $\vec{a}$, in the direction $\vec{u}$.

Definition 8.3 (Directional derivative)**.** The *directional derivative* of a function f at $\vec{a}$ in the direction $\vec{u}$ is defined as the derivative of the function $f(\vec{a} + t\vec{u})$ at $t = 0$. If it exists, it is denoted by

$$\frac{\partial f}{\partial \vec{u}}(\vec{a}) = \lim_{t \to 0} \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}\,. \tag{8.23}$$

Remark 8.1. If we think of $\vec{x}(t) = \vec{a} + t\vec{u}$ as a particle moving on a straight line, and the function f as an observable, say the pressure in the atmosphere, then $\partial_{\vec{u}} f(\vec{a})$ is the rate of change of the pressure in the direction that the particle is moving.

Remark 8.2. Note that if we choose

$$\vec{u} = (1, 0) \tag{8.24}$$

then

$$\frac{\partial f}{\partial \vec{u}} = \frac{\partial f}{\partial x} \tag{8.25}$$

is the partial derivative of f in x. Similarly $\partial_y f$ is the directional derivative in the direction $\vec{u} = (0, 1)$.

Theorem 8.4 (Formula for the directional derivative)**.** *Suppose f is differentiable at $\vec{a}$. Then the directional derivative at $\vec{a}$ in any direction $\vec{u}$ exists, and is given by*

$$\frac{\partial f}{\partial \vec{u}}(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u}\,. \tag{8.26}$$

Proof. Since f is differentiable at $\vec{a}$, we know that (8.13) holds for any $\vec{h}$, in particular for $\vec{h} = t\vec{u}$:

$$f(\vec{a} + t\vec{u}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot t\vec{u} + E(\vec{a} + t\vec{u})\,. \tag{8.27}$$

Hence

$$\frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t} = \nabla f(\vec{a}) \cdot \vec{u} + \frac{E(\vec{a} + t\vec{u})}{t}\,, \tag{8.28}$$

and the formula follows by taking the limit $t \to 0$. □

The above formula also provides a geometric interpretation of the gradient of a function: Since for any vectors $\vec{a}$, and $\vec{b}$,

$$|\vec{a} \cdot \vec{b}| \leq |\vec{a}||\vec{b}| \tag{8.29}$$

with equality when $\vec{a}$ and $\vec{b}$ are colinear, we have in particular with $|\vec{u}| = 1$ that

$$\Big|\frac{\partial f}{\partial \vec{u}}(\vec{a})\Big| \leq |\nabla f(\vec{a})| \tag{8.30}$$

with equality when $\nabla f(\vec{a})$ and $\vec{u}$ are colinear. This means that $\nabla f(\vec{a})$ points in the direction of the steepest increase of f at $\vec{a}$, and its magnitude is that rate of increase of f in that direction.

Exercise 8.1. Let $f(x, y) = x^2 + 5xy^2$, and $\vec{a} = (-2, 1)$.

1. Find the directional derivative of f at $\vec{a}$ in the direction of the vector $\vec{v} = (3, 4)$.

2. What is the largest directional derivative of f at $\vec{a}$, and in what direction does it occur?

Solution 8.2. We have $\nabla f(x, y) = (2x + 5y^2, 10xy)$, so that $\nabla f(-2, 1) = (1, -20)$.

Note that $\vec{v}$ is not normalised, so let us first determine $\vec{u}$ colinear to $\vec{v}$ of unit length:

$$\vec{u} = \frac{v}{|\vec{v}|} = \frac{1}{5}(3, 4)\,. \tag{8.31}$$

The directional derivative in that direction is

$$\partial_{\vec{u}} f(-2, 1) = \nabla f(-2, 1) \cdot \vec{u} = \frac{3 - 80}{5} = -\frac{77}{5}\,. \tag{8.32}$$

The largest directional derivative is

$$|\nabla f(\vec{a})| = \sqrt{401}\,,$$

and it occurs in the direction $(1, -20)/\sqrt{401}$.

Problems

1. For each of the following functions $f(x,y)$ find the linear function $l(x,y)$ whose graph is the tangent plane to the graph of f at the point $(1,-2,f(1,-2))$.
   a) $f(x,y) = x^2y + \sin(\pi xy)$
   b) $f(x,y) = \frac{xy}{x^2+y^2}$

2. Compute the gradient $\nabla f(\vec{x})$ of the following functions.
   a) $f(\vec{x}) = \frac{1}{|\vec{x}|} \qquad (\vec{x} \neq 0)$
   b) $f(\vec{x}) = (\vec{a} \cdot \vec{x})^2$

3. For each of the following functions $f(x,y)$ compute the directional derivative of f at the point $(-1,2)$ in the direction $(\frac{3}{5}, \frac{4}{5})$.
   a) $f(x,y) = x^2y + \sin(\pi xy)$
   b) $f(x,y) = \frac{xy}{x^2+y^2}$

4. Suppose $f(\vec{x})$ and $g(\vec{x})$ are differentiable at $\vec{a}$. Does that imply that $f+g$, and fg are differentiable at $\vec{a}$? Find a formula for the gradient of fg, and $f+g$ at $\vec{a}$.

Additional: Chain rule

Further Reading

(Spivak, *Calculus*, Chapter 9, 10) (Folland, *Advanced Calculus*, Chapter 2.3)

Real Analysis: Advanced (MAST20033)Vector Calculus: Advanced (MAST20032)

Functions of one variable

Let us first give a direct proof of the chain rule.

Proof of Theorem 7.4. The idea is write the difference quotient as

$$\frac{f(g(a+h))-f(g(a))}{h}=\frac{f(g(a+h))-f(g(a))}{g(a+h)-g(a)}\frac{g(a+h)-g(a)}{h} \tag{8.1}$$

which however we can only do when $g(a+h)-g(a)\neq 0$. This can fail even for simple choices of g (say g is the constant function), so define instead:

$$\phi(h)=\begin{cases}\frac{f(g(a+h))-f(g(a))}{g(a+h)-g(a)} & \text{, when } g(a+h)-g(a)\neq 0\\ f'(g(a)) & \text{, when } g(a+h)-g(a)=0\end{cases} \tag{8.2}$$

We want to show that

$$\lim_{h\to 0}\phi(h)=f'(g(a))\,. \tag{$*$}$$

With this statement we can complete the proof, because

$$\frac{f(g(a+h))-f(g(a))}{h}=\phi(h)\cdot\frac{g(a+h)-g(a)}{h} \tag{8.3}$$

even if $g(a+h)-g(a)=0$ (in which case both sides are zero), and taking the limit then implies the formula.

It remains to show $(*)$, namely that for every $\varepsilon>0$ there exists a $\delta>0$ such that, if $0<|h|<\delta$, then

$$|\phi(h)-f'(g(a))|<\varepsilon. \tag{8.4}$$

So let $\varepsilon>0$. Since f is differentiable at $g(a)$, we can certainly find a $\eta>0$ such that, if $|k|<\eta$, then

$$\Big|\frac{f(g(a)+k)-f(g(a))}{k}-f'(g(a))\Big|<\varepsilon\,. \tag{8.5}$$

Moreover since g is differentiable at a, hence continuous at a, we can find $\delta > 0$, so that, if $|h| < \delta$, then

$$|g(a+h) - g(a)| < \eta. \tag{8.6}$$

Hence whenever $|h| < \delta$, and $k = g(a+h) - g(a) \neq 0$, we have

$$\begin{aligned} \Big|\phi(h) - f'(g(a))\Big| &= \Big|\frac{f(g(a+h)) - f(g(a))}{g(a+h) - g(a)} - f'(g(a))\Big| \\ &= \Big|\frac{f(g(a)+k) - f(g(a))}{k} - f'(g(a))\Big| < \varepsilon \end{aligned} \tag{8.7}$$

and the same statement, $|\phi(h) - f'(g(a))| < \varepsilon$, is obviously also true for all h with $k = g(a+h) - g(a) = 0$, in view of the definition of $\phi(h)$. This proves $(*)$.

□

Functions of several variables

The chain rule is also related to the directional derivative: In (8.22) we have "parametrized" the points on a straight line in $\mathbb{R}^2$.

More generally, the equation

$$\vec{x}(t) = \vec{g}(t) = (g_1(t), \ldots, g_n(t)) \tag{8.8}$$

where $g_i : i = 1, 2, \ldots, n$ are functions of t, represents a **parametrized curve** in $\mathbb{R}^n$.

Remark 8.1. For $n = 3$, we can view $\vec{x} = \vec{g}(t)$ as the trajectory of a particle, which at time t has the position $\vec{g}(t)$. Then

$$\vec{g}'(t) = (g_1'(t), g_2'(t), g_3'(t)) \tag{8.9}$$

is the velocity of the particle at time t.

Theorem 8.1 (Chain rule)**.** *Suppose that $\vec{g}(t) = (g_1(t), \ldots, g_n(t))$, and $g_i : i = 1, 2, \ldots, n$ are differentiable functions at t = a. Suppose moreover that $f(x_1, \ldots, x_n)$ is differentiable at $\vec{b} = \vec{g}(a)$. Then the function $(f \circ g)(t)$ is differentiable at t = a, and its derivative is given by*

$$(f \circ g)'(a) = \nabla f(\vec{b}) \cdot \vec{g}'(a)\,. \tag{8.10}$$

Proof. Since f is differentiable at $\vec{g}(a)$, we have

$$\begin{aligned} f(\vec{g}(a+u)) =& f(\vec{g}(a+u) - \vec{g}(a) + \vec{g}(a)) \\ =& f(\vec{g}(a)) + \nabla f(\vec{g}(a)) \cdot \vec{h} + E_1(\vec{h})\,, \end{aligned} \tag{8.11}$$

where

$$\begin{aligned} \vec{h} =& \vec{g}(a+u) - \vec{g}(a) \\ =& \vec{g}'(a)u + \vec{E}_2(u)\,, \qquad \vec{g}'(a) = (g_1'(a), \ldots, g_n'(a))\,, \end{aligned} \tag{8.12}$$

because each $g_i(t) : i = 1, \ldots, n$ is differentiable at $t = a$. Hence

$$f(\vec{g}(a+u)) = f(\vec{g}(a)) + \nabla f(\vec{g}(a)) \cdot \vec{g}'(a)u + \nabla f(\vec{g}(a)) \cdot \vec{E}_2(u) + E(\vec{h}), \tag{8.13}$$

and the chain rule follows, because the last two terms go to zero as u tends to zero faster than u. $\square$

Exercise 8.1. Use this approach to give another proof of the chain rule for functions of one variable.

Module IV.

Mean value theorem

Note 9.

L'Hôpital's rule

The aim of this lecture is to prove:

Theorem 9.1 (L'Hôpital's rule)**.** *Suppose that*

$$\lim_{x\to a} f(x) = 0 \qquad \lim_{x\to a} g(x) = 0 \tag{9.1}$$

and suppose that $\lim_{x\to a} f'(x)/g'(x)$ *exists. Then* $\lim_{x\to a} f(x)/g(x)$ *exists, and*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} \,. \tag{9.2}$$

There are many variations of L'Hôpital's rule; see for example Problems 6 below.

The proof in turn relies on the **mean value theorem**, which says that given a continuous function on $[a, b]$, which is differentiable on (a, b), there is a some $x \in (a, b)$ such that

$$f'(x) = \frac{f(b) - f(a)}{b - a} \,. \tag{9.3}$$

Geometrically, this means that there is *some tangent line* to the graph of f, which is paralllel to the line between $(a, f(a))$ and $(b, f(b))$; see Figure 9.1.

Before we discuss these concepts an example how L'Hôptial's rule is applied:

Example 9.1*.* The theorem allows us to determine the limit of the function

$$f(x) = \frac{\sin(x)}{x} \qquad (x \neq 0) \tag{9.4}$$

near $a = 0$. Since $\sin'(x) = \cos(x)$, and $x' = 1$, and moreover

$$\lim_{x\to 0} \frac{\cos(x)}{1} = 1 \,, \tag{9.5}$$

the assumptions of L'Hôptial's rule are satisfied, and we conclude that

$$\lim_{x\to 0} \frac{\sin(x)}{x} = \lim_{x\to 0} \frac{\cos(x)}{1} = 1 \,. \tag{9.6}$$

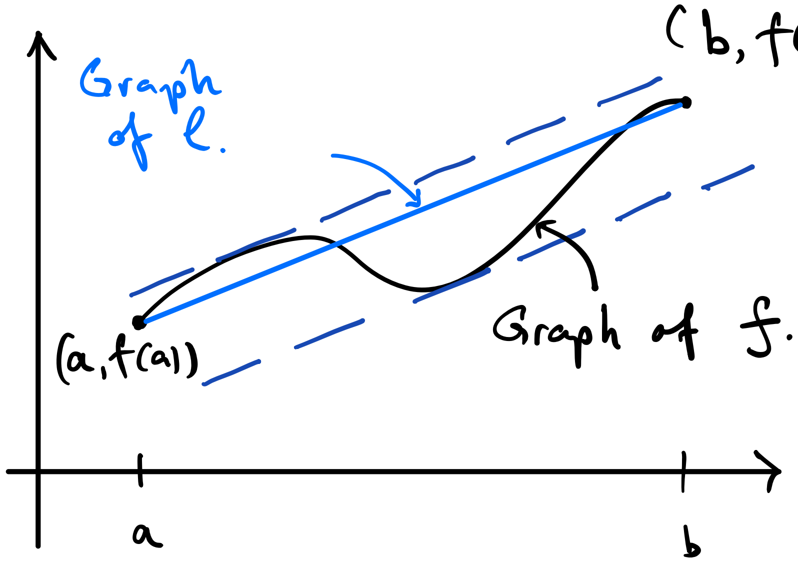

Figure 9.1.: Geometric interpretation of the mean value theorem.

9.1. Mean Value Theorem

The theorems about derivatives we have seen in the previous lectures give us some information about the derivative in terms of the function itself. It is more difficult to infer knowledge of a function in terms of its derivative. For example, it is easy to see that $f'(x) = 0$ for the constant function $f(x) = c$, but how do you prove that if $f'(x) = 0$ then f must be constant?

The mean value theorem allows us to draw this conclusion. It is probably the most important theorem about derivatives and has many consequences. We first discuss a special case:

Theorem 9.2 (Rolle's Theorem). *If f is continuous on $[a, b]$ and differentiable on (a, b), and $f(a) = f(b)$, then there is a number x in (a, b) such that $f'(x) = 0$.*

Proof. Since f is continuous on $[a, b]$ it has a maximum point, and a minimum point in $[a, b]$.

If the maximum, or minimum, occur at $x \in (a, b)$, then $f'(x) = 0$ at this point. More precisely, we are using here that if $x \in (a, b)$ is a *local* maximum (or minimum), then necessarily $f'(x) = 0$. (See Additional for a proof of this statement.)

If both the maximum *and* minimum points lie on the boundary, then since $f(a) = f(b)$ they must be equal, and the function f is a constant, hence $f'(x) = 0$ for any $x \in (a, b)$. □

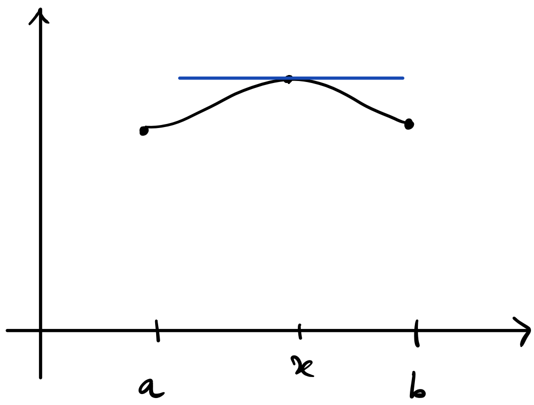

Figure 9.2.: Geometric interpretation of Rolle's theorem

Exercise 9.1. Draw the graph of a continuous on $[a, b]$, but not differentiable function on (a, b) for which the conclusion of Rolle's theorem is false.

Theorem 9.3 (Mean Value Theorem)**.** *If f is continuous on $[a, b]$ and differentiable on (a, b), then there is a number x in (a, b) such that*

$$f'(x) = \frac{f(b) - f(a)}{b - a}\,. \tag{9.7}$$

Proof. Recall the geometric interpretation of this statement: At the point x the slope of the tangent equals that of the line from $(a, f(a))$ to $(b, f(b))$.

Now the line from $(a, f(a))$ to $(b, f(b))$ is the graph of the linear function

$$l(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(a)\,. \tag{9.8}$$

In particular $l(a) = f(a)$, and $l(b) = f(b)$ and we set

$$h(x) = f(x) - l(x) \tag{9.9}$$

which is the height of the graph of f over the line from $(a, f(a))$ to $(b, f(b))$. We have

$$h(a) = 0\,, \qquad h(b) = 0\,. \tag{9.10}$$

Thus by Rolle's theorem there is an x in (a, b) such that

$$h'(x) = f'(x) - l'(x) = f'(x) - \frac{f(b) - f(a)}{b - a} = 0\,. \tag{9.11}$$

□

We can now prove the obvious

Corollary 9.4. *If f is defined on an interval and $f'(x) = 0$ for all x in the interval, then f is constant on the interval.*

Proof. Take any points $a < b$ in that interval, then by the mean value theorem there is an $x \in (a, b)$ with $f'(x)(b - a) = f(b) - f(a)$, but $f'(x) = 0$ so $f(a) = f(b)$. □

Exercise 9.2. If f and g are defined on the same interval and $f'(x) = g'(x)$ for all x, show that then there is a constant c such that $f(x) = g(x) + c$.

9.2. L'Hôpital's rule

We now derive several consequences of the mean value theorem.

Theorem 9.5 (Cauchy mean value theorem)**.** *Let f and g be continuous on $[a, b]$ and differentiable on (a, b), then there is a number $x \in (a, b)$ such that*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x)}{g'(x)} \tag{9.12}$$

provided $g(b) \neq g(a)$ and $g'(x) \neq 0$.

Remark 9.1. Note that the special case $g(x) = x$ is the mean value theorem, but the Cauchy mean value theorem is not a direct consequence of the mean value theorem, because while $f(b) - f(a) = f'(x)(b - a)$ for some x, and $g(b) - g(a) = g'(y)(b - a)$ for some y, x and y are not neccesarily the same.

Proof. Let

$$h(x) = \big(f(b) - f(a)\big)g(x) - f(x)\big(g(b) - g(a)\big) \tag{9.13}$$

then h is continuous on $[a, b]$ and differentiable on (a, b), and

$$h(a) = f(b)g(a) - f(a)g(b) = h(b) \tag{9.14}$$

so by Rolle's theorem, there is a number x in (a, b) such that $h'(x) = 0$. □

This theorem is the main statement we need to evaluate limits of the form

$$\lim_{x \to a} \frac{f(x)}{g(x)} \tag{9.15}$$

when $\lim_{x \to a} f(x) = 0$ and $\lim_{x \to a} g(x) = 0$.

Proof of Theorem 9.1. Recall that we assume that f and g approach the limit 0 near a, so let us define (possibly *redefine*)

$$f(a) = 0\,, \qquad g(a) = 0\,, \tag{9.16}$$

then f and g are continuous at a. Then by the Cauchy mean value theorem applied to f and g on the interval $[a, x]$, we get that there exists $a < \alpha_x < x$, such that

$$\frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\alpha_x)}{g'(\alpha_x)} \tag{9.17}$$

Note that the assumptions of Theorem 9.5 are indeed satisfied: Since the $f'(x)/g'(x)$ approaches a limit, and so in particular $g'(x) \neq 0$ near a. This also shows that $g(x) \neq 0$ near a, because if $g(x) = 0$ for some $x > a$, then by the mean value theorem there would exist a $y \in (a, x)$ with $g'(y) = 0$, again contradicting that $g'(x) \neq 0$ near a.

Furthermore it follows that

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(\alpha_x)}{g'(\alpha_x)}, \tag{9.18}$$

because α_x tends to 0 as x goes to 0. More precisely, we know that the limit

$$l = \lim_{x \to a} \frac{f'(x)}{g'(x)} \tag{9.19}$$

exists. This means that for any $\varepsilon > 0$, we can find $\delta > 0$, so that, if $|y - a| < \delta$, then

$$\Big| \frac{f'(y)}{g'(y)} - l \Big| < \varepsilon\,. \tag{9.20}$$

Therefore, whenever $|x - a| < \delta$,

$$\Big| \frac{f(x)}{g(x)} - l \Big| = \Big| \frac{f'(\alpha_x)}{g'(\alpha_x)} - l \Big| < \varepsilon \tag{9.21}$$

because $\alpha_x \in (a, x)$, and so in particular $|\alpha_x - a| < |x - a| < \delta$.

□

Problems

1. A function is **increasing** on an interval if $f(a) < f(b)$ whenever a and b are two numbers in the interval with $a < b$. Similarly for a **decreasing** function.

   Show that if $f'(x) > 0$ for all x in an interval, then f is increasing on the interval.

2. a) Prove that if $f'(x) \geq M$ for all x in $[a, b]$ then $f(b) \geq f(a) + M(b - a)$
   b) Prove that if $f'(x) \leq M$ for all x in $[a, b]$ then $f(b) \leq f(a) + M(b - a)$
   c) Formulate a similar theorem when $|f'(x)| \leq M$ for all x in $[a, b]$

3. a) Suppose that $f'(x) > g'(x)$ for all x and that $f(a) = g(a)$. Show that $f(x) > g(x)$ for $x > a$ and $f(x) < g(x)$ for $x < a$

b) Show by an example that these conclusions do not follow without the hypothesis $f(a) = g(a)$.

4. Find all functions f such that

   a) $f'(x) = \sin(x)$

   b) $f''(x) = x^3$.

5. What is wrong with the following use of L'Hôptial's rule:

$$\lim_{x\to 1} \frac{x^3 + x - 2}{x^2 - 3x + 2} = \lim_{x\to 1} \frac{3x^2 + 1}{2x - 3} = \lim_{x\to 1} \frac{6x}{2} = 3 \tag{9.22}$$

   Find the correct limit.

6. Prove the following variations of L'Hôpital's rule (with the much the same reasoning as in the proof of Theorem 9.1).

   a) If $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$, and $\lim_{x\to a} f'(x)/g'(x) = \infty$, then

$$\lim_{x\to a} f(x)/g(x) = \infty\,.$$

   b) If $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = 0$, and $\lim_{x\to\infty} f'(x)/g'(x) = l$, then $\lim_{x\to\infty} f(x)/g(x) = l$.

   Hint: Consider $\lim_{x\to 0^+} f(1/x)/g(1/x)$.

Note 10.

Inverse functions

The graph of f^{-1} is the graph of f reflected across the diagonal line consisting of all points (x, x). For f^{-1} to be a function it is necessary that, geometrically, no horizontal line intersects the graph of f twice, and this property has a name:

Definition 10.1. A function f is **one-to-one** if $f(a) \neq f(b)$ if $a \neq b$.

Example 10.1. Say $f(x) = x^3$. The f^{-1} is the function that assigns to $y = x^3$ the number unique number x, that is

$$f^{-1}(y) = \sqrt[3]{y}\,. \tag{10.1}$$

More generally, the fact that $f^{-1}(x)$ is the number y such that $f(y) = x$ can be restated as: $f(f^{-1}(x)) = x$ for every point x in the domain f^{-1}, or alternatively $f^{-1}(f(x)) = x$ for every point in the domain of f.

We know that all increasing, and decreasing functions are one-to-one.

Exercise 10.1. Show that if f is increasing, then f^{-1} is also increasing.

Exercise 10.2. A function f is increasing if and only if $-f$ is decreasing.

However, it is not true that every one-to-one function is either increasing or decreasing.

Example 10.2. The function

$$f(x) = \begin{cases} x^2 & 0 < x < 1 \\ \frac{1}{x-1} + 1 & x > 1 \end{cases} \tag{10.2}$$

is a continuous one-to-one function, which is neither increasing, nor decreasing. However, it is of course increasing on the interval $(0, 1)$, and decreasing on the interval $(1, \infty)$.

This is true more generally: A continuous function f which is one-to-one on an interval is either increasing or decreasing on that interval (we will not prove this here); (see the **additional notes** to Module IV).

The purpose of this note is to record some very general properties of the inverse of one-to-one functions. For example, if f is continuous, does that mean that f^{-1} is continuous? Moreover, if f is differentiable, does that mean that f^{-1} is differentiable, and if so what is the derivative?

A glance at the graph of a one-to-one function suggests the answer, see Figure 10.1: Say L is the tangent line to the graph of f at the point $(a, f(a))$. Then the tangent line L' to the graph of f^{-1} at the point $(f(a), a)$ is obtained by reflecting L across the diagonal,

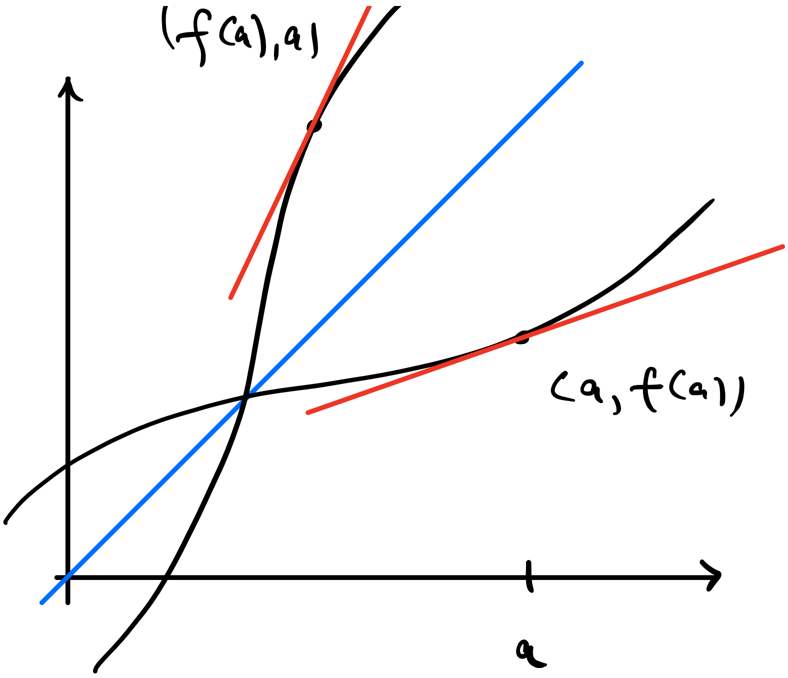

Figure 10.1.: Tangent to the inverse function.

and so the slope of L' is the reciprocal of the slope of L. In other words, this suggests that:

$$(f^{-1})'(f(a)) = \frac{1}{f'(a)}\,, \tag{10.3}$$

or alternatively,

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}\,. \tag{10.4}$$

There is another reason, this formula should be true: We know that

$$f\Big(f^{-1}(y)\Big) = y \tag{10.5}$$

for all y in the domain of f^{-1}, so by the chain rule:

$$f'\Big(f^{-1}(y)\Big) \cdot (f^{-1})'(y) = 1 \tag{10.6}$$

This argument is not a proof because it presupposes that we *know* that f^{-1} is differentiable, but it does tell us *if* f, and f^{-1} are differentiable, then $(f^{-1})'$ must be given by this formula.

This argument also tells us:

Corollary 10.1. *If f is a continuous one-to-one function defined on an interval and $f'(f^{-1}(a)) = 0$, then f^{-1} is* not *differentiable at a.*

Proof. Indeed, if f^{-1} *were* differentiable at a, then (10.6) would imply $0 = 1$. □

Example 10.3. The function $f(x) = x^3$ is continuous and one-to-one, and satisfies $f'(0) = 0$, and indeed f^{-1} is not differentiable at $0 = f^{-1}(0)$. (Draw a picture!)

Now finally, the *positive* results:

Theorem 10.2 (Continuity of the inverse)**.** *If f is continuous and one-to-one on an interval, then f^{-1} is also continuous.*

This is surprisingly cumbersome to show and we will not go into the proof here; (see the **additional notes** to Module IV).

Theorem 10.3 (Differentiability of the inverse)**.** *Let f be a continuous one-to-one function defined on an interval, and suppose that f is differentiable at $f^{-1}(b)$, with $f'(f^{-1}(b)) \neq 0$, then f^{-1} is differentiable at b, and*

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}\,. \tag{10.7}$$

Proof. In order to prove the theorem, we need to look at the difference quotient

$$\frac{f^{-1}(b+h) - f^{-1}(b)}{h} \tag{10.8}$$

where $b = f(a)$. For given h, let us choose k, depending on h, so that $b + h = f(a + k)$. Then we can also write this as

$$\frac{a+k-a}{f(a+k) - f(a)} = \frac{k}{f(a+k) - f(a)}\,. \tag{10.9}$$

Moreover, since f is differentiable at a, we have

$$f(a+k) - f(a) = f'(a)k + E(k)\,, \tag{10.10}$$

where $E(k)$ has the property that $\lim_{k\to 0} E(k)/k = 0$.

Note that k is given by $k = f^{-1}(b+h) - a = f^{-1}(b+h) - f^{-1}(b)$, and since f^{-1} is continuous at b (by Theorem 10.2), we know that k approaches 0 as h tends to 0.

Therefore

$$\lim_{h\to 0} \frac{f^{-1}(b+h) - f^{-1}(b)}{h} = \frac{1}{f'(a)}\,. \tag{10.11}$$

□

Example 10.4. Consider for any $n \in \mathbb{N}$,

$$f_n(x) = x^n \tag{10.12}$$

For n odd, this function is continuous and one-to-one, and for n even it is so if we take the domain to be $[0, \infty)$. We have

$$f_n^{-1}(x) = \sqrt[n]{x} = x^{1/n} \tag{10.13}$$

whose domain is $\mathbb{R}$ when n is odd, and $[0, \infty)$ if n is even. By Theorem 10.3 we have, for $x \neq 0$,

$$(f_n^{-1})'(x) = \frac{1}{nx^{(n-1)/n}} = \frac{1}{n}x^{1/n-1} \tag{10.14}$$

Hence if $f(x) = x^a$ for any integer a, or a the reciprocal of a natural number, then $f'(a) = na^{a-1}$. In fact, for any rational number $a = m/n$, we can write

$$f(x) = x^{m/n} = (x^{1/n})^m \tag{10.15}$$

which entails by the chain rule:

$$f'(x) = m(x^{1/n})^{m-1} \cdot \frac{1}{n}x^{1/n-1} = \frac{m}{n}x^{m/n-1}\,. \tag{10.16}$$

Problems

1. Find f^{-1} for each of the following functions f.
   a) $f(x) = x^3 + 1$
   b) $f(x) = (x - 1)^3$
2. Describe the graph of f^{-1} when
   a) f is increasing and always positive.
   b) f is increasing and always negative.
3. Prove that if f and g are one-to-one, then $f \circ g$ is also one-to-one. Find a formula for $(f \circ g)^{-1}$ in terms of f^{-1} and g^{-1}.
4. On which intervals $[a, b]$ will the following functions be one-to-one:
   a) $f(x) = x^3 - 3x^2$
   b) $f(x) = (1 + x^2)^{-1}$.
5. Suppose f is a one-to-one function and that f^{-1} has a derivative which is nowhere 0. Prove that f is differentiable.

   Hint: Apply Theorem 10.3
6. Find a formula for $(f^{-1})''(x)$ using the chain rule.

Additional: Critical points, and continuity of the inverse

10.1. Critical points

Further Reading

(Spivak, *Calculus*, Chapter 11) (Folland, *Advanced Calculus*, Chapter 2.8)

Vector Calculus: Advanced (MAST20032)

Let us first consider a function f of one variable.

Definition 10.1. A point x is a **local maximum point** for a function f, if there is some $\delta > 0$ such that $f(x) \geq f(y)$ for every y, with $x - \delta < y < x + \delta$. The number $f(x)$ itself is called the **local maximum value** of f on A. Similary for the **local minimum point and value**.

Theorem 10.1. *Let f be any function defined on (a, b). If x is a local maximum (or minimum) point for f on (a, b), and f is differentiable at x, then $f'(x) = 0$.*

Proof. If x is a local maximum point, then the difference quotient

$$\frac{f(x+h) - f(x)}{h} \tag{10.1}$$

is ≤ 0 for $h > 0$, and ≥ 0 for $h < 0$, because x is a local maximum point. Consequently,

$$\lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h} \leq 0\,, \qquad \lim_{h \to 0^-} \frac{f(x+h) - f(x)}{h} \geq 0 \tag{10.2}$$

and since f is differentiable both limits exist and are equal. Hence $f'(x) = 0$.

□

However, the converse is *not* true: A function f whose derivative $f'(x) = 0$ is zero at a point x does *not* necessarily have a minimum or maximum at that point. A simple example is the function $f(x) = x^3$, for which $f'(0)$, yet it does not have a minimum or maximum anywhere.

Definition 10.2. A **critical point** of a function f is a number x such that $f'(x) = 0$. The number $f(x)$ itself is called a **critical value**.

We have a similar notion for functions $f(x_1, x_2)$ of two variables $\vec{x} = (x_1, x_2)$.

Definition 10.3. A point $\vec{a} = (a_1, a_2)$ in the domain of $f(x_1, x_2)$ is a **critical point** if

$$\nabla f(\vec{a}) = 0\,. \tag{10.3}$$

Similarly, we can prove that if $f(x_1, x_2)$ has a **local maximum** at a point $\vec{a} = (a_1, a_2)$, in the sense that for some $\delta > 0$,

$$f(\vec{x}) \leq f(\vec{a}) \qquad |\vec{x} - \vec{a}| < \delta\,, \tag{10.4}$$

the $\vec{a}$ is a critical point of f. Indeed, consider the function

$$g(t) = f(\vec{a} + t\vec{u}) \tag{10.5}$$

then by assumption g has a local maximum at $t = 0$. If moreover f is differentiable at $\vec{a}$, then

$$g'(0) = \frac{\partial f}{\partial \vec{u}}(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u}\,. \tag{10.6}$$

From Theorem 10.1 we already know that $g'(0) = 0$, and since that holds for any vector $\vec{u}$, with $|\vec{u}| = 1$, we conclude that

$$\nabla f(\vec{a}) = 0\,. \tag{10.7}$$

10.2. Continuity and differentiability of the inverse

Further Reading

(Spivak, *Calculus*, Chapter 12)

We have seen that functions can be defined as pairs of numbers. The pairs of numbers (x, y) consist of points x in the domain of f and the values $y = f(x)$.

Definition 10.4. For any function f, the **inverse** of f, denoted by f^{-1}, is the set of pairs (y, x) for which the pair (x, y) is in f.

However, we have seen that what makes a collection of pairs (a, b) a function f, is that for each point a in the domain, there is a unique number b such that (a, b) is in f. So for f^{-1} to be a function, we need that for each $y = f(x)$ there is a unique number x such that $f(x) = y$. In other words, f needs to be **one-to-one**.

A function f which is one-to-one has an inverse function f^{-1}. The inverse function f^{-1} is itself one-to-one, and $(f^{-1})^{-1} = f$. In the pair (a, b) is in f, then $b = f(a)$; moreover if f is one-to-one, then (b, a) is in f^{-1}, and $a = f^{-1}(b)$.

We have seen that increasing, or decreasing functions are one-to-one.

Let us prove Theorem 10.2 *in the case that the domain of f is an open interval, and f is increasing on that interval.*

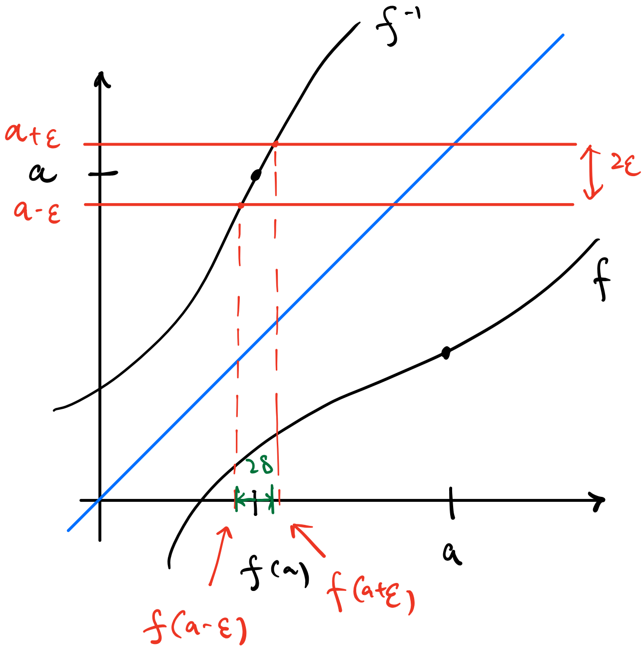

Figure 10.1.: Proof of the continuity of the inverse.

Proof of Theorem 10.2. Suppose f is increasing on an open interval, and a is a point in that interval. Let $b = f(a)$.

We want to show that

$$\lim_{y \to b} f^{-1}(y) = f^{-1}(b) \tag{10.8}$$

This means that we need to show that for any $\varepsilon > 0$, there is a $\delta > 0$, so that,

$$\text{if } |y - b| < \delta \text{ , then } |f^{-1}(y) - f^{-1}(b)| < \varepsilon \,, \tag{10.9}$$

or equivalently,

$$\text{if } f(a) - \delta < y < f(a) + \delta \text{ , then } a - \varepsilon < f^{-1}(y) < a + \varepsilon \,. \tag{10.10}$$

Now consider consider the points $(f(a+\varepsilon), a+\varepsilon)$, $(f(a-\varepsilon), a-\varepsilon)$ on the graph of f^{-1}; see Figure. We can choose $\delta > 0$ so that

$$f(a-\varepsilon) \leq f(a) - \delta < f(a) < f(a) + \delta \leq f(a+\varepsilon) \,. \tag{10.11}$$

With this choice of δ, we have that for all y, if $f(a) - \delta < y < f(a) + \delta$, then

$$f(a-\varepsilon) < y < f(a+\varepsilon). \tag{10.12}$$

We will now use that since f is increasing, also f^{-1} is increasing. Therefore:

$$a - \varepsilon = f^{-1}(f(a-\varepsilon)) < f^{-1}(y) < f^{-1}(f(a+\varepsilon)) = a + \varepsilon \,. \tag{10.13}$$

□

Additional Problems

1. In Note 8, we have seen in Theorem 8.3 a sufficient criterion for differentiability of a function of two variables, but we deferred the proof because it relies on the Mean Value Theorem. This exercise guides you through the proof of Theorem 8.3:

   Suppose the partial derivatives of a function $f(x, y)$ exist at every point. We want to show that at every point (a, b),

   $$f(a+h, b+k) = f(a, b) + \nabla f(a, b) \cdot (h, k) + E(h, k) \tag{10.14}$$

   where

   $$\nabla f(a, b) = (\partial_x f(a, b), \partial_y f(a, b)) \qquad \text{and} \qquad \lim_{(h,k)\to 0} \frac{E(h, k)}{|(h, k)|} \,. \tag{10.15}$$

   In order to prove that write

   $$E(h, k) = f(a+h, b+k) - f(a, b) - \partial_x f(a, b)h - \partial_y f(a, b)k \tag{10.16}$$

   and apply the mean value theorem twice. Finally use that the partial derivatives are *continuous functions* to evaluate the limit.

Module V.

Integration

Note 11.

The fundamental theorem of Calculus

Further Reading

(Spivak, *Calculus*, Chapter 15)

Real Analysis: Advanced (MAST20033)

Given a continuous function $f \geq 0$ on $[a,b]$ we can talk about the "area under the graph of f". In fact, this concept can be made precise for a larger class of *bounded* functions which are called **integrable** (in particular they need not be nonnegative), and the number $\int_a^b f$ which formalizes the concept of "area" is called the **integral**.

In this subject, we will not *define* the integral, or attempt to discuss this notion with the same level of care as we have treated the topics of continuity, and differentiability, for instance. Consequently, we cannot hope to *prove* any of the fundamental theorems about integration, most importantly the **fundamental theorems of Calculus** stated below, which relate integration and differentiation.

In this note, we will only state these theorems as *facts*.

Basic properties of the integral. The basic properties of the integral, which we denote interchangeably by

$$\int_a^b f(x)\mathrm{d}x \quad \text{or} \quad \int_a^b f \tag{11.1}$$

are that

$$\int_a^b f = \int_a^c f + \int_c^b f \qquad (\text{for any } a < c < b) \tag{11.2}$$

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g \tag{11.3}$$

$$\int_a^b f = c\int_a^b f \qquad (\text{for any } c \in \mathbb{R}) \tag{11.4}$$

Recall that given a differentiable function f, we obtained a *new* function f' by differentiation. In a similar way, given an integrable function f, a *new* function can be obtained by integration. In fact, given any continuous function f on $[a,b]$, consider the function

$$F(x) = \int_a^x f \qquad (x \in [a,b]) \tag{11.5}$$

It turns out this function is *itself* continuous in x, and more importantly always *differentiable*:

Theorem 11.1 (First fundamental theorem of Calculus)**.** *If f is continuous on $[a,b]$, then F is differentiable on (a,b), and*

$$F'(x) = f(x)\,. \tag{11.6}$$

Remark 11.1. The theorem says that given *any* continuous function on $[a,b]$ (in fact, more generally *any* merely integrable function), there *always* exists a function F, whose derivative is f; namely (11.5).

Also note that if G is defined by

$$G(x) = \int_x^b f \tag{11.7}$$

then $G(x) = \int_a^b f - \int_a^x f$, so consequently $G'(c) = -f(c)$.

As a consequence, if it is known that a function f is the derivative of another function g, the evaluation of its integral is a triviality:

Corollary 11.2. *If f is continuous on $[a,b]$ and $f = g'$ for some function g, then*

$$\int_a^b f = g(b) - g(a)\,. \tag{11.8}$$

Proof. Let $F(x)$ be defined by (11.5), then $F' = f = g'$ on $[a,b]$. Consequently by the mean value theorem $F = g + c$ for some constant c, cf. Corollary 9.4. In fact, $F(a) = 0 = g(a) + c$, and thus $F(x) = g(x) - g(a)$, in particular for $x = b$. □

Example 11.1. If $f(x) = x^2$, and $g(x) = x^3/3$, then $g'(x) = x^2 = f(x)$, so

$$\int_a^b x^2 \mathrm{d}x = (b^3 - a^3)/3 \tag{11.9}$$

Exercise 11.1. Compute the integral of $f(x) = x^n$ on the interval $[a,b]$ for any natural number $n \in \mathbb{N}$.

Example 11.2. Also for $f(x) = x^{-n}$, where $n \in \mathbb{N}$, $n \neq 1$, we know that we can find g with $g'(x) = f(x)$, at least for $x \neq 0$: $g(x) = (-n+1)^{-1} x^{-n+1}$. Thus for $0 < a < b$

$$\int_a^b x^{-n} \mathrm{d}x = \frac{1}{n-1}\Big(\frac{1}{a^{n-1}} - \frac{1}{b^{n-1}}\Big)\,. \tag{11.10}$$

The exception $n = 1$ in the above example is significant. While there is no monomial whose derivative is the function $1/x$, we do know, *by the fundamental theorem of Calculus*, that there exists a function $g(x)$ whose derivative is $f(x) = 1/x$ for $x > 0$, namely

$$g(x) = \int_a^x \frac{1}{t} \mathrm{d}t \tag{11.11}$$

For $a = 1$ this serves as the *definition* of the logarithm; (see the **additional notes** to Module V).

Problems

1. The fundamental theorem of Calculus, together with the chain rule, allow us to compute derivatives of a variety of functions defined in terms of integrals.
   Example 11.3. Let us compute the derivative of the function

$$f(x) = \int_a^{\sin x} \frac{1}{1+\sin^2(t)}\mathrm{d}t \tag{11.12}$$

   We can view $f = F \circ \sin$ as the composition of sin with F, where

$$F(x) = \int_a^{x} \frac{1}{1+\sin^2(t)}\mathrm{d}t \tag{11.13}$$

   Hence by the chain rule we have

$$f'(x) = F'(\sin(x))\cos(x) = \frac{\cos(x)}{1+\sin^2(\sin(x))}\,. \tag{11.14}$$

   Find the derivatives of the following functions

   a) $f(x) = \sin\Big(\int_a^x \frac{1}{1+\sin^2(t)}\mathrm{d}t\Big)$

   b) $F(x) = \int_a^{x^3} \sin^3(t)\mathrm{d}t$

   c) $F(x) = \sin\Big(\int_0^x \big(\int_0^y \sin^3(t)\mathrm{d}t\big)\mathrm{d}y\Big)$

2. Find $(f^{-1})'(0)$ if

$$f(x) = \int_0^x 1 + \sin\big(\sin(t)\big)\mathrm{d}t \tag{11.15}$$

   Hint: Do not try to evaluate the integral!

3. a) Find F' if

$$F(x) = \int_0^x x f(t)\mathrm{d}t \tag{11.16}$$

   Hint: The answer is *not* $xf(x)$.

   b) Prove that if f is continuous, then

$$\int_0^x f(u)(x-u)\mathrm{d}u = \int_0^x \Big(\int_0^u f(t)\mathrm{d}t\Big)\mathrm{d}u \tag{11.17}$$

   c) Prove that moreover

$$\int_0^x f(u)(x-u)^2\mathrm{d}u = 2\int_0^x \int_0^{u_2} \Big(\int_0^{u_1} f(t)\mathrm{d}t\Big)\mathrm{d}u_1\Big)\mathrm{d}u_2 \tag{11.18}$$

4. Show that if h is continuous, and f and g are differentiable, and

$$F(x) = \int_{f(x)}^{g(x)} h(t)\mathrm{d}t\,, \tag{11.19}$$

then $F'(x) = h(g(x))g'(x) - h(f(x))f'(x)$.

5. Suppose $f > 0$ on $[a, b]$. Find

$$\int_a^b \frac{f'(t)}{f(t)}\mathrm{d}t \tag{11.20}$$

Note 12.

The "simplest" differential equations

The simplest example of a *differential equation* is

$$y'(x) = f(x)\,. \tag{12.1}$$

Here we are looking for a *function* y whose derivative is f. The fundamental theorem of Calculus says that *for any continuous function* $f(x)$ *this differential equation has a solution*, namely

$$y(x) = \int_a^x f(t)\mathrm{d}t\,. \tag{12.2}$$

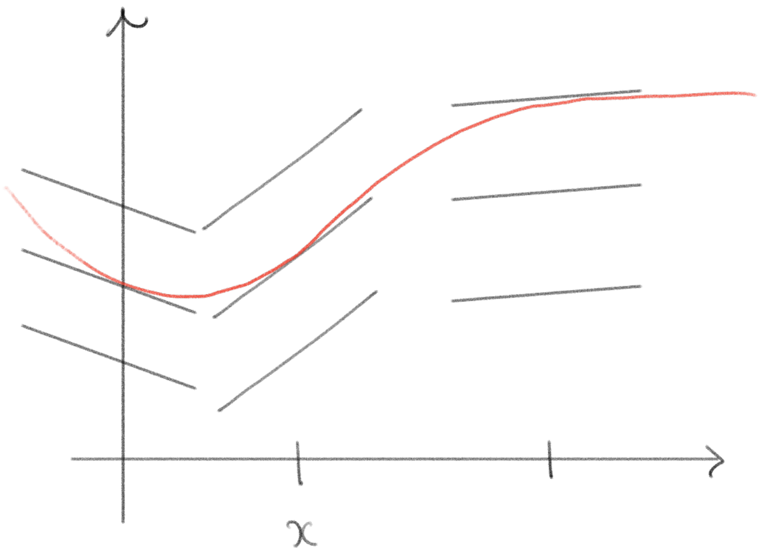

Figure 12.1.: The function f in (12.1) geometrically prescribes a slope, and gives rise to the simplest example of a direction field. Solutions are functions whose graph are tangential to the indicated slopes at every point.

Another example is the equation

$$y' = y\,. \tag{12.3}$$

Here we are looking for a *function* whose derivative equals itself.

Exercise 12.1. Draw the direction field for (12.3) as in Figure 12.1.

An example of a *solution* to this equation is the **exponential function**

$$\exp(x) = e^x\,. \tag{12.4}$$

The exponential can be *defined* as the *inverse of the logarithm*; see Note 13 below.

$$\log(x) = \int_1^x \frac{1}{t}\mathrm{d}t \quad (x > 0)\,, \qquad \exp(x) = \log^{-1}(x)\,. \tag{12.5}$$

The statement that the exponential function solves the differential equation (12.3) then follows from the formula for the derivative of the inverse given in Note 10:

$$\exp'(x) = (\log^{-1})'(x) = \frac{1}{\log'(\exp(x))} = \exp(x)\,. \tag{12.6}$$

Of course e^x is not the only function with that property, for example also $f(x) = c\exp(x)$ satisfies the relation $f'(x) = f(x)$ for any constant $c \in \mathbb{R}$. However, these are all:

Theorem 12.1. *If f is differentiable and satisfies*

$$f'(x) = f(x) \qquad \textit{for all } x \in \mathbb{R} \tag{12.7}$$

then there is a number c such that

$$f(x) = ce^x \qquad \textit{for all } x \in \mathbb{R} \tag{12.8}$$

Remark 12.1. In other words, the solutions of the differential equation $y' = y$ are all of the form $y(x) = ce^x$ for some constant $c \in \mathbb{R}$.

Proof. Let $g(x) = e^{-x}f(x)$, then

$$g'(x) = \frac{e^x f'(x) - f(x)e^x}{(e^x)^2} = 0 \tag{12.9}$$

Therefore there is a number c such that for all x:

$$g(x) = \frac{f(x)}{e^x} = c \tag{12.10}$$

□

An important example of a *second order* differential equation is

$$y'' + y = 0\,. \tag{12.11}$$

It is easy to verify that the trigonometric functions are solutions, namely for both

$$f(x) = \cos(x)\,, \text{ and } \qquad f(x) = \sin(x)\,, \tag{12.12}$$

we have $f'' + f = 0$, but they differ in their values of f and f' at $x = 0$. We have, respectively, for these two functions

$$f(0) = 1\,, f'(0) = 0\,; \text{ and } \qquad f(0) = 0\,, f'(0) = 1\,. \tag{12.13}$$

It may at first be surprising that $\cos(x)$ and $\sin(x)$ are the *only* solutions to (12.11) with these values at 0.

Lemma 12.2. *Suppose that f is twice differentiable and that*

$$f'' + f = 0 \tag{12.14}$$

$$f(0) = 0 \qquad f'(0) = 0\,. \tag{12.15}$$

Then $f = 0$.

Proof. Since f satisfies the equation $f'' + f = 0$, we compute

$$\Big((f')^2 + f^2\Big)' = 2f'\Big(f'' + f\Big) = 0 \tag{12.16}$$

so $(f')^2 + f^2$ must be constant, and evaluated at $x = 0$ it is 0, hence

$$(f'(x))^2 + (f(x))^2 = 0 \tag{12.17}$$

for all x. □

This also means that if

$$f'' + f = 0 \tag{12.18}$$

$$f(0) = a \qquad f'(0) = b \tag{12.19}$$

then

$$f(x) = a\cos(x) + b\sin(x)\,. \tag{12.20}$$

Indeed, if we define

$$g(x) = f(x) - a\cos(x) - b\sin(x) \tag{12.21}$$

then g also satisfies $g'' + g = 0$, and $g(0) = 0$, and $g'(0) = 0$, from which we conclude with the Lemma that $g(x) = 0$.

An unexpected consequence are the addition theorems for trigonometric functions.

Proposition 12.3.

$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y) \tag{12.22}$$

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y) \tag{12.23}$$

Proof. For fixed y, set

$$f(x) = \sin(x+y)\,. \tag{12.24}$$

Then

$$f''(x) + f(x) = 0 \tag{12.25}$$

and

$$f(0) = \sin(y) \qquad f'(0) = \cos(y)\,. \tag{12.26}$$

So by the argument above,

$$f(x) = \sin(y)\cos(x) + \cos(y)\sin(x)\,. \tag{12.27}$$

□

In particular we note that $\cos(2x) = \cos^2(x) - \sin^2(x)$, and $\sin(2x) = 2\cos(x)\sin(x)$.

Note 13.

Logarithm and exponential function

Further Reading

(Spivak, *Calculus*, Chapter 18), with more motivation for the definition of the logarithm. (Apostel, *Calculus I*, Chapter 8.3), with an emphasis on the exponential functions as the solution to a differential equation.

13.1. Logarithm

The logarithm is an example of a function that is defined by an integral.

Definition 13.1 (Logarithm)**.** For $x > 0$, we set

$$\log(x) = \int_1^x \frac{1}{t} \mathrm{d}t \tag{13.1}$$

Exercise 13.1. Sketch the graph of the logarithm.

Proposition 13.1. *If $x, y > 0$, then*

$$\log(xy) = \log(x) + \log(y) \tag{13.2}$$

Proof. Note that by the fundamental theorem of calculus $\log'(x) = 1/x$. Now choose a number $y > 0$, and let $f(x) = \log(xy)$. Then

$$f'(x) = \log'(xy)y = \frac{y}{xy} = \frac{1}{x} \tag{13.3}$$

which says that $f' = \log'$. This implies that there is a number c such that

$$f(x) = \log(x) + c \tag{13.4}$$

for all $x > 0$, and we can find c by evaluating

$$f(1) = \log(y) = \log(1) + c = c \tag{13.5}$$

and therefore

$$\log(xy) = f(x) = \log(x) + c = \log(x) + \log(y) \, . \tag{13.6}$$

Since this is true for all $y > 0$, the theorem is proved. □

Exercise 13.2. Show by induction that if n is a natural number and $x > 0$, then

$$\log(x^n) = n \log(x) \tag{13.7}$$

Corollary 13.2. *If $x, y > 0$, then*

$$\log(\frac{x}{y}) = \log(x) - \log(y)\,. \tag{13.8}$$

Proof. This is true because

$$\log(x) = \log(\frac{x}{y} y) = \log(x/y) + \log(y)\,. \tag{13.9}$$

□

The function $\log(x)$ is clearly increasing but since $\log'(x) = 1/x$ the slope gets very small when x is large, and consequently $\log(x)$ grows more and more slowly. It is not immediately clear if the function is bounded or unbounded. However, for any $n \in \mathbb{N}$,

$$\log(2^n) = n \log(2) \tag{13.10}$$

and $\log 2 > 1$; similarly

$$\log(2^{-n}) = -n \log(2)\,. \tag{13.11}$$

Thus by the intermediate value theorem the logarithm takes on any value $t \in \mathbb{R}$.

13.2. Exponential function

We have seen that the logarithm is increasing on $(0, \infty)$ and takes all values in $\mathbb{R}$. Therefore the inverse function $\log^{-1}$ exists and its domain is $\mathbb{R}$. This function is the *exponential function.*

Definition 13.2 (Exponential function)**.** For any real number x, we set

$$\exp(x) = \log^{-1}(x)\,. \tag{13.12}$$

Theorem 13.3. *For all numbers x,*

$$\exp'(x) = \exp(x) \tag{13.13}$$

Moreover, for any two numbers x and y,

$$\exp(x + y) = \exp(x)\exp(y)\,. \tag{13.14}$$

Proof. With the formula for the derivative of the inverse from Lecture 10,

$$\exp'(x) = (\log^{-1})'(x) = \frac{1}{\log'(\exp(x))} = \exp(x)\,. \tag{13.15}$$

Moreover, since $x = \log(a)$, $y = \log(b)$ for some $a, b > 0$,

$$\exp(x + y) = \exp(\log(a) + \log(b)) = \exp(\log(ab)) = ab = \exp(x)\exp(y)\,. \tag{13.16}$$

□

The number $e = \exp(1)$ is called **Euler's number**. For any number x, we *define* $e^x = \exp(x)$; see Additional.

Exponential growth. Another important observation is that "the exponential grows faster than any polynomial".

Theorem 13.4. *For any natural number n,*

$$\lim_{x\to\infty} \frac{e^x}{x^n} = \infty \tag{13.17}$$

Proof. Let us first show that

$$\lim_{x\to\infty} e^x = \infty\,. \tag{13.18}$$

From the definition of the logarithm we can infer that

$$\log(x) < x \qquad (x > 0)\,. \tag{13.19}$$

Indeed, for $0 < x < 1$ we have $\log(x) < 0$; moreover $\log(1) = 0 < 1$, and for $x > 1$,

$$\log(x) = \int_1^x \frac{\mathrm{d}t}{t} \leq (x-1) < x\,. \tag{13.20}$$

Therefore

$$x = \exp(\log(x)) < e^x\,, \tag{13.21}$$

which in particular implies (13.18).

Next we prove that

$$\lim_{x\to\infty} \frac{e^x}{x} = \infty\,. \tag{13.22}$$

Since

$$\frac{e^x}{x} = \frac{1}{2}\Big(\frac{e^{x/2}}{x/2}\Big)e^{x/2} \tag{13.23}$$

and in view of (13.21) the factor in parenthesis is bounded from below by 1, this statement also follows from (13.18).

Similarly we can now write

$$\frac{e^x}{x^n} = \Big(\frac{e^{x/n}}{x}\Big)^n = \frac{1}{n^n}\Big(\frac{e^{x/n}}{x/n}\Big)^n \tag{13.24}$$

which then proves the statement of the Theorem by (13.22).

□

Problems

1. a) Check that the derivative of $\log\circ f$ is f'/f.

   Note: The derivative of $\log\circ f$ is called the *logarithmic derivative* and is sometimes easier to compute than f', because taking the logarithm turns products into sums. The formula says that multiplying $(\log\circ f)'$ by f recovers f', and this process of finding the derivative of f is called *logarithmic differentiation*.

b) Use logarithmic differentiation to find the derivative of the following functions

   i. $f(x) = (1+x)(1+e^{x^2})$

   ii. $f(x) = \frac{e^x - e^{-x}}{e^{2x}(1+x^3)}$

2. Draw the graph of each of the following functions

   a) $f(x) = e^{x+1}$

   b) $f(x) = e^x + e^{-x}$

   c) $f(x) = e^x - e^{-x}$

   Are these functions the solutions to a differential equation?

Note 14.

Methods of integration

Definition 14.1 (Primitive)**.** A function F satisfying $F' = f$ is called the *primitive* of f.

Example 14.1. If $F(x) = x \log x - x$, then $F'(x) = \log x$. So F is a primitive of the logarithm. Consequently, by the fundamental theorem of calculus,

$$\int_a^b \log(x)\mathrm{d}x = b \log b - a \log a - b + a\,. \tag{14.1}$$

A continuous function f always has a primitive, namely

$$F(x) = \int_a^x f(t)\mathrm{d}t\,. \tag{14.2}$$

However, in this lecture we will try to find a primitive which can be written in terms of elementary functions, namely the trigonometric functions and their inverses, and the logarithmic and exponential functions, and rational functions formed thereof.

Remark 14.1. Elementary primitives usually cannot be found. For example, there is no elementary function F such that

$$F'(x) = e^{-x^2}\,. \tag{14.3}$$

The basic methods for finding elementary primitives are actually theorems which allow us to express primitives of one function in terms of primitives of other functions. To integrate we will therefore need a list of primitives for *some* functions, and such a list can be obtained simply by differentiating various well-known functions.

Definition 14.2. For the primitive of a function f we often use the notation

$$\int f(x)\mathrm{d}x \qquad \text{or} \int f\,. \tag{14.4}$$

These are also called *indefinite integrals*, in contrast to *definite integrals* of a function f with primitive F for which we adopt the notation

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a) = F(x)\big|_a^b\,.$$

Example 14.2. We can verify the following formulas by differentiating the right hand sides:

$$\int a\mathrm{d}x = ax \tag{14.5}$$

$$\int x^n \mathrm{d}x = \frac{x^{n+1}}{n+1} \qquad (n \neq -1) \tag{14.6}$$

$$\int \frac{1}{x}\mathrm{d}x = \log x \tag{14.7}$$

$$\int e^x \mathrm{d}x = e^x \tag{14.8}$$

$$\int \sin x \mathrm{d}x = -\cos x \tag{14.9}$$

$$\int \cos x \mathrm{d}x = \sin x \tag{14.10}$$

$$\int \sec^2 x \mathrm{d}x = \tan x \tag{14.11}$$

$$\int \frac{\mathrm{d}x}{1+x^2} = \arctan x \tag{14.12}$$

$$\int \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \arcsin x \tag{14.13}$$

14.1. Integration by parts, and substitution

We will now learn two important methods which can be used to express the primitive of one function in terms of the primitive of another.

Exercise 14.1. Convince yourself *by using differentiation* that a primitive of $f+g$ can be obtained by adding a primitive of f to a primitive of g, and a primitive of $c \cdot f$ can be obtained by multiplying a primitive of f by c. This is expressed by the two general formulas:

$$\int f+g = \int f + \int g \tag{14.14}$$

$$\int c \cdot f = c \int f\,. \tag{14.15}$$

The product rule of differentiation yields the following important theorem:

Theorem 14.1 (Integration by parts)**.** *If f' and g' are continuous, then*

$$\int fg' = fg - \int f'g \tag{14.16}$$

$$\int_a^b f(x)g'(x) = f(x)g(x)\big|_a^b - \int_a^b f'(x)g(x)\mathrm{d}x \tag{14.17}$$

Proof. The formula for the indefinite integrals states that the primitive of the function $f'g + fg'$ is fg, which is of course follows immediatly from the product rule:

$$(fg)' = f'g + fg' \tag{14.18}$$

The formula for the definite integral follows if we integrate both side of this equation on the interval $[a, b]$. □

Example 14.3.

$$\int xe^x \mathrm{d}x = xe^x - e^x \tag{14.19}$$

Example 14.4.

$$\int x \sin x \mathrm{d}x = -x\cos(x) + \int \cos x = -x\cos(x) + \sin(x) \tag{14.20}$$

Example 14.5.

$$\int \log x = x\log x - \int x \cdot (1/x)\mathrm{d}x = x\log x - x \tag{14.21}$$

Example 14.6. The computation of the primitive of $\log(x)/x$ is an example where the result is obtained in two steps:

$$\int \frac{1}{x}\log(x)\mathrm{d}x = \log(x)^2 - \int \log(x)\frac{1}{x}\mathrm{d}x \tag{14.22}$$

$$\int \frac{\log x}{x}\mathrm{d}x = \frac{1}{2}(\log x)^2 \tag{14.23}$$

Example 14.7. Any previously computed primitive can be used for integration by parts:

$$\begin{aligned}\int (\log(x))^2 \mathrm{d}x =& \int (\log x)(\log x)\mathrm{d}x \\ =&\log(x)\int \log(x) - \int \frac{1}{x}\Big(x\log x - x\Big)\mathrm{d}x \\ =&\log(x)\Big(x\log x - x\Big) - \int (\log(x) - 1)\mathrm{d}x \\ =&x(\log x)^2 - 2x\log(x) + 2x\end{aligned} \tag{14.24}$$

The next method is a consequence of the chain rule of differentiation.

Theorem 14.2 (Substitution formula). *If f and g' are continuous, then*

$$\int_{g(a)}^{g(b)} f(y)\mathrm{d}y = \int_a^b f(g(x))g'(x)\mathrm{d}x \tag{14.25}$$

$$\int_{g(a)}^{g(b)} f = \int_a^b (f \circ g) \cdot g' \tag{14.26}$$

Proof. Let F be a primitive of f, namely for some c,

$$F(x) = \int_c^x f(y)\mathrm{d}y\,, \tag{14.27}$$

then $F' = f$, and

$$(F \circ g)' = (F' \circ g) \cdot g' = (f \circ g) \cdot g'\,, \tag{14.28}$$

and so after integration, by the fundamental theorem of Calculus,

$$\int_a^b (f \circ g)(x) \cdot g'(x)\mathrm{d}x = (F \circ g)|_a^b = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f\,. \tag{14.29}$$

□

Example 14.8.

$$\int_a^b \sin^5(x)\cos(x)\mathrm{d}x = \frac{1}{6}\sin^6(x)|_a^b \tag{14.30}$$

because with $f(x) = x^5$, and $g(x) = \sin(x)$, this integral is of the form

$$\int_a^b f(g(x))g'(x)\mathrm{d}x = \int_{g(a)}^{g(b)} f(y)\mathrm{d}y = (F \circ g)|_a^b \tag{14.31}$$

where $F(x) = x^6/6$ is the primitive of f.

Example 14.9.

$$\int_a^b \tan(x)\mathrm{d}x = -\int_a^b \frac{-\sin(x)}{\cos(x)}\mathrm{d}x = -\log\cos(b) + \log(\cos(a)) \tag{14.32}$$

Example 14.10.

$$\int_a^b \frac{\mathrm{d}x}{x\log x} = \log\log(x)|_a^b \tag{14.33}$$

We are usually interested in primitives rather than definite integrals, but if we can find $\int_a^b f(x)\mathrm{d}x$ for all a, and b, then we can certainly find $\int f$. For example, from (14.30) it follows that

$$\int \sin^5(x)\cos(x)\mathrm{d}x = \frac{\sin^6 x}{6}\,. \tag{14.34}$$

It is quite uneconomical to find the primitive by first evaluating a definite integral. Instead we have the following procedure:

Substitution procedure. Find a function $g(x)$ so that after setting

$$u = g(x) \qquad \mathrm{d}u = g'(x)\mathrm{d}x \tag{14.35}$$

only the variable u appears. Then find a primitive in terms of u, and substitute $g(x)$ back in for u.

Example 14.11. Consider again Example 14.10. Set

$$u = \log x \qquad \mathrm{d}u = \frac{1}{x}\mathrm{d}x \tag{14.36}$$

then

$$\int \frac{1}{x\log x}\mathrm{d}x = \int \frac{\mathrm{d}u}{u} = \log u \tag{14.37}$$

and substitution back in for u gives the answer $\log\log(x)$.

Similarly for the other examples above:

Example 14.12. To evaluate (14.30), set

$$u = \sin(x) \qquad \mathrm{d}u = \cos(x)\mathrm{d}x \tag{14.38}$$

so

$$\int \sin^5(x)\cos(x)\mathrm{d}x = \int u^5 \mathrm{d}u = \frac{1}{6}u^6 = \frac{1}{6}\big(\sin(x)\big)^6 \tag{14.39}$$

Example 14.13. To evaluate

$$\int \frac{x}{1+x^2}\mathrm{d}x \tag{14.40}$$

set

$$u = 1 + x^2 \qquad \mathrm{d}u = 2x\mathrm{d}x \tag{14.41}$$

so this integral equals

$$\int \frac{1}{2}\frac{\mathrm{d}u}{u} = \frac{1}{2}\log u = \frac{1}{2}\log(1+x^2) \tag{14.42}$$

14.2. Applications of the substitution formula

In many cases a factor $g'(x)$ is easily recognized, and integration by substition is straightforward.

Exercise 14.2. Find

$$\int e^{\sin x}\cos(x)\mathrm{d}x \tag{14.43}$$

$$\int \frac{e^x}{\sqrt{1-e^{2x}}}\mathrm{d}x \tag{14.44}$$

Example 14.14.

$$\int e^{3x}\mathrm{d}x = \frac{e^{3x}}{3} \tag{14.45}$$

$$\int \cos(4x)\mathrm{d}x = \frac{\sin(4x)}{4} \tag{14.46}$$

More interesting uses of the substitution formula appear when the factor $g'(x)$ does *not* appear.

Example 14.15. Consider

$$\int \frac{1+e^x}{1-e^x}\mathrm{d}x\,. \tag{14.47}$$

The obvious substitution to try is

$$u = e^x \qquad \mathrm{d}u = e^x\mathrm{d}x \tag{14.48}$$

and even though this factor does not appear in the integral we are led to

$$\int \frac{1+e^x}{1-e^x}\mathrm{d}x = \int \frac{1+u}{1-u}\frac{1}{u}\mathrm{d}u\,. \tag{14.49}$$

This can be integrated easily once we recognise that

$$\frac{1+u}{1-u}\frac{1}{u} = \frac{2}{1-u} + \frac{1}{u}\,, \tag{14.50}$$

hence

$$\int \frac{1+e^x}{1-e^x}\mathrm{d}x = -2\log(1-e^x) + \log(e^x) \tag{14.51}$$

Alternatively we could have set

$$u = e^x \qquad x = \log u \qquad \mathrm{d}x = \frac{1}{u}\mathrm{d}u \tag{14.52}$$

then immediately

$$\int \frac{1+e^x}{1-e^x}\mathrm{d}x = \int \frac{1+u}{1-u}\frac{1}{u}\mathrm{d}u\,. \tag{14.53}$$

This trick works more generally: Suppose we make the substitution

$$u = g(x) \tag{14.54}$$

and say we are in the situation that g is one-to-one, at least for all x under consideration, in particular $g' \neq 0$, then we can solve

$$x = g^{-1}(u)\,. \tag{14.55}$$

In order to find

$$\int f(g(x))\mathrm{d}x \tag{14.56}$$

we would conventionally write

$$\mathrm{d}u = g'(x)\mathrm{d}x \tag{14.57}$$

so

$$\int f(g(x))\mathrm{d}x = \int f(g(x))\frac{1}{g'(x)}g'(x)\mathrm{d}x = \int f(u)\frac{1}{g'(g^{-1}(u))}\mathrm{d}u\,. \tag{14.58}$$

Now note that

$$\frac{1}{g'(g^{-1}(u))} = (g^{-1})'(u) \tag{14.59}$$

so we could have obtained the identity

$$\int f(g(x))\mathrm{d}x = \int f(u)(g^{-1})'(u)\mathrm{d}u \tag{14.60}$$

equally well with the substitution rule

$$x = g^{-1}(u) \qquad \mathrm{d}x = (g^{-1})'(u)\mathrm{d}u\,. \tag{14.61}$$

Example 14.16. Consider the integral

$$\int \frac{e^{2x}}{\sqrt{e^x+1}}\mathrm{d}x\,. \tag{14.62}$$

Set

$$u = \sqrt{e^x+1}\,, \tag{14.63}$$

then

$$u^2 = e^x + 1 \tag{14.64}$$

$$x = \log(u^2-1) \qquad \mathrm{d}x = \frac{2u}{u^2-1}\mathrm{d}u \tag{14.65}$$

hence

$$\begin{aligned}\int \frac{e^{2x}}{\sqrt{e^x+1}}\mathrm{d}x &= \int \frac{(u^2-1)^2}{u}\frac{2u}{u^2-1}\mathrm{d}u = 2\int (u^2-1)\mathrm{d}u \\ &= \frac{2}{3}u^3 - 2u = \frac{2}{3}(e^x+1)^{\frac{3}{2}} - 2(e^x+1)^{\frac{1}{2}}\end{aligned} \tag{14.66}$$

Finally let us look at some examples for the **integration of trigonometric functions by substitution**. When integrating a monomials in trigonometric functions it is useful to remember the formulas from Prop. 12.3, in particular

$$\cos(2x) = \cos^2(x) - \sin^2(x) \tag{14.67}$$

from which we obtain $\cos(2x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x)$ or

$$\cos^2(x) = \frac{\cos(2x)+1}{2} \qquad \sin^2(x) = \frac{1-\cos(2x)}{2}\,. \tag{14.68}$$

Example 14.17. Consider the integral

$$\int \sqrt{1-x^2}\ \mathrm{d}x \tag{14.69}$$

If we set

$$x = \sin(u) \tag{14.70}$$

then $\sqrt{1-x^2} = \cos(x)$ simplifies, so we are led to the substitution

$$u = \arcsin(x)\,. \tag{14.71}$$

Then

$$\int \sqrt{1-x^2}\ \mathrm{d}x = \int \sqrt{1-\sin^2(u)}\cos(u)\mathrm{d}u = \int \cos^2(u)\mathrm{d}u\,. \tag{14.72}$$

This integral can be evaluated using that $\cos^2(u) = (1+\cos(2u))/2$ and we find that

$$\int \cos^2(u)\mathrm{d}u = \frac{u}{2} + \frac{\sin(2u)}{4} \tag{14.73}$$

and substituting back in $u = \arcsin(x)$ we have an expression for the primitive of (14.69):

$$\begin{aligned}\int \cos^2(u)\mathrm{d}u =& \frac{\arcsin(x)}{2} + \frac{\sin(2\arcsin(x))}{4} \\ =& \frac{\arcsin(x)}{2} + \frac{x\sqrt{1-x^2}}{2}\end{aligned} \tag{14.74}$$

We can also use the formulas (14.68) to integrate:

$$\int \sin^n x\,\mathrm{d}x \qquad \int \cos^n x\,\mathrm{d}x \tag{14.75}$$

Example 14.18. We see immediately that

$$\int_0^{2\pi} \cos^2(x)\mathrm{d}x = \pi \tag{14.76}$$

Example 14.19. Let us consider the case $n = 3$:

$$\begin{aligned}\int \sin^3(x)\mathrm{d}x =& \int \sin(x)(1 - \cos^2(x))\mathrm{d}x \\ =& -\cos(x) + \frac{1}{3}\cos^3(x)\end{aligned} \tag{14.77}$$

This approach works for any $n = 2k + 1$ odd.

Example 14.20. Finally let us compute

$$\begin{aligned}\int_0^{2\pi} \cos^4(x)\mathrm{d}x =& \int_0^{2\pi} \Big(\frac{\cos(2x)+1}{2}\Big)^2 \mathrm{d}x \\ =& \int_0^{2\pi} \frac{1}{4}\cos^2(2x) + \frac{1}{2}\cos(2x) + \frac{1}{4}\mathrm{d}x \\ =& \frac{1}{4}\int_0^{2\pi} \Big(\frac{\cos(4x)+1}{2} + 1\Big)\mathrm{d}x = \frac{3\pi}{4}\end{aligned} \tag{14.78}$$

Problems

1. Find elementary expressions for the following primitives.

   a)
$$\int \frac{\sqrt[5]{x^3} + \sqrt[6]{x}}{\sqrt{x}} \mathrm{d}x$$

   b)
$$\int \frac{\mathrm{d}x}{\sqrt{x-1} + \sqrt{x+1}}$$

   c)
$$\int \frac{\mathrm{d}x}{a^2 + x^2}$$

   d)
$$\int \frac{\mathrm{d}x}{\sqrt{a^2 - x^2}}$$

2. Solve by substitution.

   a)
$$\int x e^{-x^2} \mathrm{d}x$$

   b)
$$\int \frac{e^x \mathrm{d}x}{e^{2x} + 2e^x + 1}$$

   c)
$$\int \log(\cos(x)) \tan x \mathrm{d}x$$

3. Solve by integration by parts.

   a)
$$\int x^3 e^{x^2} \mathrm{d}x$$

   b)
$$\int (\log x)^3 \mathrm{d}x$$

   c)
$$\int \frac{\log(\log x)}{x} \mathrm{d}x$$

4. Find the following primitives in elementary terms using substitution.

   a)
$$\int e^x \sin e^x \mathrm{d}x$$

b)

$$\int x\sqrt{1-x^2}\mathrm{d}x$$

c)

$$\int \frac{\log(\log(x))}{x\log(x)}\mathrm{d}x$$

5. The following integrations involve substitutions of various types. There is no general rule, but try to substitute for an expression which appears frequently and prominently. Also remember that it usually helps to express x directly in terms of u.

a)

$$\int \frac{\mathrm{d}x}{1+\sqrt{x+1}}$$

b)

$$\int \frac{\mathrm{d}x}{1+e^x}$$

c)

$$\int \frac{\mathrm{d}x}{\sqrt{x}+\sqrt[3]{x}}$$

d)

$$\int \frac{1}{\sqrt{1+e^x}}\mathrm{d}x$$

e)

$$\int \frac{1}{2+\tan(x)}\mathrm{d}x$$

Module VI.

Ordinary differential equations

Note 15.

First order linear differential equations

In the previous lectures we have already encountered the notion of a solution to a **differential equation**, and the **initial value problem**.

The notion of a *solution* is more familiar for *algebraic* equations. For example the equation $x^2 - 1 = 0$ has as its solution $x = \pm 1$, which means inserting the values $x = \pm 1$ turns the equation into a true statement.

For a *differential* equation the *solutions* are **functions**, which upon inserting turn the equation into a true statement.

Definition 15.1 (Solutions to first order differential equations)**.** Let $U \subset \mathbb{R}^2$ and

$$f : U \to \mathbb{R}$$

be a continuous real valued function.

(i) The expression

$$y' = f(x, y) \tag{15.1}$$

is called an (explicit) *first order differential equation.*

(ii) A function $\varphi : I \to \mathbb{R}$ is a *solution* to (15.1) if φ is *continuously differentiable* on an interval I (namely φ is differentiable and the derivative φ' is a continuous function; we write $\varphi \in C^1(I)$), where the domain $I \subset \mathbb{R}$ is an interval such that $(x, \varphi(x)) \in U$ for all $x \in I$, and

$$\varphi'(x) = f(x, \varphi(x))\,. \tag{15.2}$$

(iii) Let $(x_0, y_0) \in U$. We say $\varphi \in C^1(I)$ is a solution to the *initial value problem*

$$y' = f(x, y) \qquad y(x_0) = y_0 \tag{15.3}$$

if $\varphi : I \to \mathbb{R}$ is a solution in the above sense, and $x_0 \in I$ and $\varphi(x_0) = y_0$.

Remark 15.1. The general theory of differential equations addresses the question under which conditions on f there *exist* solutions to (15.1), and when they are *unique*. It is also of interest when the solutions can be expressed in explicit terms.

Remark 15.2. In practise we often write $y(x)$ for the solution, but conceptually it is important to distinguish between the *unknown* y, and the *solution* $\varphi(x)$.

Remark 15.3. We use the notation $\mathrm{C}^0(I)$ to denote "the space of continuous functions on the interval I". For example, $\varphi \in \mathrm{C}^0(I)$, where $I = (a, b)$, means that φ is continuous on (a, b). Similarly, $\mathrm{C}^1(I)$ denotes the "space of continuously differentiable functions on I". So $\varphi \in \mathrm{C}^1(I)$, where $I = (a, b)$, means that φ is differentiable on (a, b), and $\varphi' \in \mathrm{C}^0(I)$.

In Module V we have seen the exponential as the solution to an initial value problem. Let us prove this characterisation of the exponential in yet another way, and thereby give an example of an *existence and uniqueness theorem* for a differential equation.

Proposition 15.1. *Let $a, y_0 \in \mathbb{R}$. Then $f : I \to \mathbb{R}$ is a solution to the initial value problem*

$$y' = ay \qquad y(0) = y_0 \tag{15.4}$$

if and only if

$$f(x) = y_0 e^{ax} . \tag{15.5}$$

Proof. Clearly, with $f(x)$ given by (15.5) we have

$$f'(x) = y_0 a e^{ax} = af(x) \tag{15.6}$$

and $f(0) = y_0$, so $f(x)$ solves the initial value problem. Conversely, let $f(x)$ be a solution to the initial value problem (15.4), and set $g(x) = f(x)e^{-ax}$ on the interval I where f is defined. Then

$$g'(x) = f'(x)e^{-ax} - af(x)e^{-ax} = 0 . \tag{15.7}$$

Therefore (by the Mean Value Theorem) $g(x) = g(0)$, hence

$$f(x)e^{-ax} = g(x) = g(0) = f(0) = y_0 . \tag{15.8}$$

□

15.1. Linear first order equations

A special case are the **linear differential equations** of first order, which correspond to the case when the function f in Defintion 15.1 depends *linearly* on y.

Definition 15.2. Let $P, Q : I \to \mathbb{R}$ be continuous functions on an interval $I \subset \mathbb{R}$. A *linear first order differential equation* is an expression of the form

$$y' + P(x)y = Q(x) \tag{15.9}$$

and we say the equation is *homogeneous* if $Q = 0$, and *inhomogeneous* otherwise.

Assuming for a moment that the homogeneous equation (15.9), with $Q = 0$, has a solution $\varphi \neq 0$, we see that we can integrate the homogeneous equation in the form

$$\frac{\varphi'}{\varphi} = -P(x) \tag{15.10}$$

which yields for any $(x_0, x) \subset I$,

$$\log \frac{|\varphi(x)|}{|\varphi(x_0)|} = -\int_{x_0}^{x} P(t)\mathrm{d}t . \tag{15.11}$$

Proposition 15.2. *Let $I \subset \mathbb{R}$ be an interval, $P \in C^0(I)$ and $x_0 \in I$, and $y_0 \in \mathbb{R}$. Then $\varphi : I \to \mathbb{R}$ is a solution to the initial value problem*

$$y' + Py = 0 \qquad y(x_0) = y_0 \tag{15.12}$$

if and only if

$$\varphi(x) = y_0 \exp\Big[-\int_{x_0}^{x} P(t)\mathrm{d}t\Big] \tag{15.13}$$

Proof. By differentiating it is easily verified that (15.13) solves (15.12). Conversely, if $\varphi : I \to \mathbb{R}$ is a solution to (15.12), set

$$\psi(x) = \varphi(x) \exp\Big[\int_{x_0}^{x} P(t)\mathrm{d}t\Big]\,, \tag{15.14}$$

then $\psi'(x) = 0$, hence $\psi(x) = \psi(0) = y_0$. □

One can find the solutions to the *inhomogeneous* equation by the method of **variation of constants**: Consider the function

$$\varphi(x) = \varphi_0(x)e^{-G(x)} \qquad G(x) = \int_{x_0}^{x} P(t)\mathrm{d}t \tag{15.15}$$

which is obtained from the solution to the *homogeneous* equation by replacing the constant y_0 by a function $\varphi_0 \in C^1(I)$. We will now derive a condition for the function $\varphi_0(x)$ for $\varphi(x)$ to be a solution to (15.9). We compute

$$\varphi'(x) = \varphi_0'(x)e^{-G(x)} - \varphi_0(x)P(x)e^{-G(x)} = \varphi_0'(x)e^{-G(x)} - P(x)\varphi(x) \tag{15.16}$$

hence for $\varphi(x)$ to solve (15.9) we need:

$$\varphi_0' = e^{G(x)}Q(x)\,. \tag{15.17}$$

Integrating gives a formula for $\varphi_0(x)$ in terms of the known functions $P(x)$, and $Q(x)$.

Theorem 15.3. *Let $I \subset \mathbb{R}$ be an interval, P, Q continuous functions on the interval I.*

1. *Let G be a primitive of P, and H a primitive of $e^G Q$. Then φ is a solution to (15.9) if and only if*
$$\varphi(x) = e^{-G(x)}H(x)\,. \tag{15.18}$$

2. *Let $x_0 \in I$, $y_0 \in \mathbb{R}$. Then $\varphi : I \to \mathbb{R}$ is a solution to the initial value problem*
$$y' + P(x)y = Q(x) \qquad y(x_0) = y_0 \tag{15.19}$$
if and only if
$$\varphi(x) = e^{-\int_{x_0}^{x} P(t)\mathrm{d}t}\Big(y_0 + \int_{x_0}^{x} e^{\int_{x_0}^{t} P(\tau)\mathrm{d}\tau}Q(t)\mathrm{d}t\Big)\,. \tag{15.20}$$

Exercise 15.1. Prove the first part of the theorem. Also verify that (15.20) solves the initial value problem (15.19).

Example 15.1. Let us find all solutions of the equation

$$xy' + (1-x)y = e^{2x} \tag{15.21}$$

on the interval $I = (0, \infty)$.

This equation takes the form (15.9) with $P(x) = 1/x - 1$ and $Q(x) = e^{2x}/x$. We shall express the solutions φ_{y_0} in terms of the initial value y_0 at $x_0 = 1$. We first compute

$$G(x) = \int_1^x P(t)\mathrm{d}t = \log(x) - (x-1) \tag{15.22}$$

and thus $e^{-G(x)} = e^{x-1}/x$ and $e^{G(x)} = xe^{-(x-1)}$, and so

$$\varphi(x; y_0) = \frac{e^{x-1}}{x}\Big(y_0 + \int_1^x e^{-(t-1)}e^{2t}\mathrm{d}t\Big) = \frac{e^x}{x}\Big(y_0 e^{-1} + e^x - e^1\Big) \tag{15.23}$$

Note that these solutions are unbounded as x tends to 0, unless $y_0/e + 1 - e = 0$ in which case $\lim_{x\to 0}\varphi(x; y_0) = 1$. Here we used that e^x is well approximated by the linear function $1 + x$ near $x = 0$.

Problems

1. Solve the following initial value problems:
   a) $y' - 3y = e^{2x}$ with $y(0) = 0$
   b) $y' + y = e^{2x}$ with $y(0) = 1$

2. Find all solutions of $y' \sin(x) + y\cos(x) = 1$ on the interval $(0, \pi)$. Prove that exactly one of these solutions has a finite limit as $x \to 0$, and another has a finite limit as $x \to \pi$.

3. We can view the following as a linear inhomogeneous differential equation:

$$y' = x + y \tag{15.24}$$

   What are its solutions?

Note 16.

Separable differential equations

In this lecture we consider a somewhat more challenging class of differential equations: These are equations as in Definition 15.1 where the function f on the right hand side is in fact a *product* of a function of x and a function of y.

More precisely, a **separable** differential equation is an equation of the form

$$y' = f(x)g(y) \tag{16.1}$$

where f, and g are continuous functions on intervals I, and J respectively, and the corresponding **initial value problem** is

$$y' = f(x)g(y)\,, \qquad y(x_0) = y_0\,, \tag{16.2}$$

for any values $x_0 \in I$, and $y_0 \in J$.

Let us first make the assumption

$$g(y) \neq 0 \qquad y \in J \tag{16.3}$$

and illustrate the behaviour that may occur if g has a zero with an example later.

We have not proven yet that a solution to (16.2) exists, but we will obtain such a proof by first *assuming* there is a solution, and deriving an explicit formula for the solution. Indeed, suppose φ is a solution to the initial value problem (16.2), then in view of the assumption (16.3), we obtain

$$\int_{x_0}^{x} \frac{\varphi'(t)}{g(\varphi(t))}\mathrm{d}t = \int_{x_0}^{x} f(t)\mathrm{d}t \tag{16.4}$$

We can apply the substitution rule of Theorem 14.2 to write the left hand side as

$$\int_{x_0}^{x} \frac{\varphi'(t)}{g(\varphi(t))}\mathrm{d}t = \int_{y_0}^{\varphi(x)} \frac{\mathrm{d}u}{g(u)} \tag{16.5}$$

where we have used that $\varphi(x_0) = y_0$. Introducing the notation

$$G(y) = \int_{y_0}^{y} \frac{\mathrm{d}u}{g(u)}\,, \qquad F(x) = \int_{x_0}^{x} f(t)\mathrm{d}t\,, \tag{16.6}$$

we can write the relation (16.4) as

$$G(\varphi(x)) = F(x)\,. \tag{16.7}$$

The function G is continously differentiable and *strictly monotone* on J because g is continuous and $g \neq 0$ on J. Hence G has an inverse, G^{-1}, and we find

$$\varphi(x) = (G^{-1} \circ F)(x) \tag{16.8}$$

We have derived here a *necessary* condition for the solution: If there is a solution φ to the initial value problem then it must of this form. Conversely, we can now argue as promised that given f, and g, we can define the functions F, and G as in (16.6) and verify that with φ *defined by* (16.8),

$$\varphi'(x) = \frac{1}{G'((G^{-1} \circ F)(x))} F'(x) = g(\varphi(x)) f(x) \tag{16.9}$$

and $\varphi(x_0) = G^{-1}(0) = y_0$.

In summary, we have proven that the solution to the initial value problem (16.2) is precisely given by (16.8). In the same way one proves:

Theorem 16.1. *Suppose f, and g are continuous functions on intervals I, and J, respectively and $g \neq 0$ on J. Then φ is a solution to* (16.1) *if and only if*

$$\varphi(x) = (G^{-1} \circ F)(x) \qquad x \in \tilde{I} \tag{16.10}$$

where F and G are primitives of f, and $1/g$ respectively, and $\tilde{I}$ an interval.

Example 16.1. Let us consider the differential equation

$$y' = \frac{x}{y} \tag{16.11}$$

on the upper half plane $y > 0$. We sketch the direction field as in Figure 16.1.

With the above notation $f(x) = x$, and $g(y) = 1/y > 0$, and $F(x) = x^2/2$ and $G(y) = y^2/2$ are primitives of f, and $1/g$, respectively. For $y > 0$, G is indeed invertible, and the inverse given by $G^{-1}(t) = \sqrt{2t}$. Thus the solutions are

$$\varphi(x) = \sqrt{2(x^2/2 + C)} = \sqrt{x^2 + 2C} \tag{16.12}$$

for some constant C, which is defined on

$$\tilde{I} = \begin{cases} (-\infty, \infty) & \text{if } C > 0 \\ (-\infty, 0) \text{ or } (0, \infty) & \text{if } C = 0 \\ (-\infty, -\sqrt{-2C}) \text{ or } (\sqrt{-2C}, \infty) & \text{if } C < 0\,. \end{cases} \tag{16.13}$$

Note that for all solutions

$$\lim_{x \to \pm\infty} \frac{\varphi(x)}{x} = \pm 1\,, \tag{16.14}$$

which confirms the qualitative behaviour inferred from the direction field. We may sketch the solutions and see the three types of integral curves depending on whether $C < 0$, $C = 0$, and $C > 0$.

Exercise 16.1. Solve the initial value problem (16.2) for (16.11).

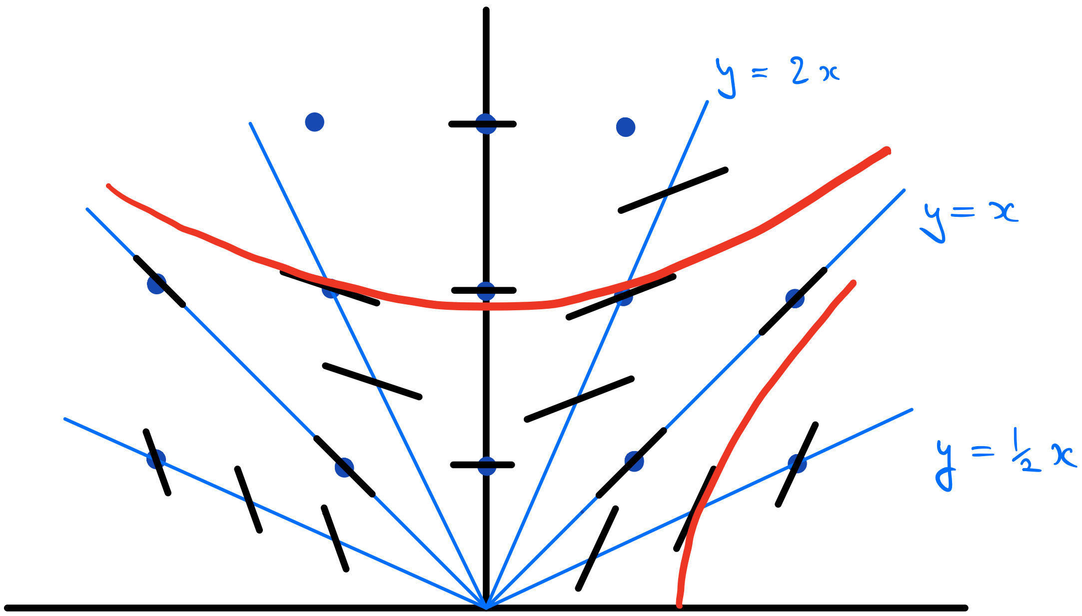

Figure 16.1.: Direction field of the equation $y' = x/y$.

Problems

1. Determine the solutions to the differential equation

$$y' = x + y\,. \tag{16.15}$$

Hint: This is not a separable equation, but it can be *reduced* to this case, by considering the equation satisfied by

$$z(x) = x + y(x)\,. \tag{16.16}$$

2. Consider a differential equation of the from

$$y' = f(y/x) \tag{16.17}$$

Show that if φ is a solution to (16.17), then

$$z(x) = \frac{\varphi(x)}{x} \tag{16.18}$$

is a solution to

$$z' = \frac{f(z) - z}{x}\,. \tag{16.19}$$

Conversely, show that if ψ is a solution to (16.19), then $\varphi(x) = x\psi(x)$ is a solution to (16.17).

Note 17.

Examples of first order equations

17.1. Examples of separable equations

We have discussed separable first order equations in general. Let us look at one more example, and others will be discussed in the tutorials.

Example 17.1. Consider the non-linear equation $xy' + y = y^2$. By inspection we see that $y = 0$, and $y = 1$ are solutions. The remaining solutions, when $y(y-1) \neq 0$, and $x \neq 0$, satisfy

$$\frac{y'(x)}{y(y-1)} = \frac{1}{x} \tag{17.1}$$

That means that a solution $y = \varphi(x)$ satisfies

$$\int \frac{\varphi'(x)}{\varphi(x)(\varphi(x)-1)} \mathrm{d}x = \int \frac{\mathrm{d}x}{x}\,. \tag{17.2}$$

We can use partial fractions to rewrite the integrand

$$\frac{1}{y(y-1)} = \frac{1}{y-1} - \frac{1}{y} \tag{17.3}$$

so by the substitution rule,

$$\int \frac{\varphi'(x)}{\varphi(x)(\varphi(x)-1)} \mathrm{d}x = \int \frac{\mathrm{d}y}{y(y-1)} = \ln|y-1| - \ln|y| \tag{17.4}$$

with $y = \varphi(x)$, and on the right hand side $\int \mathrm{d}x/x = \ln|x|$. Thus for some constant C,

$$\ln\Big|\frac{y-1}{y}\Big| = \ln|x| + C \qquad y = \varphi(x) \tag{17.5}$$

which gives $|(y-1)/y| = e^C|x|$, or $(y-1)/y = Kx$ for some constant K. We have

$$\frac{y-1}{y} - 1 = -\frac{1}{y} = Kx - 1 \tag{17.6}$$

which finally gives the formula for the solutions

$$\varphi_K(x) = \frac{1}{1-Kx} \qquad (x \in I_K) \tag{17.7}$$

where $I_K = (-\infty, 1/K)$, or $I_K = (1/K, \infty)$ depending on the parameter $K \in \mathbb{R}$, $K \neq 0$. Note that $K = 0$ corresponds to the solution $y = 1$.

Exercise 17.1. Find the solution passing through any given point (x_0, y_0). Sketch all solutions to the differential equation (17.1). Are there points with several solutions passing through them?

17.2. Reductions to separable equations

The following two types of equations can be transformed to separable equations.

The first type are equations of the form

$$y' = f(ax + by + c) \tag{17.8}$$

Here the direction field is constant on straight lines, and we can pass from $y(x)$ to the new unknown

$$u(x) = ax + by(x) + c\,. \tag{17.9}$$

Then u satisfies

$$u' = a + by' = a + bf(u) \tag{17.10}$$

which is an equation of the form $u' = g(u)$. Conversely, any solution to (17.10) gives rise to a solution of the original equation (17.8) using the relation (17.9).

Exercise 17.2. Verify this!

Example 17.2.

$$y' = (x + y)^2 \tag{17.11}$$

We find that $u(x) = x + y(x)$ satisfies $u' = 1 + u^2$, hence

$$\arctan(u(x)) = x + C \tag{17.12}$$

for some constant C. The solutions of (17.11) are thus given by

$$y(x) = \varphi_C(x) = \tan(x + C) - x\,. \tag{17.13}$$

The second type are equations of the form

$$y' = f\Big(\frac{y}{x}\Big)\,. \tag{17.14}$$

Here the transformation $u(x) = y(x)/x$ yields

$$y' = u + xu' = f(u)\,, \tag{17.15}$$

which is a separable equation for the new unknown u,

$$u' = \frac{f(u) - u}{x}\,. \tag{17.16}$$

Exercise 17.3. Verify that any solution (17.16) yields a solution to the (17.14).

This second type of equations are examples of differential equations which are *homogeneous of degree zero*; (see the **additional notes** to Module VI).

17.3. Loss of uniqueness

In our treatment of separable equation in Note 16, we have assumed that

$$y' = f(x)g(y) \qquad \text{where } g(y) \neq 0 \tag{17.17}$$

Let us consider an example where where $g(y)$ does have a zero.

Consider the differential equation

$$y' = \sqrt{|y|}\,, \tag{17.18}$$

which corresponds to the case $f(x) = 1$, and $g(y) = \sqrt{|y|}$.

Exercise 17.4. Sketch the direction field.

Consider first the region $y > 0$. Here the procedure discussed in Note 16 applies with $g(y) = \sqrt{y} > 0$. We find $F(x) = x$ and $G(y) = 2\sqrt{y}$ are primitives, and thus the solutions in the upper half plane are

$$\varphi(x) = \frac{1}{4}(x + C)^2\,, \qquad x \in (-C, \infty)\,. \tag{17.19}$$

Exercise 17.5. Show that

$$\varphi(x) = -\frac{1}{4}(-x + C)^2\,, \qquad x \in (-\infty, C)\,. \tag{17.20}$$

are the solutions in the lower half plane $y < 0$.

Note that the solutions in the lower half plane all arrive on the $y = 0$ axis and their graphs have vanishing slope there. Similarly the solutions in the upper half plane begin on the $y = 0$ axis and have vanishing slope there. Furthermore $\varphi(x) = 0$ is a solution to (17.18). This means that the solutions to (17.18) lose their uniqueness property precisely at the points where $g(y) = 0$.

Exercise 17.6. Write down an explicit family of solutions to (17.18) whose graphs pass through the point $(a, 0)$, $a \in \mathbb{R}$.

Remark 17.1. Finally note that the function $g(y)$ is continuous but not differentiable. It turns out that uniqueness holds for differential equations $y' = g(y)$ even at points where $g(y) = 0$ provided g is *differentiable*.

Problems

1. Find formulas for the solutions of the following differential equations.
   a) $(x + 1)y' + y^2 = 0$
   b) $y' = (y - 1)(y - 2)$
   c) $(x - 1)y' = xy$

2. Solve the following initial value problem:

$$y' = \frac{e^{-y^2}}{y(2x + x^2)}, \qquad y(2) = 0. \tag{17.21}$$

3. Find all solutions to the following differential equation:

$$y' = (x - y + 3)^3 \tag{17.22}$$

4. Integrate the following differential equations:
   a) $y' = -x/y$
   b) $y' = 1 + y/x$
   c) $y' = \frac{x^2+2y^2}{xy}$

Additional: Isoclines and Homogeneity

Further Reading

(Apostel, *Calculus I*, Chapter 8, Section 20-21, 23, 25–27)

Differential Equations (MAST20030)

Definition 17.1. A function $f(x, y)$ is *homogeneous of degree* λ, if

$$f(tx, ty) = t^{\lambda} f(x, y) \qquad (t > 0)\,. \tag{17.1}$$

The type of functions often appear in differential equations.

Exercise 17.1. An example with $f(x, y)$ homogeneous of degree 2 is

$$y' = x^2 + y^2\,. \tag{17.2}$$

Sketch the direction field, and determine qualitatively the solution passing through the origin $(0, 0)$.

In the construction of a direction field it is often useful to locate first those points at which the slope y' has a constant value c.

Definition 17.2. A level set of $f(x, y)$ is called an **isocline** of the differential equation $y' = f(x, y)$.

Exercise 17.2. The equation

$$y' = x + y \tag{17.3}$$

is an example of differential equation with $f(x, y)$ homogeneous of degree 1. Show that the isoclines of the differential equation form a one-parameter family of straight lines. Plot the isoclines corresponding to the constant slopes 0, $\pm 1/2$, ± 1, $\pm 3/2$, ± 2. With the aid of the isoclines construct a direction field and sketch the integral curve passing through the origin. One of the integral curves is also an isocline; find this curve.

In the context of separable equations, the case of functions which are homogeneous of degree 0 are distinguished:

Consider

$$y' = f(x, y) \tag{17.4}$$

with the property that

$$f(tx, ty) = f(x, y) \qquad (t > 0)\,, \tag{17.5}$$

namely the right hand side is invariant under rescalings of the variables (x, y). An equation with this property is said to be **homogeneous of degree zero**.

Example 17.1. The following equations are homogeneous degree zero:

$$y' = \frac{y-x}{y+x} \tag{17.6}$$

$$y' = \Big(\frac{x^2+y^2}{xy}\Big)^3 . \tag{17.7}$$

All of these equations are separable: Indeed, any first-order equation with the property (17.5) can be written as

$$y' = f(1, y/x) . \tag{17.8}$$

Exercise 17.3. Explain why.

Thus introducing a new unknown $v = y/x$, we have $y = vx$, $y' = v'x + v$, and thus (17.8) becomes again

$$xv' = f(1, v) - v \tag{17.9}$$

which is *separable*, and can be solved as in Note 16.

Exercise 17.4. Show that also the substitution $v = x/y$ transforms a differential equation which is homogeneous of degree zero into a first-order equation for v which is separable.

Example 17.2. Let us solve the differential equation (17.6). As indicated above, the equation transforms to

$$xv' = \frac{v-1}{v+1} - v = -\frac{v^2+1}{v+1}, \qquad v = y/x , \tag{17.10}$$

which we may integrate to get

$$\int \frac{v\mathrm{d}v}{v^2+1} + \int \frac{\mathrm{d}v}{v^2+1} = -\int \frac{\mathrm{d}x}{x} \tag{17.11}$$

$$\frac{1}{2}\ln(v^2+1) + \arctan(v) = -\ln|x| + C \tag{17.12}$$

which shows that for every solution $y = \varphi(x)$ there is a constant C such that

$$\frac{1}{2}\ln\Big(\varphi(x)^2 + x^2\Big) + \arctan\Big(\frac{\varphi(x)}{x}\Big) = C . \tag{17.13}$$

We have seen examples of differential equations (17.8) which are homogeneous of degree zero in the sense of (17.5). Let us explore some of their properties in greater generality.

Exercise 17.5. Show that straight lines through the origin are isoclines of the differential equations which are homogeneous of degree zero.

Exercise 17.6. Demonstrate this property for the equation $y' = -2y/x$. Sketch the isoclines and the direction field.

Exercise 17.7. Given that all straight lines are isoclines, and the slope of the direction field is unchanged along an isocline, we may guess that the integral curves are *similar*, in the sense that if

$$G = \{(x, \varphi(x)) : x \in I\} \tag{17.14}$$

is the graph of a solution, then so is

$$kG = \{(kx, k\varphi(x)) : x \in I\} \tag{17.15}$$

for *any* $k > 0$. (This transformation is a concentric scaling, or *similarity transformation.*) Prove that this is indeed true.

Example 17.3. An example of a differential equation where the geometric property of the previous exercise is quite obvious is

$$y' = -x/y \tag{17.16}$$

whose integral curves are concentric circles given by $x^2 + y^2 = C$ for some constant $C > 0$.

Note 18.

Linear differential equations of second order with constant coefficients

The equation

$$y'' = -k^2 y \tag{18.1}$$

describes the *oscillations* of a particle on a line around the origin $y = 0$, when a force proportional to its displacement pulls it back to its equilibrium position at $y = 0$.[1]

This is an example of a *second order* equation, and could we viewed as a *system of first order equations*: Introducing the unknowns $y_1 = y$ and $y_2 = y'$ we could rewrite this equation as

$$y_1' = y_2\,, \qquad y_2' = -k^2 y_1\,. \tag{18.2}$$

We can think of $\vec{y} = (y_1, y_2)$ as a point in the plane, and view the solution $\vec{\varphi}(t)$ as a point moving through 2-dimensional space $\mathbb{R}^2$ with velocity $\vec{v}$:

$$\vec{\varphi}'(t) = \vec{v}(\vec{\varphi}(t)) \tag{18.3}$$

Here $\vec{v}(y_1, y_2) = (y_2, -k^2 y_1)$ is a **vectorfield**.

Exercise 18.1. A *vectorfield* in $\mathbb{R}^2$, in comparison to a *direction field* which has a line attached to each point, has in addition to the slope also a *magnitude* at each point. In fact, with $\vec{v}(y_1, y_2) = (v_1(y_1, y_2), v_2(y_1, y_2))$ the magnitude at the point (y_1, y_2) is $|v| = \sqrt{v_1^2 + v_2^2}$. Sketch the vectorfield $\vec{v}$ of (18.2).

Exercise 18.2. Show that in the case $k = 1$ concentric circles are solutions to (18.2). The fact that $y_1^2/2 + y_2^2/2$ is constant along the solution curve can be interpreted as the law of *conservation of energy.* What is the situation if $k \neq 1$?

While recasting a second order equation as a first order system is a very fruitful point of view, we will not adopt this approach here, and study directly the *linear differential equation of second order with constant coefficients*:

$$y'' + ay' + by = 0 \tag{18.4}$$

More generally, a **linear differential equations of second order** is an expressions of the form

$$y'' + g(x)y' + h(x)y = r(x) \tag{18.5}$$

In the case that $r(x)$ vanishes identically we say the equation is *homogeneous*, otherwise *inhomogeneous.*

[1]For the given equation, this relationship is *linear*, and is then also referred to as *Hooke's law* in physics.

18.1. Existence of solutions by inspection

We consider first the equation (18.4) in the case that $a = 0$.

The equation $y'' = 0$. In this special case, when both coefficients $a = b = 0$, we see immediately by integration that any solution that satisfies this equation has constant derivative $y' = c_1$, and so integrating again

$$y(x) = c_1 x + c_2 \tag{18.6}$$

with constant c_2. Conversely, for any numbers $c_1, c_2 \in \mathbb{R}$, the function given by (18.6) is a solution, so we have found all solutions in this case.

The equation $y'' + by = 0$, when $b < 0$. Since $b < 0$, we can write $b = -k^2$ for some $k > 0$, and the differential equation takes the form

$$y'' = k^2 y\,. \tag{18.7}$$

We immediately verify that $y(x) = e^{kx}$ is a solution, and another is $y(x) = e^{-kx}$. Therefore also linear combinations of these are solutions, and we conclude: For any constants $c_1, c_2 \in \mathbb{R}$

$$y(x) = c_1 e^{kx} + c_2 e^{-kx} \tag{18.8}$$

is a solution. We will prove below that these are in fact *all* solutions in this case.

The equation $y'' + by = 0$, when $b > 0$. In this case let us again write $b = k^2$, then the equation (18.4) takes the form

$$y''(x) = -k^2 y\,. \tag{18.9}$$

One may recognize that this relation is satisfied by the function $y(x) = \cos(kx)$, and also $y(x) = \sin(kx)$. We find again a general solution by forming a linear combination:

$$y(x) = c_1 \cos(kx) + c_2 \sin(kx) \tag{18.10}$$

This does *not* show that any solution of (18.9) is of this form, but we have actually already given a proof of that in Note 12 in the case $k = 1$.

Exercise 18.3. Can you generalise the argument we have given in Lemma 12.2 to show that *any* solution to (18.9) is of the form (18.10)?

The case $a \neq 0$. The cases considered above actually cover the case $a \neq 0$ as well, in the sense that we can *reduce* the problem (18.4) to the problem with $a = 0$.

The idea is to consider solutions of (18.4) of the form

$$y(x) = u(x)v(x)\,. \tag{18.11}$$

A simple calculation gives that

$$y'' + ay' + by = (v'' + av' + bv)u + (2v' + av)u' + vu'' \tag{18.12}$$

and so by choosing v such that $v' = -av/2$ — which can evidently be arranged by choosing $v(x) = e^{-ax/2}$ — the coefficient to the u' term disappears, and with this choice the equation reduces to

$$y'' + ay' + by = \Big(u'' + \frac{4b - a^2}{4}u\Big)v\,. \tag{18.13}$$

Since $v(x) = e^{-ax/2} \neq 0$, we have shown:

Proposition 18.1. *Let $y(x) = u(x)\exp(-ax/2)$. Then the function y satisfies* (18.4) *if and only if the function u satisfies the differential equation*

$$u'' + \frac{4b - a^2}{4}u = 0 \tag{18.14}$$

18.2. General form of the solutions

Theorem 18.2. *Let $b \in \mathbb{R}$. Every solution $\varphi(x)$ of the equation*

$$y'' + by = 0 \tag{18.15}$$

is a linear combination of two functions f_1, f_2, where

- *when $b = 0$: $f_1(x) = 1$, $f_2(x) = x$*
- *when $b < 0$: set $b = -k^2$ and $f_1(x) = e^{kx}$, $f_2(x) = e^{-kx}$*
- *when $b > 0$: set $b = k^2$ and set $f_1(x) = \cos(kx)$, and $f_2(x) = \sin(kx)$.*

In other words, for every solution $\varphi(x)$ of the equation (18.15) *there are constants $c_1, c_2 \in \mathbb{R}$ so that*

$$\varphi(x) = c_1 f_1(x) + c_2 f_2(x)\,. \tag{18.16}$$

Proof. Let $\varphi(x)$ be a solution to (18.15), and let $f_1(x)$, and $f_2(x)$ be chosen as in the statement of the theorem, depending on the value of b. Then also

$$\psi(x) = \varphi(x) - c_1 f_1(x) - c_2 f_2(x) \tag{18.17}$$

is a solution to (18.15).

Let us choose the constants c_1, and c_2 so that $\psi(0) = 0$, and $\psi'(0) = 0$. This amounts to solving the equations:

$$\varphi(0) = c_1 f_1(0) + c_2 f_2(0) \tag{18.18a}$$

$$\varphi'(0) = c_1 f_1'(0) + c_2 f_2'(0) \tag{18.18b}$$

In the case $b = 0$, these equations reduce to

$$\varphi(0) = c_1 \qquad \varphi'(0) = c_2 \,. \tag{18.19}$$

Now in the case $b < 0$, these two equations read:

$$\begin{aligned} \varphi(0) =& c_1 + c_2 \\ \varphi'(0) =& c_1 k - c_2 k \end{aligned}$$

which we can solve to find

$$\begin{aligned} c_1 =& \frac{\varphi(0) + \varphi'(0)/k}{2} \\ c_2 =& \frac{\varphi(0) - \varphi'(0)/k}{2} \end{aligned}$$

Similary in the case $b > 0$.

Thus it remains to show that the only solution to the initial value problem

$$y'' + by = 0 \qquad y(0) = 0 \qquad y'(0) = 0 \tag{18.20}$$

is the trivial solution $y = 0$.

In the case $b = 1$ we have already demonstrated this in Note 12; see Lemma 12.2.

Exercise 18.4. Generalize this argument to the case $b > 0$.

In the case $b = 0$ this follows from integrating twice.

Exercise 18.5. Show that if f is twice differentiable and satisfies $f''(x) = 0$, and moreover $f(0) = f'(0) = 0$, then $f(x) = 0$.

The case $b < 0$ is covered by the more general considerations below; (see the **additional notes** to Module VI).

Exercise 18.6. Read the additional material on uniqueness of solutions to the initial value problem and give a proof of the above statement that the only solution to (18.20) is the trivial solution, in the special case that $b = -k^2$.

□

Exercise 18.7. In view of Proposition 18.1 it is now also possible to state a theorem that gives a characterisation of all solutions to (18.4), for any values of $a, b \in \mathbb{R}$, depending on the sign of the so-called **discriminant** of the differential equation (18.4):

$$\Delta = a^2 - 4b \,. \tag{18.21}$$

State it!

Problems

1. Find explicitly the solution to the following initial value problem:
   a) $y'' + ky = 0$, $y(0) = 0$, $y'(0) = y_1$, where $k > 0$, and $y_1 \in \mathbb{R}$ are fixed constants.
   b) $y'' + ay' = 0$, $y(0) = 1$, $y'(0) = 0$, where $a \in \mathbb{R}$ is fixed.

2. Show that the solution (18.16) in the case $b > 0$ can also be written as

$$\varphi(x) = c \sin(kx + \delta) \tag{18.22}$$

   where $c = \sqrt{c_1^2 + c_2^2}$, and δ is to be determined.

3. Find all solutions to the following differential equations:
   a) $y'' + 4y' = 0$
   b) $y'' + 2y' + y = 0$

4. Find all values of the constant k so that $y'' + ky = 0$ has a non-trivial solution $y = \varphi_k(x)$ for which $\varphi_k(0) = \varphi_k(1) = 0$.

Additional: Uniqueness of solutions to the initial value problem

We will give a proof of Theorem 18.2 above, in a way that applies to all cases, and already outlines an approach to the general second order differential equation (18.5) (not necessarily with constant coefficients).

In this section we will prove that there exists *one and only one* solution to the following

Initial value problem. Let $a, b \in \mathbb{R}$, and $x_0, y_0, y_1 \in \mathbb{R}$. The initial value problem for (18.4) is to find a solution $\varphi \in C^2(\mathbb{R})$ to the problem

$$y'' + ay' + by = 0\,, \tag{18.1a}$$

$$y(x_0) = y_0 \qquad y'(x_0) = y_1\,. \tag{18.1b}$$

In particular, we want to prove that the *only* solution to this initial value problem with $x_0 = y_0 = y_1 = 0$ is $y = 0$.

We begin with the following observation:

Proposition 18.1 (Wronskian determinant)**.** *Let $a, b \in \mathbb{R}$ and $\varphi_1, \varphi_2 \in C^2(\mathbb{R})$ be any two solutions to the differential equation* (18.1a)*. Let the* Wronskian *be defined as the determinant*

$$W(\varphi_1, \varphi_2) := \begin{vmatrix} \varphi_1 & \varphi_2 \\ \varphi_1' & \varphi_2' \end{vmatrix} = \varphi_1\varphi_2' - \varphi_1'\varphi_2 \tag{18.2}$$

then for some $c \in \mathbb{R}$ it holds:

$$W(\varphi_1, \varphi_2)(x) = ce^{-ax}\,. \tag{18.3}$$

Note this means in particular that *either* $W(\varphi_1, \varphi_2)(x) = 0$ for *all* $x \in \mathbb{R}$ *or none.*

Proof. Let us denote for short by $W(x) = W(\varphi_1, \varphi_2)(x)$ and note that $W \in C^1(\mathbb{R})$ because $\varphi_1, \varphi_2 \in C^2(\mathbb{R})$. We have

$$W'(x) = \varphi_1\varphi_2'' - \varphi_1''\varphi_2 = \varphi_1(-a\varphi_2' - b\varphi_2) - (-a\varphi_1' - b\varphi_1)\varphi_2 = -aW(x) \tag{18.4}$$

therefore by Proposition 15.1,

$$W(x) = W(0)e^{-ax}\,. \tag{18.5}$$

□

Exercise 18.1. With f_1, and f_2 as defined in Theorem 18.2, compute $W(f_1, f_2)$ in all cases $b > 0$, $b = 0$, $b < 0$. In particular, verify that $W(f_1, f_2) \neq 0$ in all cases!

Theorem 18.2 (Uniqueness of solutions)**.** *Let $b \in \mathbb{R}$, and $x_0, y_0, y_1 \in \mathbb{R}$, and suppose $\varphi_1, \varphi_2 \in C^2(\mathbb{R})$ are two solutions to the initial value problem*

$$y'' + by = 0\,, \qquad y(x_0) = y_0\,, \qquad y'(x_0) = y_1\,. \tag{18.6}$$

Then

$$\varphi_1(x) = \varphi_2(x) \qquad x \in \mathbb{R}\,. \tag{18.7}$$

Proof. The difference $\varphi(x) = \varphi_1(x) - \varphi_2(x)$ is a solution to the same equation with trivial initial values, namely $\varphi(x_0) = \varphi'(x_0) = 0$. Assume that $\varphi(x_1) \neq 0$ for some $x_1 \in \mathbb{R}$.

The idea is now to write down a solution to the initial value problem

$$y'' + by = 0\,, \qquad y(x_1) = 0\,, \qquad y'(x_1) = \varphi(x_1)\,. \tag{18.8}$$

We can do this as follows: We know that

$$\psi(x) = c_1 f_1(x) + c_2 f_2(x) \tag{18.9}$$

is a solution, and we can arrange for the initial conditions by choosing

$$c_1 f_1(x_1) + c_2 f_2(x_1) = 0 \tag{18.10}$$
$$c_1 f_1'(x_1) + c_2 f_2'(x_1) = \varphi(x_1) \tag{18.11}$$

This system is solvable for c_1, and c_2, because

$$\det \begin{pmatrix} f_1(x_1) & f_2(x_1) \\ f_1'(x_1) & f_2'(x_1) \end{pmatrix} = W(f_1, f_2)(x_1) \neq 0 \tag{18.12}$$

In fact, we obtain

$$c_1 = -\frac{\varphi(x_1) f_2(x_1)}{W(x_1)} \qquad c_2 = \frac{\varphi(x_1) f_1(x_1)}{W(x_1)}\,. \tag{18.13}$$

Since φ and ψ are solutions to $y'' + by = 0$ we know from Proposition 18.1 that the Wronskian determinant either vanishes for all $x \in \mathbb{R}$ or none. However,

$$\begin{aligned} W(\varphi, \psi)(x_0) &= \varphi(x_0)\psi'(x_0) - \varphi'(x_0)\psi(x_0) = 0 \\ W(\varphi, \psi)(x_1) &= \varphi(x_1)\psi'(x_1) - \varphi'(x_1)\psi(x_1) = \varphi(x_1)^2 > 0\,, \end{aligned}$$

which is a contradiction, hence $\varphi(x) = \varphi_1(x) - \varphi_2(x) = 0$ for all $x \in \mathbb{R}$. □

Remark 18.1. In the proof we have made the following important observation: In order to find the solution to the initial value problem (18.1a,18.1b) it suffices to find *any* two solutions of (18.1a) whose Wronski determinant does not vanish at a point. Indeed, if $\varphi_1, \varphi_2 \in C^2(\mathbb{R})$ are two solutions to (18.1a) such that $W(x_1) \neq 0$ for some $x_1 \in \mathbb{R}$ where

$W(x) = W(\varphi_1, \varphi_2)(x)$, then $W(x) \neq 0$ for all $x \in \mathbb{R}$ by virtue of Proposition 18.1, and we can set for any $x_0, y_0, y_1 \in \mathbb{R}$,

$$c_1 = \frac{1}{W(x_0)}(y_0\varphi_2'(x_0) - y_1\varphi_2(x_0)) \tag{18.14a}$$

$$c_2 = \frac{1}{W(x_0)}(-y_0\varphi_1'(x_0) + y_1\varphi_1(x_0))\,. \tag{18.14b}$$

Then, with this choice of constants,

$$\varphi(x) = c_1\varphi_1(x) + c_2\varphi_2(x) \tag{18.15}$$

is a solution to the initial value problem (18.1b).

Exercise 18.2. Verify this claim!

Exercise 18.3. The proof of the uniqueness of solutions we have given actually equally applies to the initial value problem (18.1a), (18.1b), when $a \neq 0$. Indeed, Proposition 18.1 is valid in that case! Find the part of the proof of Theorem 18.2 where we have actually used that $a = 0$, and adapt it to include the case $a \neq 0$!

Problems

1. Let φ_1, and φ_2 be two solutions to differential equation (18.1a), and assume φ_1 is not identically zero.
   a) Prove that $W(\varphi_1, \varphi_2)(0) = 0$ if and only if φ_2/φ_1 is constant.
   b) Suppose φ_2/φ_1 is not constant. Let φ be any solution to (18.1a). Show that there exists constants c_1, c_2 such that

   $$c_1\varphi_1(0) + c_2\varphi_2(0) = \varphi(0) \qquad c_1\varphi_1'(0) + c_2\varphi_2'(0) = \varphi'(0)\,. \tag{18.16}$$

   Moreover, show that

   $$\varphi(x) = c_1\varphi_1(x) + c_2\varphi_2(x)\,. \tag{18.17}$$

Additional: The space of solutions

Further Reading

(Apostel, *Calculus I*, Section 8.1-8.4) and (Apostel, *Calculus I*, Section 8.8-8.17)

Differential Equations (MAST20030)

First order equations

One can phrase the statement of Proposition 15.2 in the languge of **Linear Algebra** as follows: A **linear map** $T : C^1(I) \to C^0(I)$ is defined by

$$T(y) = y' + Py\,. \tag{18.1}$$

A continuously differentiable function φ is a solution to the homogeneous equation (15.12) if and only if it lies in the **kernel** of T, $T(\varphi) = 0$. We have seen that the kernel as a subspace of $C^1(I)$ is *one dimensional*, and spanned by e^{-G} where G is a primitive of P:

$$\ker(T) = \{\varphi \in C^1(I) : \varphi = ce^{-G} \text{ for some } c \in \mathbb{R}\} \tag{18.2}$$

Similarly φ is a solution to the inhomogeneous equation (15.9) if and only if $T(\varphi) = Q$. According to Theorem 15.3 we have that the space of solutions is

$$\{\varphi \in C^1(I) : \varphi = \varphi_{in} + Ce^{-G}, C \in \mathbb{R}\} = \varphi_{in} + \ker(T) \tag{18.3}$$

where φ_{in} is a solution to the inhomogeneous problem (15.9). The set of solutions to the inhomogeneous equation is thus span an *affine subspace* of $C^1(I)$ of dimension 1.

Second order equations

It is useful to use the language of linear algebra to describe the set of solutions to the homogeneous equation (18.4).

Theorem 18.1 (Space of solutions)**.** *Let $a, b \in \mathbb{R}$, and $T : C^2(\mathbb{R}) \to C^0(\mathbb{R})$ be defined by*

$$T(f) = f'' + af' + bf \tag{18.4}$$

Then we have

(i) *T is a* linear *map.*

(ii) *For $f \in C^2(\mathbb{R})$, f is a solution to* (18.1a) *if and only if $f \in \ker(T)$. In particular the set of solutions to the homogeneous second order differential equation with constant coefficients is a* linear subspace *of $C^2(\mathbb{R})$.*

(iii) *If $f_1, f_2 \in \ker T$ then f_1, f_2 are* linearly independent *if and only if $W(f_1, f_2)(x) \neq 0$ for some $x \in \mathbb{R}$ (or equivalently $W(f_1, f_2)(x) = 0$ for all $x \in \mathbb{R}$).*

(iv) $\dim \ker(T) = 2$.

(v) *If $f_1, f_2 \in \ker(T)$ are linearly independent, then we say $\{f_1, f_2\}$ are a* fundamental system *for the differential equation* (18.1a)*, and*

$$\ker(T) = \{f \in C^2(\mathbb{R}) : \textit{there exists constants } c_1, c_2 \in \mathbb{R} \textit{ such that } f = c_1 f_1 + c_2 f_2\} \tag{18.5}$$

is the space of solutions to (18.1a) *defined on* $\mathbb{R}$.

We omit most of the proof except for *(iii,iv)* which are instructive.

Proof of (iii). If f_1, f_2 are linearly dependent, say $f_1 = \lambda f_2$, then clearly $W(f_1, f_2) = 0$. Conversely suppose that $W(f_1, f_2) = 0$. If either f_1 or f_2 vanish identically then they are linearly dependent, so we can assume that $f_1 \neq 0$ and $f_2 \neq 0$. (Meaning f_1, f_2 are not the "zero function" $0(x) = 0$, $x \in \mathbb{R}$.) Then by Theorem 18.2 f_1 and f_2 cannot have trivial initial values at $x = 0$. Consider the case that $f_1'(0) = f_2'(0) = 0$. Then we must have $f_1(0) \neq 0$, and $f_2(0) \neq 0$, and we can define

$$f(x) = f_2(0) f_1 - f_1(0) f_2 \in \ker(T) \tag{18.6}$$

which satisfies $f(0) = 0$ and $f'(0) = W(f_1, f_2)(0) = 0$, so $f = 0$ again by Theorem 18.2, which shows that f_1, f_2 are linearly dependent. In the case that either $f_1'(0) \neq 0$ or $f_2'(0) \neq 0$ we can define

$$f = f_2'(0) f_1 - f_1'(0) f_2 \in \ker(T) \tag{18.7}$$

and find $f(0) = W(0) = 0$ and $f'(0) = 0$ which implies that $f = 0$ by Theorem 18.2 hence f_1, f_2 are also linearly dependent in this case.

□

Proof of (iv). Let f_1, f_2 be the solutions to (18.1b) with $(y_0, y_1) = (1, 0)$, and $(y_0, y_1) = (0, 1)$ respectively (and $x_0 = 0$). Then $W(f_1, f_2)(0) = 1$ and so f_1, f_2 are linearly independent by (iii). Hence $\dim \ker(T) \geq 2$. Given $f \in \ker(T)$, then we claim that $f = f(0) f_1 + f'(0) f_2$. Indeed $g = f - f(0) f_1 + f'(0) f_2 \in \ker(T)$, and $g(0) = f(0) - f(0) = 0$, and $g'(0) = f'(0) - f'(0) = 0$, hence $g = 0$ identically by Theorem 18.2. Thus we have shown that f_1, f_2 is a basis for $\ker(T)$, so $\dim \ker(T) = 2$. □

Module VII.

Complex numbers and complex exponentials

Note 19.

Complex numbers

In Module VI we have found the solutions to the differential equation $y'' + by = 0$ "by inspection." What that means is that we have taken a guess that a certain function is a solution, or as one also says *we make the ansatz*:

$$f(x) = e^{\lambda x} \qquad (x \in \mathbb{R}) \tag{19.1}$$

and we hope to find a number λ so that f is indeed a solution.[1]

Now inserting this ansatz into the differential equation we find that f is a solution to $y'' + by = 0$ if

$$\lambda^2 + b = 0 \tag{19.2}$$

In the case when $b < 0$ there are two solutions, namely $\lambda = \pm\sqrt{|b|}$, and they correspond to the two fundamental solutions $e^{\pm\sqrt{|b|}x}$. But what about the case $b > 0$? Does the equation

$$\lambda^2 + 1 = 0 \tag{19.3}$$

have a solution? There is no *real number* that solves this equation, but as we will see in this lecture it does have the **complex** solutions $\lambda = i$ and $\lambda = -i$. The corresponding fundamental solutions $e^{\pm ix}$ — of which we still have to make sense of! — are in fact related to the *trigonometric functions* that we have encountered in Section 18.1 as the fundamental solutions to $y'' + by = 0$ when $b > 0$.

More generally, we verify that the ansatz (19.1) gives a solution to (18.4) if λ is a zero of the **characteristic polynomial**:

$$\lambda^2 + a\lambda + b = 0 \tag{19.4}$$

Depending on the sign of the discriminant Δ — see (18.21) — this equation has either real solutions ($\Delta > 0$, or $\Delta = 0$) or complex roots ($\Delta < 0$), corresponding to the three systems of fundamental solutions that we have encountered in Theorem 18.2.

[1] This of course does not explain *how* one arrives at this particular ansatz. But it is clear that polynomials cannot provide a solution, because differentiation reduces its order. The exponential function already led to success for the equation $y' = y$, so why not try again! This procedure of trial and error is also called *heuristics*, which comes from Greek for "searching."

19.1. Imaginary numbers and quadratic equations

Since $\lambda^2 + 1 > 0$ for all $x \in \mathbb{R}$ there cannot be a real number that solves (19.3). The name "imaginary number" reflects that at first the "number" i that solves $i^2 + 1 = 0$ was *invented.* While mysterious, this number allows us to solve every quadratic equation:

$$ax^2 + bx + c = 0 \tag{19.5}$$

Formally the solutions are

$$x = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \qquad x = \frac{-b - \sqrt{b^2 - 4ac}}{2a}, \tag{19.6}$$

which seems to make sense even if $\Delta = b^2 - 4ac < 0$. Consider for example the case $a = b = c = 1$, then this formula says

$$x = \frac{-1 + \sqrt{-3}}{2} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i, \qquad x = \frac{-1 - \sqrt{-3}}{2} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i \tag{19.7}$$

where we have taken $\sqrt{-3} = \sqrt{3i^2} = \sqrt{3}i$.

Exercise 19.1. Verify that these are formally solutions to $x^2 + x + 1 = 0$.

These are examples of **complex numbers**, which can always be written as $a + bi$ for some *real numbers* a, and b, with the convetion that $a = a + 0i$, and $i = 0 + 1i$. The laws of arithmetic and the relation $i^2 = -1$ then show that

$$(a + bi) + (c + di) = (a + c) + (b + d)i \tag{19.8a}$$

$$(a + bi) \cdot (c + di) = (ac - bd) + (ad + bd)i \tag{19.8b}$$

Remark 19.1. An equation like $(1 + 2i)(3 + 1i) = 1 + 7i$ may thus be regarded as an abbreviation for two equations $1 \cdot 3 - 2 \cdot 1 = 1$, and $1 \cdot 1 + 2 \cdot 3 = 7$.

Any complex number $a + bi \neq 0$ has a (multiplicative) inverse denoted by $(a + bi)^{-1}$, which is given by:

$$\frac{1}{a + bi} = \frac{1}{a + bi}\frac{a - bi}{a - bi} = \frac{a - bi}{a^2 + b^2} \tag{19.9}$$

19.2. Complex plane

It is clear that we can view complex numbers $z = a + bi$ as points (a, b) in the plane, where the points $(a, 0)$ on the horizontal axis are called the *real axis*, and the vertical axis consisting of the points $(0, b)$ is called the *imaginary axis.* In fact, we can use this point of view to *define* complex numbers ; (see the **additional notes** to Module VII).

Definition 19.1. For any complex number $z = x + iy$ $(x, y \in \mathbb{R})$, the **conjugate** is defined by

$$\overline{z} = x - iy, \tag{19.10}$$

and the **modulus** (or **absolute value**) is defined as

$$|z| = \sqrt{x^2 + y^2}. \tag{19.11}$$

Exercise 19.2. Interpret these definitions geometrically.

Exercise 19.3. Let z, and w be complex numbers. Show that

1. $\overline{\overline{z}} = z$
2. $\overline{z} = z$ if z is real
3. $\overline{z+w} = \bar{z} + \bar{w}$
4. $\overline{-z} = -\overline{z}$
5. $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$
6. $\overline{z^{-1}} = \overline{z}^{-1}$
7. $|z|^2 = z\overline{z}$
8. $|z \cdot w| = |z| \cdot |w|$

A less straight-forward but very important statement is the following triangle inequality; (see the **additional notes** to Module VII): For any complex numbers z, and w it holds

$$|z + w| \leq |z| + |w| \,. \tag{19.12}$$

Addition and multiplication both have geometric interpretations in the complex plane.

For the interpretation of multiplication note that for any complex number $z \neq 0$ we can write

$$z = |z| \frac{z}{|z|} \tag{19.13}$$

where $|z|$ is a positive factor, and $z/|z|$ is a complex number with unit absolute value. Since any complex number with unit modulus can be written in the form $\cos\theta + i\sin\theta$, we see that

$$z = r(\cos\theta + i\sin\theta) \tag{19.14}$$

where $r = |z| > 0$, and $\theta \in \mathbb{R}$ (which is *not* unique, because if θ_0 is one possibility, then so are $\theta + 2\pi k$, for any $k \in \mathbb{Z}$); θ is called the **argument** of z.

Exercise 19.4. Show that the product of two nonzero complex numbers $z = r(\cos\theta + \sin\theta)$, and $w = s(\cos\phi + i\sin\phi)$ is

$$z \cdot w = rs(\cos(\theta + \phi) + i\sin(\theta + \phi)) \tag{19.15}$$

and give a geometric interpretation in the complex plane.

It is an easy matter to derive from (19.15), by induction, a formula for z^n, for any given complex number z: If z is given by (19.14), then

$$z^n = |z|^n \big(\cos(n\theta) + i\sin(n\theta)\big) \,. \tag{19.16}$$

This formula is also known as de Moivre's theorem, and can be used to compute the nth roots of a complex number; (see the **additional notes** to Module VII).

Problems

1. Find the absolute value and argument(s) of each of the following complex numbers:
   a) $3 + 4i$
   b) $(3 + 4i)^{-1}$
   c) $(1 + i)^5$
2. Solve the following equations:
   a) $x^2 + ix + 1 = 0$
   b) $x^4 + x^2 + 1 = 0$
   c) $x^2 + 2ix - 1 = 0$
3. Describe the set of all complex numbers z such that
   a) $\bar{z} = -z$
   b) $\bar{z} = z^{-1}$
   c) $|z - a| = |z - b|$
4. Prove that $|z| = |\bar{z}|$ and that the real part of z can be written as $(z + \bar{z})/2$, while the imaginary part is $(z - \bar{z})/2i$.
5. Consider a polynomial $p(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_0$ with *real* coefficients $a_0, \ldots, a_{n-1} \in \mathbb{R}$. Suppose $a + ib$ (with real a, b) is a root of p, $p(a + ib) = 0$. Show that then also $a - ib$ is a root of the polynomial. (Thus the complex roots of the equation $p(z) = 0$ always occur in pairs, and the number of such roots is even.)
6. Find all the 4th roots of i.

Additional: Complex numbers

Further Reading

(Spivak, *Calculus*, Chapter 25) or (Apostel, *Calculus I*, Chapter 9)

The discussion suggests that we can arrive at a sensible definition of complex numbers if we view them as *pairs of real numbers*:

Definition 19.1 (Complex numbers)**.** A **complex number** z is an ordered pair of real numbers (a, b), where a is called the **real part**, and b the **imaginary part** of z. The set of complex numbers is denoted by $\mathbb{C}$. If (a, b) and (c, d) are complex numbers, we define addition and multiplication by (19.8), namely

$$(a, b) + (c, d) = (a + c, b + d) \tag{19.1a}$$

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc) \tag{19.1b}$$

Moreover, we define

$$i = (0, 1) \tag{19.2}$$

Remark 19.1. We can identify complex numbers $(a, 0)$ with the real number $a \in \mathbb{R}$. Moreover $i^2 = (0, 1) \cdot (0, 1) = (-1, 0)$, and so we have

$$(a, b) = (a, 0) + (0, b) = (a, 0) + b(0, 1) = a + bi\,. \tag{19.3}$$

We will not verify explicitly that $\mathbb{C}$ satisfy all the properties of a number system, or more precisely the *axioms of a field*; see for example (Spivak, *Calculus*, Chapter 1, and Chapter 25). But let us figure out how to compute the multiplicative inverse. For (a, b) let us find (x, y) such that

$$(a, b) \cdot (x, y) = (1, 0) \tag{19.4}$$

For this to be true we need $ax - by = 1$, and $ay + bx = 0$, which has the solutions $x = a/(a^2 + b^2)$, and $y = -b/(a^2 + b^2)$. This proves (19.9)

Proof of the triangle inequality

Proposition 19.1. *Let z, and w be complex numbers, then*

$$|z + w| \leq |z| + |w|\,. \tag{19.5}$$

Proof. It is easy to see that this is true when $z = \lambda w$ for some real number λ. So let us assume that $z \neq \lambda w$ for any $\lambda \in \mathbb{R}$, and that $w \neq 0$. Then for all $\lambda \in \mathbb{R}$,

$$\begin{aligned} 0 < |z - \lambda w|^2 =& (z - \lambda w) \cdot (\bar{z} - \lambda \bar{w}) \\ =& |z|^2 + \lambda^2 |w|^2 - \lambda(z\bar{w} + w\bar{z}) \end{aligned} \tag{19.6}$$

Since $z\bar{w} + w\bar{z}$ is real (verify this!), the right hand side is a quadratic in λ with real coefficients, which by the inequality cannot have a zero. Therefore its discriminant must be negative:

$$(z\bar{w} + w\bar{z})^2 - 4|w|^2|z|^2 < 0 \tag{19.7}$$

From this inequality it follows that

$$\begin{aligned} |z + w|^2 =& (z + w) \cdot (\bar{z} + \bar{w}) = |z|^2 + |w|^2 + z\bar{w} + w\bar{z} \\ <& |z|^2 + |w|^2 + 2|w||z| = (|z| + |w|)^2 \end{aligned} \tag{19.8}$$

which implies the stated inequality. □

Consequences of de Moivre's theorem

The formula (19.15) can be used to show by induction that for any nonzero complex number $z = r(\cos\theta + i \sin\theta)$ we have (and this holds for any argument θ of z):

$$z^n = |z|^n (\cos(n\theta) + i \sin(n\theta)) \tag{19.9}$$

This formula is also known as *de Moivre's theorem*, and has the important consequence:

Proposition 19.2. *Every nonzero complex number has exactly n complex n^{th} roots.*

Proof. The statement is that for any complex number $w = s(\cos\phi + i \sin\phi) \neq 0$, and any natural number n there are precisely n different complex numbers $z = r(\cos\theta + \sin\theta)$ satisfying $z^n = w$. So by de Moivre's theorem this happens if and only if

$$r^n = s \tag{19.10}$$

$$\cos(n\theta) + i \sin(n\theta) = \cos(\phi) + i \sin\phi \tag{19.11}$$

So from the first equation $r = \sqrt[n]{s}$, and from the second for some integer k,

$$n\theta = \phi + 2\pi k \tag{19.12}$$

which has the solutions θ_k. However, it remains to find out how many of these are *distinct*. Since any integer k can be written as $k = nq + k'$ for some integer q, and some integer k' between 0 and $n - 1$, we see that

$$\theta_k = \frac{\phi}{n} + 2\pi q + \frac{2\pi k'}{n} = \theta_{k'} + 2\pi q \tag{19.13}$$

and so θ_k and $\theta_{k'}$ are the arguments of the same root z in the complex plane. Therefore there are n distinct nth roots

$$z = \sqrt[n]{s}(\cos\theta_k + i\sin\theta_k) \qquad k = 0, 1, 2, \ldots, n-1\,. \tag{19.14}$$

□

Exercise 19.1. The proof actually shows us a method to compute the nth root of a complex number. Use it to compute the three cube roots of i.

Note 20.

Hyperbolic functions

Let us return to the differential equation

$$y'' - y = 0\,. \tag{20.1}$$

We have learned in Note 18 that every solution to (20.1) is of the from

$$y(x) = c_1 e^x + c_2 e^{-x} \tag{20.2}$$

for some constants $c_1, c_2 \in \mathbb{R}$. We now define the following special solutions, corresponding to the choice of constants $c_1 = \frac{1}{2}$, and $c_2 = \pm\frac{1}{2}$:

$$\cosh(x) := \frac{1}{2}(e^x + e^{-x}) \tag{20.3}$$

$$\sinh(x) := \frac{1}{2}(e^x - e^{-x})\,. \tag{20.4}$$

Remark 20.1. We may view these functions as the unique solutions to the initial value problem

$$y'' - y = 0 \qquad y(0) = y_0 \qquad y'(0) = y_1 \tag{20.5}$$

with $(y_0, y_1) = (1, 0)$, and $(y_0, y_1) = (0, 1)$, respectively. Indeed,

$$\cosh(0) = 1 \qquad \cosh'(0) = 0 \tag{20.6}$$

$$\sinh(0) = 0 \qquad \sinh'(0) = 1\,. \tag{20.7}$$

In particular, we find that the Wronskian at $x = 0$ is

$$W(\cosh(\cdot), \sinh(\cdot)) = \cosh(0)\sinh'(0) - \cosh'(0)\sinh(0) = 1 \tag{20.8}$$

which tells us that these two solutions are *linearly independent*, namely every solution to (20.1) can be written as a linear combination of these; cf. Theorem 18.1 in Note 18, additional material to Module VI.

20.1. Basic properties of hyperbolic functions

We easily verify the identity

$$\cosh^2(x) - \sinh^2(x) = 1 \tag{20.9}$$

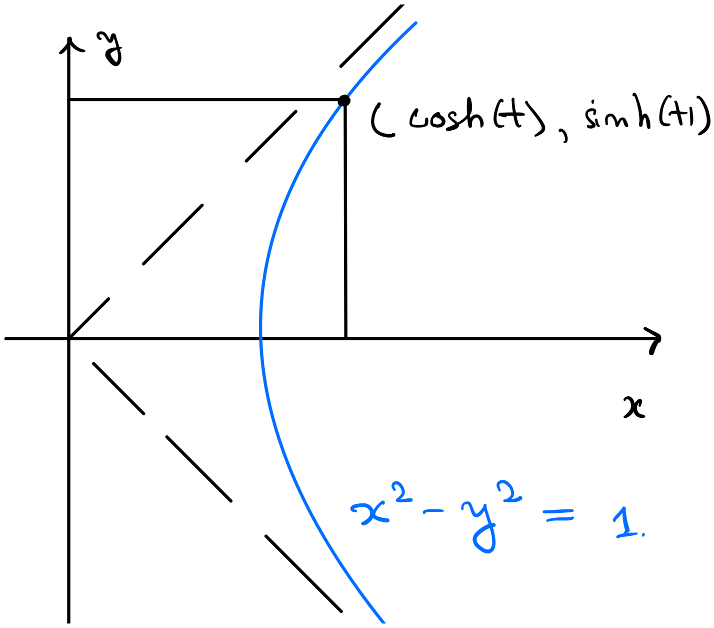

Figure 20.1.: Geometric interpretation of the hyperbolic functions.

which means geometrically that

$$\{(x, y) : x = \cosh(t), y = \sinh(t) \text{ for some } t \in \mathbb{R}\}$$

is one branch of the hyperbola

$$x^2 - y^2 = 1\,.$$

This is the origin of the name **hyperbolic functions**, and we call $\cosh(\cdot)$ the **hyperbolic cosine**, and $\sinh(\cdot)$ the **hyperbolic sine**.

Exercise 20.1. The geometric interpretation of the hyperbolic functions is not to be confused with the graphs of the hyperbolic functions. Sketch the graphs of the hyperbolic cosine and sine!

The hyperbolic cosine is an example of an *even* function, while the hyperbolic sine is an example of an *odd* function:

$$\cosh(-x) = \cosh(x) \tag{20.10}$$

$$\sinh(-x) = -\sinh(x) \tag{20.11}$$

Recall that

Definition 20.1. A function f is called **even** if $f(-x) = f(x)$ for all $x \in \mathbb{R}$, and it is called **odd** if $f(-x) = -f(x)$ for all $x \in \mathbb{R}$.

Exercise 20.2. Give other examples of even and odd functions.

Exercise 20.3. Show that for all $x, y \in \mathbb{R}$:

$$\sinh(x \pm y) = \sinh(x)\cosh(y) \pm \cosh(x)\sinh(y) \tag{20.12}$$

Also note that the hyperbolic functions are each other's derivatives:

$$\cosh'(x) = \sinh(x) \tag{20.13}$$

$$\sinh'(x) = \cosh(x) \tag{20.14}$$

20.2. Inverse hyperbolic functions

The hyperbolic sine is a strictly increasing function because $\sinh'(x) = \cosh(x) > 0$ for all $x \in \mathbb{R}$. Similarly the hyperbolic cosine is strictly increasing on the interval $[0, \infty)$. Indeed

$$\cosh'(x) = \sinh(x) = \frac{1}{2}(e^x - e^{-x}) > 0 \qquad (x > 0) \tag{20.15}$$

Therefore the inverse functions $\sinh^{-1}(x)$, and $\cosh^{-1}(x)$, exist on the intervals $[1, \infty)$, and $\mathbb{R}$, respectively, and are also called **area hyperbolic cosine** $\operatorname{arcosh}(\cdot)$, and **area hyperbolic sine** $\operatorname{arsinh}(\cdot)$.

Proposition 20.1. *The area hyperbolic functions can be expressed as*

$$\operatorname{arcosh}(x) = \log\left(x + \sqrt{x^2 - 1}\right) \qquad (x \in [1, \infty)) \tag{20.16}$$

$$\operatorname{arsinh}(x) = \log\left(x + \sqrt{x^2 + 1}\right) \qquad (x \in \mathbb{R}) \tag{20.17}$$

Proof. For $x > 1$ we have that

$$\begin{aligned} \operatorname{arcosh}'(x) &= \frac{1}{\cosh'(\operatorname{arcosh}(x))} = \frac{1}{\sinh(\operatorname{arcosh}(x))} \\ &= \frac{1}{\sqrt{\cosh^2(\operatorname{arcosh}(x)) - 1}} = \frac{1}{\sqrt{x^2 - 1}} \end{aligned} \tag{20.18}$$

We find the same result when differentiating

$$\left(\log\left(x + \sqrt{x^2 - 1}\right)\right)' = \frac{1}{x + \sqrt{x^2 - 1}}\left(1 + \frac{x}{\sqrt{x^2 - 1}}\right) = \frac{1}{\sqrt{x^2 - 1}} \tag{20.19}$$

Therefore the difference $g(x) = \operatorname{arcosh}(x) - \log(x + \sqrt{x^2 - 1})$ is a continuous function on $[1, \infty)$ whose derivative exists on $(1, \infty)$ and is $g'(x) = 0$. Therefore $g(x)$ is constant, and

$$g(x) = g(1) = \operatorname{arcosh}(1) + \log(1) = \operatorname{arcosh}(\cosh(0)) + 0 = 0 \qquad (x \geq 1) \tag{20.20}$$

□

Exercise 20.4. Write down an analogous proof for (20.17).

Exercise 20.5. The **hyperbolic tangent** is defined by

$$\tanh(x) := \frac{\sinh(x)}{\cosh(x)} \qquad (x \in \mathbb{R}) \tag{20.21}$$

Show that $\tanh(\cdot)$ takes values in $(-1, 1)$, and

$$\tanh'(x) = 1 - \tanh^2(x) = \frac{1}{\cosh^2(x)} \tag{20.22}$$

Sketch the graph of the hyperbolic tangent, and show that $\tanh(\cdot)$ is invertible. The inverse function $\operatorname{artanh}(\cdot)$ is defined on $(-1, 1)$ and is called the **area hyperbolic tangent**. Verify that

$$\operatorname{artanh}'(x) = \frac{1}{1 - x^2} \qquad (x \in (-1, 1))\,, \tag{20.23}$$

and recall the partial fraction decomposition of Note 14 to write

$$\frac{1}{1 - x^2} = \frac{1}{2}\Big[\frac{1}{1 - x} + \frac{1}{1 + x}\Big] \tag{20.24}$$

Finally integrate this identity to find an explicit formula for the area hyperbolic tangent function in terms of the logarithm.

We have seen that

$$\int \frac{1}{\sqrt{x^2 - 1}}\mathrm{d}x = \operatorname{arcosh}(x) \tag{20.25}$$

$$\int \frac{1}{\sqrt{1 + x^2}}\mathrm{d}x = \operatorname{arsinh}(x) \tag{20.26}$$

In general, the primitive of integrals involving $\sqrt{x^2 - 1}$ may often be found using the substitution $x = \cosh(u)$; cf. Note 14. Similarly the substitution $x = \sinh(u)$ often works for integrals involving $\sqrt{x^2 + 1}$.

Example 20.1. To find the primitive

$$\int \frac{\mathrm{d}x}{x\sqrt{x^2 - 1}} \tag{20.27}$$

we make the substitution $x = \cosh(u)$, then $\mathrm{d}x = \sinh(u)\mathrm{d}u$, so

$$\int \frac{\mathrm{d}x}{x\sqrt{x^2 - 1}} = \int \frac{\mathrm{d}u}{\cosh(u)} \tag{20.28}$$

and using the definition of the hyperbolic cosine, together with the substitution $e^u = y$, $e^u\mathrm{d}u = \mathrm{d}y$ we find

$$\int \frac{\mathrm{d}u}{\cosh(u)} = \int \frac{2\mathrm{d}u}{e^u + e^{-u}} = \int \frac{2e^u\mathrm{d}u}{1 + e^{2u}} = \int \frac{2\mathrm{d}y}{1 + y^2} = \arctan(y) \tag{20.29}$$

and thus in view of Proposition 20.1, with $u = \operatorname{arcosh}(x)$,

$$\int \frac{\mathrm{d}x}{x\sqrt{x^2 - 1}} = \arctan(e^u) = \arctan\big(x + \sqrt{x^2 - 1}\big)\,. \tag{20.30}$$

20.3. The catenary curve

A famous application of hyperbolic functions is the solution the following problem:

Catenary problem What shape does a chain of uniform density take under its own weight when suspended between two points?

The solution to this problem invokes the theory of separable differential equations from Note 16, and involves hyperbolic functions as discussed above; (see the **additional notes** to Module VII).

Problems

1. Compute the following integrals using hyperbolic substitutions:

   a)
   $$\int \frac{\mathrm{d}x}{x\sqrt{1+x^2}} \,. \tag{20.31}$$

   b)
   $$\int \frac{\mathrm{d}x}{x^2\sqrt{x^2-1}} \tag{20.32}$$

   c)
   $$\int \sqrt{x^2+1}\mathrm{d}x \tag{20.33}$$

   d)
   $$\int \sqrt{x^2-1}\mathrm{d}x \tag{20.34}$$

Note 21.

Second order differential equations

21.1. Homogeneous equation and complex exponentials

We have considered the special case

$$y'' - y = 0 \tag{21.1}$$

of a linear differential equation, which we solved by making the *ansatz* $f(x) = e^{\lambda x}$. This led to the equation

$$\lambda^2 - 1 = 0 \tag{21.2}$$

which has the solutions $\lambda = \pm 1$, corresponding to the linearly independent solutions $e^{\pm x}$. Linear combinations of these define the hyperbolic functions

$$\cosh(x) = \frac{1}{2}(e^x + e^{-x}) \qquad \sinh(x) = \frac{1}{2}(e^x - e^{-x}) \tag{21.3}$$

which we have studied in Note 20.

Recall from Note 18 that the special case

$$y'' + y = 0 \tag{21.4}$$

has the trigonometric functions as its solutions. The same ansatz $f(x) = e^{\lambda x}$ leads — as we have already seen in Lecture 19 — to the equation

$$\lambda^2 + 1 = 0 \tag{21.5}$$

which now has the *complex* solutions $\lambda = \pm i$, corresponding to the *complex valued* solutions $e^{\pm ix}$.

These complex valued functions need to be discussed separately; (see the **additional notes** to Module VII). However, given that $e^{\pm ix}$ is a complex valued solution, we can already infer that their *real and imaginary parts* are also solutions. Since for any complex number w the real part is given by $\mathrm{Re}(w) = (w + \bar{w})/2$, and the imaginary part is given by $\mathrm{Im}(w) = (w - \bar{w})/2i$, and given that moreover $\overline{e^z} = e^{\bar{z}}$ for any complex number z, we obtain that the real and imaginary parts of e^{ix} are given, respectively, by

$$\mathrm{Re}(e^{ix}) = \frac{1}{2}(e^{ix} + e^{-ix}), \qquad \mathrm{Im}(e^{ix}) = \frac{1}{2i}(e^{ix} - e^{-ix}). \tag{21.6}$$

Therefore they are, respectively, the solutions of the initial value problems,

$$y'' + y = 0\,, \qquad y(0) = 1\,, \qquad y'(0) = 0 \tag{21.7}$$

$$y'' + y = 0\,, \qquad y(0) = 0\,, \qquad y'(0) = 1\,. \tag{21.8}$$

In view of the uniqueness theorem for solutions to the initial value problem — see Theorem 18.2 — and the discussion in Section 18.1, we can infer that they must be precisely the trigonometric functions:

$$\cos(x) = \frac{1}{2}\big(e^{ix} + e^{-ix}\big)\,, \qquad \sin(x) = \frac{1}{2i}\big(e^{ix} - e^{-ix}\big)\,. \tag{21.9}$$

Moreover, by the very definition of the trigonometric functions as the real and imaginary part we obtain **Euler's identity**:

$$e^{ix} = \cos(x) + i\sin(x) \tag{21.10}$$

We will discuss more properties of the trigonometric functions separately below.

Remark 21.1. For example, we see directly from the definition that

$$\cos^2(x) + \sin^2(x) = (\mathrm{Re}(e^{ix})^2 + \big(\mathrm{Im}(e^{ix})^2\big) = e^{ix}\overline{e^{ix}} = e^0 = 1\,, \tag{21.11}$$

and so geometrically the set $\{(\cos(\theta), \sin(\theta) : \theta \in [0, 2\pi]\}$ is a circle.

Let us now return to the general case (18.4) of a *homogeneous* linear second order differential equation with constant coefficients:

$$y'' + ay' + by = 0 \tag{21.12}$$

As we have seen in Note 19, the corresponding *characteristic polynomial* is (19.4), namely the algebraic condition for $f(x) = e^{\lambda x}$ to be a solution is that λ is a root of

$$Q(\lambda) = \lambda^2 + a\lambda + b\,. \tag{21.13}$$

This polynomial may not have any real zeros, depending on the sign of the *discriminant* $\Delta = a^2 - 4b$. Indeed λ is a solution to $Q(\lambda) = 0$ if and only if

$$\lambda = -\frac{a}{2} \pm \frac{1}{2}\sqrt{a^2 - 4b} \tag{21.14}$$

Case $\Delta > 0$.

Exercise 21.1. Let λ_1, λ_2 be the two real solutions of (21.13). Show that $f_1(x) = e^{\lambda_1 x}$, and $e^{\lambda_2 x}$ are two linearly independent solutions of (21.12).

Case $\Delta = 0$. In this case the two roots coincide $\lambda_1 = \lambda_2 = -a/2$. Let us prove that

$$f_1(x) = e^{-ax/2} \qquad f_2(x) = xe^{-ax/2} \tag{21.15}$$

are two linearly independent solutions to (21.12).

Clearly $f_1(x)$ is a solution, and we verify that

$$f_2''(x) + af_2'(x) + bf_2(x) = -ae^{-a/2} + xQ(-a/2)e^{-ax/2} + ae^{-ax/2} = 0\,. \tag{21.16}$$

Moreover the Wronskian of the two solutions is

$$\begin{aligned} W(x) =& f_1(x)f_2'(x) - f_1'(x)f_2(x) = \\ =& e^{-ax/2}(1 - ax/2)e^{-ax/2} + (a/2)e^{-ax/2}xe^{-ax/2} = e^{-ax/2} \neq 0 \end{aligned} \tag{21.17}$$

and thus $f_1(x)$ and $f_2(x)$ are linearly independent by Theorem 18.1.

Case $\Delta < 0$. In this case we verify that

$$\lambda = \alpha \pm i\beta\,, \qquad \alpha = -a/2\,, \qquad \beta = \sqrt{-\Delta}/2 \tag{21.18}$$

are zeros of the characteristic polynomial $Q(\lambda)$. Similarly to the special case (21.4), we expect that the real and imaginary parts of $e^{\lambda x}$ are solutions to (21.12). Since

$$e^{(\alpha+i\beta)x} = e^{\alpha x}e^{i\beta x}\,, \qquad \overline{e^{(\alpha+i\beta)x}} = e^{\alpha x}e^{-i\beta x} = e^{(\alpha-i\beta)x} \tag{21.19}$$

we obtain

$$\mathrm{Re}\Big(e^{(\alpha+i\beta)x}\Big) = \frac{1}{2}e^{\alpha x}\Big(e^{i\beta x} + e^{-i\beta x}\Big) = e^{\alpha x}\cos(\beta x) \tag{21.20}$$

$$\mathrm{Im}\Big(e^{(\alpha+i\beta)x}\Big) = \frac{1}{2i}e^{\alpha x}\Big(e^{i\beta x} - e^{-i\beta x}\Big) = e^{\alpha x}\sin(\beta x) \tag{21.21}$$

Exercise 21.2. Show that the functions $f_1(x) = e^{\alpha x}\cos(\beta x)$, and $f_2(x) = e^{\alpha x}\sin(\beta x)$ are solutions to (21.12) in the case $\Delta < 0$.

Exercise 21.3. Show that the solutions $f_1(x)$, and $f_2(x)$ are linearly independent, i.e. that for all $x \in \mathbb{R}$,

$$W(f_1, f_2)(x) \neq 0\,. \tag{21.22}$$

Summary. In summary, to find the solutions to the homogeneous equation (21.12), one may determine the zeros of the characteristic polynomial,

$$Q(\lambda) = \lambda^2 + a\lambda + b = 0 \qquad \lambda \in \mathbb{C} \tag{21.23}$$

In the three possible cases

1. $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda \neq 0$
2. $\lambda_1 = \lambda_2 \in \mathbb{R}$

3. $\lambda_1 \in \mathbb{C}$, $\mathrm{Im}(\lambda) \neq 0$, $\lambda_2 = \overline{\lambda_1}$,

the functions f_1 and f_2 defined above form a *fundamental system*, in the sense that a twice differentiable function f is a solution to (21.12) if and only there are constants $c_1, c_2 \in \mathbb{R}$ (uniquely determined) such that $f = c_1 f_1 + c_2 f_2$.

Exercise 21.4. Convince yourself that the fundamental system of solutions found in this way is exactly the same as we have found initially in Lecture 18. In particular from Proposition 18.1 we know that any solution to (21.12) can be written as a linear combination of

$$f_1(x) = e^{-ax/2} u_1(x) \qquad f_1(x) = e^{-ax/2} u_2(x) \tag{21.24}$$

where $u_1(x)$, and $u_2(x)$ are as in Theorem 18.2 depending on the sign of the discrimant $\Delta = a^2 - 4b$.

21.2. Inhomogeneous equations

A few comments are in order about the inhomogeneous equation

$$y'' + ay' + by = r(x) \tag{21.25}$$

where a and b remain constants, but $r(x)$ is a continuous function on $(-\infty, \infty)$.

Let us first observe that if y_1 and y_2 are solutions to (21.25), then $y_1 - y_2$ is a solution to the *homogeneous* equation (21.12), and thus can be written as a linear combination of the fundamental solutions f_1 and f_2. In other words, there are constants c_1, and c_2 such that

$$y_1 - y_2 = c_1 f_1 + c_2 f_2 \,. \tag{21.26}$$

This means in particular if *one particular* solution y_1 to (21.25) is known, then *any* solution y_2 to (21.25) can be written as $y_2 = c_1 f_1 + c_2 f_2 + y_1$.

Proposition 21.1. *If y_1 is a particular solution to* (21.25)*, then the general solution to the inhomogeneous equation* (21.25) *is obtained by adding to y_1 the general solution to the homogeneous equation* (21.12)*.*

It thus suffices to find one particular solution to the inhomogeneous problem.

In general, for any continuous function $r(x)$, a particular solution can be found by the method of **variation of constants**; (see the **additional notes** to Module VII).

21.3. Special types of the inhomogeneous terms

While Theorem 22.1 provides a general method to determine a particular solution to

$$y'' + ay' + by = r(x)\,, \tag{21.27}$$

there are special cases — namely when the function $r(x)$ takes a special form — for which other methods are available. We illustrate this in the cases when $r(x)$ is a polynomial, or a polynomial times an exponential.

Case: $r(x) = p_n(x)$ **is a polynomial of degree** n. If $b \neq 0$, then a particular solution is given by a polynomial

$$g(x) = \sum_{k=0}^{n} a_k x^k \,, \tag{21.28}$$

whose coefficients can be determined successively order by order, after substituting into the equation. Moreover, if $b = 0$ but $a \neq 0$, then a polynomial of order $n+1$ satisfies the equation

$$y'' + ay' = p_n(x) \,. \tag{21.29}$$

Example 21.1. Consider the equation

$$y'' + y = x^3 \,. \tag{21.30}$$

The ansatz

$$y_1(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0 \tag{21.31}$$

leads to the equation

$$a_3 x^3 + a_2 x^2 + (6a_3 + a_1)x + 2a_2 + a_0 = x^3 \tag{21.32}$$

which is evidently satisfied provided

$$a_3 = 1 \quad a_2 = 0 \quad a_1 = -6a_3 = -6 \quad a_0 = -2a_2 = 0 \,. \tag{21.33}$$

Therefore a particular solution is $g(x) = x^3 - 6x$, and the general solution is given by

$$y(x) = c_1 \cos(x) + c_2 \sin(x) + x^3 - 6x \,. \tag{21.34}$$

Exercise 21.5. Derive the particular solution to (21.30) by using Theorem 22.1 and verify that the two methods give the same result.

Case: $r(x) = p_n(x)e^{mx}$ **where** p_n **is a polynomial of degree** n**, and** $m \in \mathbb{R}$. In this case we can always find a particular solution of the form

$$y_1(x) = u(x)e^{mx} \tag{21.35}$$

because inserting this ansatz in (21.27) gives an equation for u,

$$u'' + (2m + a)u' + (m^2 + am + b)u = p_n(x) \tag{21.36}$$

which is precisely of the from discussed above, and can be solved with a particular solution of the form

$$q(x) = \sum_{k=1}^{n} a_k x^k \,, \tag{21.37}$$

at least when $m^2 + am + b \neq 0$. If $m^2 + am + b = 0$, but $2m + a \neq 0$, then we can take q to be a polynomial of degree $n + 1$.

Example 21.2. Consider the equation

$$y'' + y = xe^{3x} \,. \tag{21.38}$$

Setting $y = ue^{3x}$ we find

$$u'' + 6u' + 10u = x \tag{21.39}$$

which we can solve with the ansatz

$$q(x) = a_1 x + a_0 \,, \tag{21.40}$$

which gives the relations

$$10a_1 x + 6a_1 + 10a_0 = x \tag{21.41}$$

which can only be satisfied if $a_1 = 1/10$, and $a_0 = -6/100$, so

$$q(x) = (5x - 3)/50 \tag{21.42}$$

and so the general solution is

$$y(x) = c_1 \cos(x) + c_2 \sin(x) + (5x - 3)e^{3x}/50 \,. \tag{21.43}$$

Problems

1. Find the general solution of each of the following differential equations. Unless defined on the whole real line, indicate the interval on which the solution is defined.
   a) $y'' - y = x$
   b) $y'' + y' = x^2 + x$
   c) $y'' - 4y = e^{2x}$
   d) $y'' - 2y' + y = x + 2xe^{2x}$
2. If $k \neq 0$, verify that the equation

   $$y'' - k^2 y = r(x) \tag{21.44}$$

   has a particular solution given by

   $$g(x) = \frac{1}{k} \int_0^x r(t) \sinh\big(k(x - t)\big) \mathrm{d}t \,. \tag{21.45}$$

1

Additional: Complex functions, and power series

Further Reading

This note is only an introduction to complex valued functions. This material is mostly beyond the scope of this subject (MAST10021). A deeper discussion, along with many of the proofs can be found in (Spivak, *Calculus*, Chapter 26, 27). For applications also see (Apostel, *Calculus I*, Chapter 9).

Complex Analysis (MAST30021)

In this note, we want to understand the meaning of the expression

$$f(x) = e^{ix} \tag{22.1}$$

that we have encountered as a formal solution to a differential equation in Note 21. f is an example of a **complex-valued** function: It assigns to each real number x a complex number $f(x)$. More generally we would like to understand e^z for any complex number z. This is an example of a **complex function**.

22.1. Complex functions

A complex function is rule

$$\begin{aligned} f :& \mathbb{C} \to \mathbb{C} \\ & z \mapsto f(z) \end{aligned} \tag{22.2}$$

which assigns to each complex number, or complex numbers in a subset $D \subset \mathbb{C}$ — the **domain** of f — another complex number.[1] We can always write

$$f(z) = u(z) + iv(z) \tag{22.3}$$

where $u(z)$ and $v(z)$ are *real* numbers, and say that f is **real-valued** if $v(z) = 0$.

[1] In much the same way that we have talked about functions in Note 2 it makes sense to give a formal defintion of a **complex function** as a collection of pairs of complex numbers, which does not contain two distinct pairs with the same first element; cf. **Additional material** to Module I.

Example 22.1. A **polynomial** is given by

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0 \tag{22.4}$$

where $a_0, \ldots, a_n \in \mathbb{C}$.

Example 22.2. The **conjugate** function is defined by

$$f(z) = \bar{z}\,. \tag{22.5}$$

Example 22.3. Since any complex number z can be written in the form $z = x + iy$, many explict examples can be written down using (22.3), such as

$$f(x + iy) = e^y \sin(x - y) + ix^3 \cos(y)\,. \tag{22.6}$$

In a similar fashion to Note 3 and 4 we can talk about the existence of a **limit**, and **continuity**:

> The statement
> $$\lim_{z \to a} f(z) = l \tag{22.7}$$
> means that for every (real) number $\varepsilon > 0$, there is a (real) number $\delta > 0$, such that, for all z, if $0 < |z - a| < \delta$, then $|f(z) - l| < \varepsilon$.

While formally the same, the geometric interpretation is different; see Fig. 22.1. The function $f(z)$ has the limit l as z approaches a, if $f(z)$ can be made to lie in a *circle of radius ε in the complex plane $\mathbb{C}$, by restricting z to lie within a circle of radius δ around a in the complex plane $\mathbb{C}$.*

Moreover a function $f(z)$ is **continuous at** a if $\lim_{z\to a} f(z) = f(a)$, and continuous if f is continuous at every a in the domain.

Remark 22.1. Compare this to the discussion of limits and continuity for functions of two variables in Note 6!

In much the same way as in Notes 3 and 4, one can prove that

$$\lim_{z \to a} z = a \tag{22.8}$$

$$\lim_{z \to a} \Big(f(z) + g(z) \Big) = \lim_{z \to a} f(z) + \lim_{z \to a} g(z) \tag{22.9}$$

$$\lim_{z \to a} f(z) \cdot g(z) = \lim_{z \to a} f(z) \cdot \lim_{z \to a} g(z) \tag{22.10}$$

$$\lim_{z \to a} \frac{f(z)}{g(z)} = \frac{\lim_{z\to a} f(z)}{\lim_{z\to a} g(z)} \qquad , \text{ if } \lim_{z \to a} g(z) \neq 0\,. \tag{22.11}$$

The story is however more delicate — but eventually more beautiful! — for differentiability. We can define that a function $f(z)$ is **differentiable at** a, if

$$f'(a) = \lim_{z \to a} \frac{f(z) - f(a)}{z - a} \tag{22.12}$$

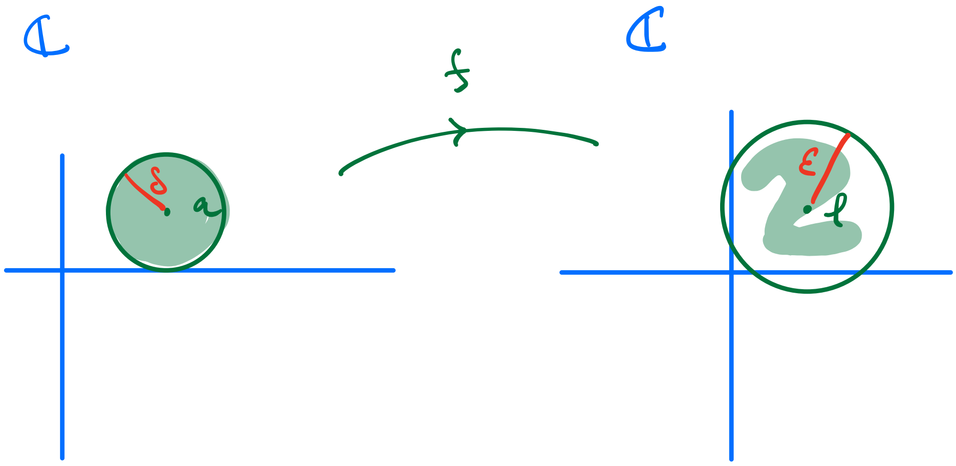

Figure 22.1.: Definition of limits of complex functions.

exists, in which case the limit on the right is denoted by $f'(z)$ on the left. While the familiar rules of differentiation from Note 7 can be verified for **rational functions**, in particular for

$$f(z) = z^n \qquad f'(z) = nz^{n-1} \tag{22.13}$$

there are many perplexing examples of functions which are simply not differentiable.

Example 22.4. Consider the conjugate function $f(z) = \bar{z}$ which we can also write in the form

$$f(x + iy) = x - iy\,. \tag{22.14}$$

Note that for $z = x + iy$, with $y = 0$, the difference quotient at $a = 0$ reads

$$\frac{f(z) - f(0)}{z - 0} = \frac{x - iy}{x + iy} = \frac{x}{x} = 1 \tag{22.15}$$

but the same expression restricted to $z = x + iy$ with $x = 0$ reads

$$\frac{f(z) - f(0)}{z - 0} = \frac{x - iy}{x + iy} = \frac{-iy}{iy} = -1\,. \tag{22.16}$$

Thus no matter how small we take $0 < |z|$, $f(z)$ does *not* approach any limit l.

Nonetheless the most important differentiable functions can be defined by means of infinite series.

22.2. Complex power series

The complex-valued function (22.1) is the most important example of a complex power series:

$$\exp(z) = e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \tag{22.17}$$

For *real numbers* $z = x$ this is an identity with the exponential function e^x defined as in Note 13:

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots \tag{22.18}$$

The precise interpretation of the expression $1 + x + x^2/2! + \dots$, and the explanation for why this is an identity is the subject of Module VIII and **??**.

In fact, the idea for the interpretation of the right hand side of (22.17) is as follows: For a given complex number z, the partial sum

$$s_n(z) = \sum_{j=0}^{n} \frac{z^j}{j!} = 1 + z + \dots + \frac{z^n}{n!} \tag{22.19}$$

is a **sequence of complex numbers**. In Note 25 we will discuss the notion of **convergence** of a sequence of numbers: Informally speaking, a sequence $\{a_n\}$ of complex numbers has a limit l, if for any small circle drawn around l in the complex plane, the sequence is eventually, namely for large n, contained in that circle. We then write

$$\lim_{n\to\infty} a_n = l\,. \tag{22.20}$$

It turns out that the sequence of numbers $\{s_n(z)\}$ of (22.19) converges for any complex number, and we denote the limit by

$$\sum_{j=0}^{\infty} \frac{z^j}{j!} = \lim_{j\to\infty} s_n(z)\,. \tag{22.21}$$

In Note **??** we will give a formal definition of convergent series along these lines.

More generally, a **complex power series** is a series of the form

$$f(z) = \sum_{n=1}^{\infty} a_n(z-a)^n = a_0 + a_1(z-a) + a_2(z-a)^2 + \dots \tag{22.22}$$

where $a, a_n \in \mathbb{C}$, and the power series centered at 0 reads

$$f(z) = \sum_{n=1}^{\infty} a_n z^n = a_0 + a_1 z + a_2 z^2 + \dots \tag{22.23}$$

An important theorem about power series is that if $f(z_0)$ converges, then $f(z)$ converges for any $|z| < |z_0|$. This means, geometrically, that if the power series converges for some z_0, then it does so inside the entire circle of radius $|z_0|$. In fact, for power series of the form (22.23) one of the following three possibilities must be true:

1. $\sum_{n=0}^{\infty} a_n z^n$ converges only for $z = 0$
2. $\sum_{n=0}^{\infty} a_n z^n$ converges absolutely for all $z \in \mathbb{C}$
3. There is a number $R > 0$ such that $\sum_{n=0}^{\infty} a_n z^n$ converges absolutely if $|z| < R$ and diverges for $|z| > R$.

(We say a series $\sum_n a_n$ "converges absolutely" if $\sum_n |a_n|$ converges.)

The number R is called the **radius of convergence** of the power series.[2] Inside the **circle of convergence** a power series defines a differentiable function. In view of the examples we have given in the previous section, this shows that power series provide a large class of differentiable function.

In fact, if the power series (22.23) has radius of convergence $R > 0$, then f is *differentiable* inside the circle of convergence $|z| < R$, and

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1} \,. \tag{22.24}$$

It follows that a power series is *infinitely differentiable* and continuous inside the circle of convergence.

As we have seen power series provide a way to *define* the **complex exponential** $\exp(z)$ as in (22.17). Then by the above mentioned results for power series $\exp'(z) = \exp(z)$. We compute in particular

$$\begin{aligned} e^{iz} =& 1 + iz + \frac{(iz)^2}{2!} + \frac{(iz)^3}{3!} + \frac{(iz)^4}{4!} + \frac{(iz)^5}{5!} + \dots \\ =& 1 + iz - \frac{z^2}{2!} - i\frac{z^3}{3!} + \frac{z^4}{4!} + i\frac{z^5}{5!} + \dots \\ =& \Big(1 - \frac{z^2}{2!} + \frac{z^4}{4!} + \dots\Big) + i\Big(z - \frac{z^3}{3!} + \frac{z^5}{5!} + \dots\Big) \\ =& \cos(z) + i\sin(z) \end{aligned} \tag{22.25}$$

where we have *defined* the complex functions $\cos(z)$ and $\sin(z)$ by:

$$\sin(z) = z - \frac{z^3}{3!} + \frac{z^5}{5!} + \dots w \tag{22.26}$$

$$\cos(z) = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} + \dots \tag{22.27}$$

then also $\sin'(z) = \cos(z)$, and $\cos'(z) = -\sin(z)$.

As for the exponential function in (22.18), the formulas (22.26) and (22.27) *are identities for real numbers* $z = x$. This will be proven in Modules VIII and **??**. In other words, the

[2] Inside the circle of convergence the power series converges absolutely, but outside it diverges. What happens *on* the circle is a more difficult question, and there are examples of series which converge on this circle, and others which do not.

definitions (22.26) and (22.27) extend statements for trigonometric functions and their **Taylor series** to the domain of complex functions.

It is also clear from the definitions, namely the power series, that

$$\sin(-z) = -\sin(z) \tag{22.28}$$
$$\cos(-z) = \cos(z) \tag{22.29}$$

and so we also have

$$e^{-iz} = \cos(z) - i\sin(z)\,. \tag{22.30}$$

From these formulas we obtain that

$$\cos(z) = \frac{1}{2}\left(e^{iz} + e^{-iz}\right) \tag{22.31}$$
$$\sin(z) = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right) \tag{22.32}$$

Note that for real z this agrees with our discussion in Note 21.

From this point of view the exponential function plays a central role in the definition of all elementary functions. We also obtain the famous relation

$$e^{i\pi} = -1\,, \tag{22.33}$$

from (22.25) with $z = \pi$, and more generally that $e^{2\pi i/n}$ is an nth root of 1.

Recall that for *real* numbers z, the values of $\sin(z)$ always lie between -1 and 1. However for *complex* z this is not true at all: Take $z = iy$, for any real y, then

$$\sin(iy) = \frac{e^{i(iy)} - e^{-i(iy)}}{2i} = \frac{e^{-y} - e^{y}}{2i} = i\sinh(y)\,, \tag{22.34}$$

which is unbounded.

It may seem that the complex functions we have considered are very special, but this is not quite true. The basic theorems of **complex analysis** show:

> *If a complex function is defined in some region A of the complex plane and is differentiable in A, then it is automatically infinitely differentiable in A. Moreover, for each point a in A the function f can be represented as a power series inside a circle centered at the point a.*

The power series that are relevant for this statement are **Taylor series** which are the topic of Module VIII and **??**.

22.3. Applications to differentiation and integration

Let now $\lambda \in \mathbb{C}$ be a complex number, and consider

$$f(x) = \exp(\lambda x)\,. \tag{22.35}$$

We have seen above that

$$
\begin{aligned}
f'(x) = \Big(\exp(\lambda x)\Big)' &= \Big(1 + \lambda x + \frac{1}{2}(\lambda x)^2 + \frac{1}{3!}(\lambda x)^3 + \dots\Big)' \\
&= \lambda + \lambda(\lambda x) + \frac{\lambda}{2}(\lambda x)^2 + \dots = \lambda \exp(\lambda x) = \lambda f(x)
\end{aligned}
\tag{22.36}
$$

or simply

$$
e^{\lambda x} = \Big(\frac{1}{\lambda} e^{\lambda x}\Big)' \tag{22.37}
$$

and

$$
\int e^{\lambda x} \mathrm{d}x = \frac{1}{\lambda} e^{\lambda x}\,. \tag{22.38}
$$

The statement is here that $\lambda^{-1}\exp(\lambda x)$ is a primitive of $\exp(\lambda x)$. Here the notion of continuity and differentiablity are to be understood for complex valued functions of a *real variable*.

Exercise 22.1. A complex valued functions of a *real variable* can always be written as

$$
f(x) = u(x) + iv(x)\,. \tag{22.39}
$$

Adapt the definition of continuity given in (22.7) to the case of a comlex-valued function of a real variable, and show that f is continuous in that sense, if and only if u and v are continuous.

Exercise 22.2. Similarly, show that f is differentiable if and only if u and v are differentiable, and

$$
f'(x) = u'(x) + iv'(x)\,. \tag{22.40}
$$

Use this to give an alternative proof of (22.38) using Euler's identity (21.10).

As an application recall the formula (21.20). We obtain

$$
\begin{aligned}
\int e^{\alpha x}\cos(\beta x)\mathrm{d}x &= \mathrm{Re}\int e^{(\alpha+i\beta)x}\mathrm{d}x \\
&= \mathrm{Re}\Big(\frac{1}{\alpha+i\beta}e^{(\alpha+i\beta)x}\Big) = \frac{1}{\alpha^2+\beta^2}\Big(\alpha e^{\alpha x}\cos(\beta x) + \beta e^{\alpha x}\sin(\beta x)\Big)
\end{aligned}
\tag{22.41}
$$

Exercise 22.3. Derive a similar formula from (21.21).

Additional Problems

1. Show that every complex number of absolute value 1 can be written as e^{iy} for some real number y.

2. For all complex numbers z and w it holds that $e^{z+w} = e^z e^w$. It also holds that $\sin(z+w) = \sin(z)\cos(w) + \cos(z)\sin(w)$. Interprete these statement in terms of series. How could we prove these?

3. Prove that $|e^{x+iy}| = e^x$ for real x and y.

4. a) Prove that exp takes on every complex value except 0.
   b) Prove that sin takes on every complex value.

5. Show that exp is not one-to-one on $\mathbb{C}$. Given $w \neq 0$ find z so that $e^z = w$.

6. Prove that

$$\sum_{k=1}^{n} e^{ikx} = e^{ix}\frac{1-e^{inx}}{1-e^{ix}} = \frac{\sin(nx/2)}{\sin(x/2)}e^{i(n+1)x/2}.$$

7. Use (21.9) to show

$$\cos^2(\theta) = \frac{1}{2}(1+\cos(2\theta)) \tag{22.42}$$

   Derive a similar formula for $\sin^2(\theta)$.

8. Prove that every trigonometric sum of the form

$$S_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^{n}(a_k\cos(kx) + b_k\sin(kx)) \tag{22.43}$$

   can be expressed as a sum of complex exponentials,

$$S_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}, \tag{22.44}$$

   where $c_k = \frac{1}{2}(a_k - ib_k)$ for $k = 1, 2, \ldots, n$.

9. a) If m and n are integers, prove that

$$\int_0^{2\pi} e^{inx}e^{-imx}\mathrm{d}x = \begin{cases} 0 & m \neq n \\ 2\pi & m \neq n\,. \end{cases} \tag{22.45}$$

   b) Use part (a) to deduce the following orthgonality relation: If m and n are integers with $m^2 \neq n^2$, then

$$\int_0^{2\pi} \sin(nx)\cos(mx)\mathrm{d}x = 0\,. \tag{22.46}$$

Additional: Catenary problem

In this note we want to solve the

Catenary problem: What shape does a flexible rope[1] of uniform density take under its own weight when suspended between two points?

The rope is in a static equilibrium meaning that at every point on the rope the total force acting on it vanishes. The rope lies in a plane and we can choose coordinates (x, y) so that the y-axis aligns with the gravitational force and the lowest point on the rope falls in the origin $(0, 0)$. At every point (x_0, y_0) on the rope there is a force $F(x_0, y_0) \in \mathbb{R}^2$ corresponding to the pull of the rope segment with $x > x_0$. This force is not yet known to us, but we know it is always tangential to the rope at (x_0, y_0). Now consider the lowest point where the "pull from the rope segment on the left" is balanced by the "pull from the rope segment on the right":

$$(-F_0, 0) + F(0, 0) = 0 \tag{22.1}$$

for some $F_0 > 0$, representing the force exerted by the rope segment with $x < 0$.

Now consider some point (x_0, y_0) on the rope on the right with $x_0 > 0$; cf. Figure 22.1. The *horizontal* component of the force exerted by the rope segment $x < x_0$ is still F_0, but *in addition* there is a *vertical* component related to the weight of the rope segment between $0 < x < x_0$, which is proportional to the total mass of this segment given by ρs, where ρ_0 is the density (mass per unit length) and s is the *arc length* of the rope segment from the lowest point. Therefore

$$(-F_0, -\rho s(x_0, y_0) g) + F(x_0, y_0) = 0 \tag{22.2}$$

where g is the gravitational constant, and by writing F in the form $F = |F|(\cos\theta, \sin\theta)$, where $|F|$ and θ a functions of (x_0, y_0) we obtain two equations

$$|F| \cos\theta = F_0 \qquad |F| \sin\theta = \rho s g\,. \tag{22.3}$$

Let us now view the rope as a graph $y = y(x)$ over the x-axis. The arc length $s(x_0, y_0)$ of the rope segment from $(0, 0)$ to (x_0, y_0) is then given by the length $s(x_0)$ of the curve $x \to (x, y(x))$, with $x \in [0, x_0]$. The length of this curve is given by

$$s(x_0) = \int_0^{x_0} \sqrt{1 + \Big(\frac{\mathrm{d}y(x)}{\mathrm{d}x}\Big)^2}\,\mathrm{d}x \tag{22.4}$$

[1]The word *catenary* comes from the latin word for *chain*, but I find it more pleasant to think about ropes. The problem was first solved in the 1690s, independently by Leibniz, Huygens, and Bernoulli.

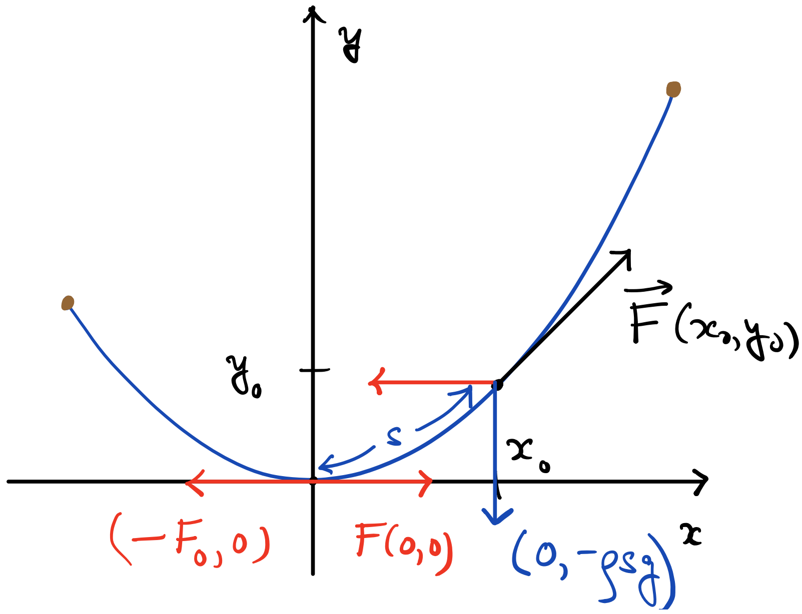

Figure 22.1.: Catenary problem.

where $m(x) = y'(x) = \frac{\mathrm{d}y}{\mathrm{d}x}(x)$ is the slope of the graph at $(x, y(x))$. Since $(\cos(\theta(x)), \sin(\theta(x))$ is the unit tangent vector to the curve at $(x, y(x))$, with the angle θ as introduced above, it follows from (22.3) that

$$m(x) = \tan(\theta(x)) = \frac{\rho s(x) g}{F_0} \tag{22.5}$$

In view of (22.4) we thus obtain the following ODE for the slope $m(x)$:

$$m'(x) = \frac{\rho g}{F_0} s'(x) = \frac{\rho g}{F_0} \sqrt{1 + m(x)^2} \tag{22.6}$$

In Note 16 we have learned that this is an example of a *separable* differential equation, which can be solved by writing

$$\int \frac{\mathrm{d}m}{\sqrt{1 + m^2}} = a \int \mathrm{d}x \tag{22.7}$$

where for short $a = \rho g / F_0$. It follows from (20.26) that

$$\operatorname{arsinh}(m(x)) = ax \tag{22.8}$$

where we used that $m(0) = 0$. Hence

$$\frac{\mathrm{d}y}{\mathrm{d}x}(x) = m(x) = \sinh(ax) \tag{22.9}$$

which we can easily integrate:

$$y(x) = \int_0^x \sinh(ax')\mathrm{d}x' = \frac{\cosh(ax) - 1}{a} \tag{22.10}$$

where we used $y(0) = 0$.

This is the solution to the catenary problem, and obviously finds applications in civil engineering, although the assumption of uniform density is rather idealized.

Exercise 22.1. Find the area under the catenary curve from $x = 0$ to $x = 1/a$.

Additional: Variation of constants

In this note we show how to obtain a particular solution to (21.25) by variation of constants.[1]

More precisely, let f_1 and f_2 be two linearly independent solutions of the homogeneous equation as discussed in Section 21.1. We are then looking for a particular solution to the inhomogeneous problem of the form

$$g(x) = c_1(x)f_1(x) + c_2(x)f_2(x)\,. \tag{22.1}$$

We compute

$$g'(x) = c_1(x)f_1'(x) + c_2(x)f_2'(x) + c_1'(x)f_1(x) + c_2'(x)f_2(x) \tag{22.2}$$

$$\begin{aligned} g''(x) = c_1(x)f_1''(x) + c_2(x)f_2''(x) + c_1'(x)f_1'(x) + c_2'(x)f_2'(x) \\ + (c_1'(x)f_1(x) + c_2'(x)f_2(x))' \end{aligned} \tag{22.3}$$

Therefore, using the fact that $f_1(x)$ and $f_2(x)$ are solutions to the homogeneous equation,

$$\begin{aligned} g''(x) + ag'(x) + bg(x) = c_1'(x)f_1'(x) + c_2'(x)f_2'(x) \\ + (c_1'(x)f_1(x) + c_2'(x)f_2(x))' + a(c_1'(x)f_1(x) + c_2'(x)f_2(x)) \end{aligned} \tag{22.4}$$

So if $c_1(x)$ and $c_2(x)$ can be chosen such that

$$c_1'(x)f_1(x) + c_2'(x)f_2(x) = 0 \tag{22.5}$$

then $g(x)$ is a solution to (21.25) provided also

$$c_1'(x)f_1'(x) + c_2'(x)f_2'(x) = r(x) \tag{22.6}$$

The point is that (22.5) and (22.6) are a linear system of equations for $c_1'(x)$ and $c_2'(x)$ which is always solvable:

$$\begin{pmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{pmatrix} \begin{pmatrix} c_1'(x) \\ c_2'(x) \end{pmatrix} = \begin{pmatrix} 0 \\ r(x) \end{pmatrix} \tag{22.7}$$

[1] This idea was first used by Bernoulli in 1697 to solve linear equations of first order, and then by Lagrange in 1774 to solve linear equations of second order.

The determinant of this matrix is precisely the Wronskian $W(f_1, f_2)(x)$, and we have shown in Theorem 18.1 that $W(f_1, f_2)(x)$ never vanishes. Therefore we can always solve this linear system and obtain

$$c_1'(x) = -\frac{f_2(x)r(x)}{W(f_1, f_2)(x)} \tag{22.8}$$

$$c_2'(x) = \frac{f_1(x)r(x)}{W(f_1, f_2)(x)}\,. \tag{22.9}$$

Theorem 22.1. *Let (f_1, f_2) be a fundamental system for the homogeneous equation* (21.12)*. Then a particular solution to the inhomogeneous equation* (21.25) *is given by the formula* (22.1)*,*

$$y_1(x) = c_1(x)f_1(x) + c_2(x)f_2(x) \tag{22.10}$$

where c_1 and c_2 are primitives of the following functions:

$$c_1 = -\int \frac{f_2(x)r(x)}{W(f_1, f_2)(x)}\mathrm{d}x\,, \qquad c_2 = \int \frac{f_1(x)r(x)}{W(f_1, f_2)(x)}\mathrm{d}x\,. \tag{22.11}$$

Example 22.1. Let us determine the general solution of the equation

$$y'' + y = \sin(2x) \tag{22.12}$$

A fundamental system of solutions for the homogeneous equation is given by

$$f_1(x) = \sin(x)\,, \qquad f_2(x) = \cos(x)\,, \tag{22.13}$$

which has the Wronskian $W(x) = -1$, and the two primitives (22.11) are

$$c_1 = \int \cos(x)\sin(2x)\mathrm{d}x = 2\int \cos^2(x)\sin(x)\mathrm{d}x = -2\cos^3(x)/3 \tag{22.14}$$

$$c_2 = -\int \sin(x)\sin(2x)\mathrm{d}x = -2\int \sin^2(x)\cos(x)\mathrm{d}x = -2\sin^3(x)/3 \tag{22.15}$$

which shows that a particular solution to (22.12) is given by

$$y(x) = -(2/3)\cos^3(x)\sin(x) - (2/3)\sin^3(x)\cos(x)\,, \tag{22.16}$$

and the general solution is given by

$$y(x) = a_1\sin(x) + a_2\cos(x) - (2/3)\cos^3(x)\sin(x) - (2/3)\sin^3(x)\cos(x)\,, \tag{22.17}$$

where a_1, and a_2 are constants.

Problems

1. Find the general solution of the equation

   $$y'' + y = \tan(x) \tag{22.18}$$

   on the interval $(-\pi, \pi)$.

2. Derive the formula (21.45) by the method of variation of constants.

Module VIII.

Taylor polynomials

Note 23.

Approximation by polynomial functions

In the previous lectures we have looked at differential equations with constant coefficients, and we have found several special functions as solutions. These functions are often given implicitly and "in practice" it would not be easy at all to compute their values.

Example 23.1. For example to compute values of the exponential function, we would first have to approximate

$$\log(x) = \int_1^x \frac{1}{t} \mathrm{d}t \tag{23.1}$$

by some upper and lower sums, and finding $e^x = \log^{-1}(x)$ would involve computing $\log(a)$ for many values of a, until $\log(a)$ is close to x, and a would then be an approximation for e^x.

In this lecture we will learn a way to approximate functions by polynomials,

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n \,, \tag{23.2}$$

in the sense that the derivatives of $p(x)$ at a point a agree with those of the given function up to a given order. Such polynomials are called **Taylor polynomials**.

Remark 23.1. There are of course other methods to approximate functions by polynomials. One could for example try to find a polynomial of degree n that passes through $n+1$ given points on the graph of the function. Or one could attempt to make the area between the function and the polynomial as small as possible. Or one could approximate a function *uniformly* by polynomials, as it is done in the "*Weierstrass approximation theorem.*"

First note that all coefficients in (23.2) can be expressed in terms of the values of p and its derivatives at 0: $p(0) = a_0$, and $p'(x) = a_1$, and more generally

$$p^{(k)}(0) = k! a_k \tag{23.3}$$

If we had begun with a polynomial in $(x - a)$, namely

$$p(x) = a_0 + a_1(x-a) + \ldots + a_n(x-a)^n \tag{23.4}$$

then the same argument shows that

$$a_k = \frac{p^{(k)}(a)}{k!} \,. \tag{23.5}$$

Definition 23.1. Suppose f is a function which is n-times differentiable at a. Then

$$P_{n,a}[f](x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \ldots + a_n(x-a)^n\,, \qquad (23.6)$$

$$\text{where } a_k = \frac{f^{(k)}(a)}{k!} \qquad (0 \le k \le n) \qquad (23.7)$$

is the **Taylor polynomial of degree n for f at a.** (We usually drop $[f]$ from the notation.)

The Taylor polynomials of the most important elementary function are extremely simple.

Example 23.2. The Taylor polynomial for $\sin(x)$ at 0 is, to order $2n+1$,

$$P_{2n+1,0}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} \qquad (23.8)$$

Example 23.3. The Taylor polynomial of $\log(x)$ at $a = 1$ is

$$P_{n,1}(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 + \ldots + \frac{(-1)^{n-1}}{n}(x-1)^n \qquad (23.9)$$

Exercise 23.1. For the example of the logarithm it is more convenient to consider the function $f(x) = \log(1+x)$. In this case compute the Taylor polynomial of f at $a = 0$.

Example 23.4. Let us compute the Taylor polynomial for $\arctan(x)$ at $a = 0$. Since

$$\arctan'(x) = \frac{1}{1+x^2}\,, \qquad (23.10)$$

we see that $\arctan'(0) = 1$, and $\arctan''(0) = 0$, so

$$P_{2,0} = x\,. \qquad (23.11)$$

However to compute the Taylor polynomials to higher order appears to be a laborious exercise. We will return to this example when we have learned more about the properties of Taylor polynomials.

Let us return to the Taylor polynomial in general form, first of degree 1, given by

$$P_{1,a}[f](x) = f(a) + f'(a)(x-a)\,. \qquad (23.12)$$

We see that by the very definition of the derivative,

$$\frac{f(x) - P_{1,a}(x)}{x-a} = \frac{f(x) - f(a)}{x-a} - f'(a) \to 0 \qquad (x \to a)\,, \qquad (23.13)$$

which tells us that $f(x) - P_{1,a}[f](x)$ not only becomes small as $x \to a$, but becomes *small compared to* $(x-a)$ *as* $x \to a$.

Example 23.5. Consider the exponential function $f(x) = e^x$, and its Taylor polynomial of degree 1,

$$P_{1,0}(x) = 1 + x\,. \tag{23.14}$$

While this is a good approximation as x approaches 0, it appears that

$$P_{2,0}(x) = 1 + x + \frac{1}{2}x^2 \tag{23.15}$$

gives an even better approximation. Indeed by L'Hôpital's rule,

$$\lim_{x\to 0} \frac{e^x - 1 - x - \frac{1}{2}x^2}{x^2} = \lim \frac{e^x - 1 - x}{2x} = \lim_{x\to 0} \frac{e^x - 1}{2} = 0\,. \tag{23.16}$$

Theorem 23.1. *Suppose f is a function for which $f'(a), \ldots, f^{(n)}(a)$ all exist. Then*

$$\lim_{x\to a} \frac{f(x) - P_{n,a}(x)}{(x-a)^n} = 0\,. \tag{23.17}$$

Proof. This is also a consequence of L'Hôpital's rule. □

We also say that "*$f(x)$ equals $P_{n,a}(x)$ up to order n at a.*" We have designed $P_{n,a}$ to have that property, but it turns out that the Taylor polynomial is *the only polynomial with this property.*

Theorem 23.2. *Let $p(x)$ and $q(x)$ be two polynomials in $(x-a)$, of degree at most n, which are equal up to order n at a in the sense that*

$$\lim_{x\to a} \frac{p(x) - q(x)}{(x-a)^n} = 0\,. \tag{23.18}$$

Then $p = q$.

Proof. The difference $r = p - q$ is a polynomial of degree at most n,

$$r(x) = b_0 + b_1(x-a) + \ldots + b_n(x-a)^n \tag{23.19}$$

and by assumption, *for all* $0 \le i \le n$,

$$\lim_{x\to a} \frac{r(x)}{(x-a)^i} = 0\,. \tag{23.20}$$

In the case $i = 0$ this implies $b_0 = 0$, so

$$r(x) = b_1(x-a) + \ldots + b_n(x-a)^n \tag{23.21}$$

and we can use the case $i = 1$ to infer that $b_1 = 0$. Continuing in this way gives

$$b_0 = b_1 = b_2 = \ldots = b_n = 0\,.$$

□

Corollary 23.3. *Let f be n-times differentiable at a, and suppose P is a polynomial in $(x-a)$ that equals f up to order n at a. Then $P = P_{n,a}[f]$ is the Taylor polynomial of f.*

In some situations this insight gives an unexpected way *to compute* the Taylor polynomial of a function.

Example 23.6. Let us return to the problem of computing the Taylor polynomial of

$$\arctan(x) = \int_0^x \frac{1}{1+t^2}\mathrm{d}t\,. \tag{23.22}$$

It is not hard to verify (multiply both sides by $(1+t^2)$!) that

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - t^6 + \ldots + (-1)^n t^{2n} + (-1)^{n+1}\frac{t^{2n+2}}{1+t^2}\,. \tag{23.23}$$

Therefore

$$\arctan(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \ldots + (-1)^n\frac{x^{2n+1}}{2n+1} + (-1)^{n+1}\int_0^x \frac{t^{2n+2}}{1+t^2}\mathrm{d}t\,, \tag{23.24}$$

and so the polynomial that appears here *is the Taylor polynomial* of degree $2n+1$, provided we can show that

$$\lim_{x\to 0}\frac{1}{x^{2n+1}}\int_0^x \frac{t^{2n+2}}{1+t^2}\mathrm{d}t = 0 \tag{23.25}$$

Since

$$0 \le \Big|\int_0^x \frac{t^{2n+2}}{1+t^2}\mathrm{d}t\Big| \le \frac{|x|^{2n+3}}{2n+3} \tag{23.26}$$

this is satisfied, and we conclude

$$P_{2n+1,0}(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \ldots + (-1)^n\frac{x^{2n+1}}{2n+3}\,. \tag{23.27}$$

This last example also shows that for $|x| \le 1$,

$$|\arctan(x) - P_{2n+1,0}(x)| \le \frac{1}{2n+1}\,, \tag{23.28}$$

which means that we use Taylor polynomials to compute $\arctan(x)$ on this interval as accurately as we like. We will now turn more generally to the question how well a Taylor polynomial $P_{n,a}[f](x)$ approximates $f(x)$, for fixed x, and different n.

Problems

1. Find the Taylor polynomials $P_{n,a}[f]$ for the following functions f:
   a) $f(x) = e^{e^x}$, $n = 3$, $a = 0$

b) $f(x) = e^{\sin(x)}$, $n = 3$, $a = 0$

c) $f(x) = \sin(x)$, degree $2n$, $a = \pi/2$

d) $f(x) = \cos(x)$, degree $2n$, $a = \pi$

2. Write each of the following polynomials in x as a polynomial in $(x - 3)$.

   a) $x^2 - 4x - 9$

   b) x^5

   Hint: It is only necessary to compute the Taylor polynomial at $a = 3$, of the same degree as the given polynomial. Why?

3. Consider the equation $x^2 = \cos(x)$, which has precisely two solutions. Use the Taylor polynomial of the cosine function of degree 3, to show that the solutions are approximately $\pm\sqrt{2/3}$, and give a bound for the error. Use a fifth degree Taylor polynomial to get a better approximation.

4. Suppose that a_i and b_i are the coefficients in the Taylor polynomials at a of f and g. Find the coefficients of the Taylor polynomials at a of the following functions in terms of a_i and b_i:

   a) $f + g$

   b) fg

   c) f'

   d) $h(x) = \int_a^x f(t)\mathrm{d}t$

5. Prove that the Taylor polynomial of $f(x) = \sin(x^2)$ of degree $4n + 2$ at 0 is

$$P_{4n+2,0}(x) = x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \ldots + (-1)^n \frac{x^{4n+2}}{(2n+1)!}\,. \tag{23.29}$$

   Hint: If P is the Taylor polynomial of degree $2n + 1$ of the sine function at 0, then $\sin(x) = P(x) + R(x)$, where $\lim_{x\to 0} R(x)/x^{2n+1} = 0$. What does this imply for the limit of $R(x^2)/x^{4n+2}$?

Note 24.

Taylor's theorem

In the previous lecture we have introduced Taylor polynomials $P_{n,a}$ as polynomial functions whose derivatives at a point a agree with those of a given function f up to a certain order n. The main theorem in this lecture will answer the question precisely in which sense a Taylor polynomial approximates the values of a given function away from the point a.

If f is a function for which $P_{n,a}[f]$ exists, we define the **remainder term** $R_{n,a}[f]$ by

$$\begin{aligned} f(x) =& P_{n,a}[f](x) + R_{n,a}[f](x) \\ =& f(a) + f'(a)(x-a) + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n,a}[f](x) \end{aligned} \tag{24.1}$$

The aim is to derive a *formula* for the remainder. We can see this might be possible from the case $n = 0$: By the fundamental theorem of Calculus

$$R_{0,a}(x) = f(x) - f(a) = \int_a^x f'(x)\mathrm{d}x\,. \tag{24.2}$$

We can proceed by integration by parts to obtain a formula in the case $n = 1$:

$$\begin{aligned} \int_a^x f'(t)\mathrm{d}t =& \int_a^x f'(t)(t-x)'\mathrm{d}t \\ =& (x-a)f'(a) - \int_a^x f''(t)(t-x)\mathrm{d}t \end{aligned} \tag{24.3}$$

which shows that

$$f(x) = f(a) + f'(a)(x-a) + R_{1,a}(x)\,, \qquad R_{1,a}(x) = \int_a^x f''(t)(x-t)\mathrm{d}t\,. \tag{24.4}$$

Exercise 24.1. Show that

$$R_{2,a}(x) = \int_a^x \frac{f^{(3)}(t)}{2}(x-t)^2\mathrm{d}t \tag{24.5}$$

by suitably integrating by parts:

$$R_{1,a}(x) = \int_a^x f''(t)\Big[-\frac{1}{2}(x-t)^2\Big]'\mathrm{d}t\,. \tag{24.6}$$

Proceeding in this way we can prove by induction the *integral form of the remainder*,

$$R_{n,a}(x) = \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n \mathrm{d}t\,. \tag{24.7}$$

We get from this the first *estimate* for the remainder:

Proposition 24.1. *Suppose f is $n+1$ times differentiable on an interval I, $a \in I$, and $|f^{(n+1)}(x)| \leq M$ for $x \in I$, with some $M > 0$. Then*

$$|R_{n,a}(x)| \leq \frac{M}{(n+1)!}|x-a|^{n+1} \tag{24.8}$$

Example 24.1. As an example consider the trigonometric functions,

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + R_{2n+1,0}(x) \tag{24.9}$$

and the remainder term is easy to estimate:

$$|R_{2n+1,0}(x)| \leq \frac{x^{2n+2}}{(2n+2)!} \tag{24.10}$$

Since $x^n/n!$ can be made arbitrarily small by choosing n large enough, for *any* x, (see Note 25) this means that that $\sin(x)$ can be computed to *any* degree of accuracy by evaluating the Taylor polynomial $P_{n,0}(x)$. For example, suppose we wish to compute $\sin(2)$ with an error of less than 10^{-4}, we just need to choose n so that

$$\frac{2^{2n+2}}{(2n+2)!} < 10^{-4} \tag{24.11}$$

In this case $n = 5$ works, and so

$$\sin(2) = 2 - \frac{2^3}{3!} + \frac{2^5}{5!} - \frac{2^7}{7!} + \frac{2^9}{9!} - \frac{2^{11}}{11!} + R \tag{24.12}$$

where $R < 10^{-4}$.

While we will be content with this estimate of the remainder, another version of Taylor's theorem gives the statement $f(x) = P_{n,a}(x) + R_{n,a}(x)$ even under the weaker assumption that $f^{(k)}$ exist up to order $0 \leq k \leq n+1$, but $f^{(n+1)}$ is not necessarily continuous. Its proof does not rely on the integral form of the remainder, but instead on a higher order version of the mean value theorem; (see the **additional notes** to Module VIII) .

Theorem 24.2 (Taylor's theorem with Lagrange remainder)**.** *Let f be a $n+1$-times differentiable function on an interval (c,d), $a \in (c,d)$, and $R_{n,a}(x)$ defined by*

$$f(x) = f(a) + f'(a)(x-a) + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n,a}(x) \tag{24.13}$$

Then for some t in between a and x,

$$R_{n,a}(x) = \frac{f^{(n+1)}(t)}{(n+1)!}(x-a)^{n+1}\,. \tag{24.14}$$

Problems

1. Prove that if $x \leq 0$, then the remainder term $R_{n,0}$ for e^x satisfies

$$|R_{n,0}| \leq \frac{|x|^{n+1}}{(n+1)!} \tag{24.15}$$

2. Prove that if $-1 < x \leq 0$, then the remainder term $R_{n,0}$ for $\log(1+x)$ satisfies

$$|R_{n,0}| \leq \frac{|x|^{n+1}}{(1+x)(n+1)}\,. \tag{24.16}$$

3. a) Prove that if $f''(a)$ exists, then

$$f''(a) = \lim_{h\to 0} \frac{f(a+h) + f(a-h) - 2f(a)}{h^2}\,. \tag{24.17}$$

   The limit on the right hand side is called the *Schwarz second derivative* of f at a. *Hint:* Use the Taylor polynomial $P_{2,a}(x)$ with $x = a+h$, and with $x = a-h$.

   b) Let $f(x) = x^2$ for $x \geq 0$, and $-x^2$ for $x \leq 0$. Show that the Schwarz second derivative of f at 0 exists, even though $f''(0)$ does not.

   c) Prove that if $f'''(a)$ exists, then

$$\frac{f^{(3)}(a)}{3} = \lim_{h\to 0} \frac{f(a+h) - f(a-h) - 2hf'(a)}{h^3} \tag{24.18}$$

4. Give another proof for the *uniqueness* of solutions to the differential equation $y'' - y = 0$ *using Taylor's theorem*. In other words, suppose $f'' - f = 0$ and $f(0) = f'(0) = 0$, and show that then $f = 0$.

Additional: Taylor's theorem

Proof of Taylor's theorem with Lagrange remainder

Recommended Reading

(Spivak, *Calculus*, Chapter 20)

Real Analysis: Advanced (MAST20033)

As mentioned above the proof of Taylor's theorem does not rely on the integral form of the remainder, but instead on a higher order version of the mean value theorem:

Lemma 24.1. *Suppose a function g is $n+1$ times differentiable on an interval I, and for some $a \in I$,*

$$g(a) = g'(a) = \ldots = g^{(n)}(a) = 0\,. \tag{24.1}$$

Then for any $x \in I$, we have

$$g(x) = \frac{g^{(n+1)}(t)}{(n+1)!}(x-a)^{n+1}\,. \tag{24.2}$$

for some t in between a and x.

Proof. We can prove this by induction in n.

In the case $n = 0$, by the mean value theorem

$$g(x) = g(x) - g(a) = g'(t)(x-a) \tag{24.3}$$

for some $t \in (a, x)$.

Now let the statement be valid for some $n = k$. We want to show it holds for $n = k+1$. So assume g is $k+2$ times differentiable, and $g(a) = \ldots = g^{(k+1)}(a) = 0$. Now by the Cauchy mean value theorem (Lecture 9) we get with $h(x) = (x-a)^{k+2}$,

$$\frac{g(x)}{(x-a)^{k+2}} = \frac{g(x)-g(a)}{h(x)-h(a)} = \frac{g'(t)}{h'(t)} = \frac{g'(t)}{(k+2)(t-a)^{k+1}} \tag{24.4}$$

for some $t \in (a, x)$. Now g' itself is $k+1$ times differentiable, and satisfies $g'(a) = \ldots = (g')^{(k)}(a) = 0$, so by our inductive assumption,

$$g'(t) = \frac{(g')^{(k+1)}(y)}{(k+1)!}(t-a)^{k+1} \tag{24.5}$$

for some y in between a and t. Therefore

$$\frac{g(x)}{(x-a)^{k+2}} = \frac{(g')^{(k+1)}(y)}{(k+2)(k+1)!} = \frac{g^{(k+2)}(y)}{(k+2)!} \tag{24.6}$$

which shows the claim for $n = k+1$.

□

Proof of Theorem 24.2. The remainder $R_{n,a} = f(x) - P_{n,a}(x)$ satisfies by its very definition,

$$R_{n,a}(a) = \ldots = R_{n,a}^{(n)}(a) = 0 \tag{24.7}$$

Thus by the above Lemma, for some t in between a and x,

$$R_{n,a}(x) = \frac{(f-P_{n,a})^{(n+1)}(t)}{(n+1)!}(x-a)^{n+1} = \frac{f^{(n+1)}(t)}{(n+1)!}(x-a)^{n+1}\,, \tag{24.8}$$

where we used that $P_{n,a}(x)$ is a polynomial of degree n, and hence $P_{n,a}^{(n+1)} = 0$. □

Local extrema

Recommended Reading

(Spivak, *Calculus*, Chapter 20)

As a consequence of Theorem 23.1 the test for local extrema can be answered even in the indefinite case. Recall that if a is a critical point of f, then f has a local minimum at a if $f''(a) > 0$, and a local maximum if $f''(a) < 0$, but no immediate conclusion can be drawn if $f''(a) = 0$. It is now clear that in this case $f^{(3)}(a)$ will give the relevant information, and moreover if also $f^{(3)}(a) = 0$, then the sign of $f^{(4)}(a)$ is significant. More generally, we can ask what happens when

$$f'(a) = f''(a) = \cdots = f^{(n-1)}(a) = 0 \tag{24.9a}$$

$$f^{(n)}(a) \neq 0\,. \tag{24.9b}$$

Theorem 24.2. *Suppose that a given function f satisfies* (24.15)*.*

1. *If n is even and $f^{(n)}(a) > 0$, then f has a local minimum at a.*

2. *If n is even and $f^{(n)}(a) < 0$, then f has a local maximum at a.*

3. *If n is odd, then f has neither a local maximum nor a local minimum at a.*

Additional Problems

1. a) Show that if $|g'(x)| \leq M|x-a|^n$ for $|x-a| < \delta$, then $|g(x) - g(a)| \leq M|x-a|^{n+1}/(n+1)$ for $|x-a| < \delta$.

   b) Use this to show that if $\lim_{x\to a} g'(x)/(x-a)^n = 0$, then

$$\lim_{x\to a} \frac{g(x) - g(a)}{(x-a)^{n+1}} = 0\,. \tag{24.10}$$

   c) Show that if $g(x) = f(x) - P_{n,a}[f](x)$, then $g'(x) = f'(x) - P_{n-1,a}[f'](x)$.

   d) Give an inductive proof of Theorem 23.1, without using L'Hôptial's rule.

2. Deduce Theorem 23.1 as a corollary of Taylor's theorem, albeit under the assumption of one more derivative.

Note 25.
Infinite Sequences

In the previous lecture we have encountered the problem whether the numbers

$$\frac{a^n}{n!} \tag{25.1}$$

for a given choice of $a \in \mathbb{R}$, become "small for n sufficiently large". The numbers $a_n = a^n/n!$ for $n \in \mathbb{N}$ are an example of an "infinite sequence" of real numbers

$$a_1, a_2, a_3, \dots \tag{25.2}$$

which more generally is denoted by

$$\{a_n\}_{n=1}^{\infty} . \tag{25.3}$$

Given that a sequence assigns to each natural number n and real number a_n we could define the concept as follows:

Definition 25.1. An **infinite sequence** of real numbers is a function whose domain is $\mathbb{N}$.

One could graph a sequence in the same way we graph a function, but it is usually more convenient to simply label the points a_n on the real number line $\mathbb{R}$.

Sequences are often defined *explicitly* by a formula for the n^{th} term, or *recursively* in the sense that a_{k+1} is given in terms a_k, or even a_l for $1 \leq l \leq k$.

Example 25.1. $a_k = k^2$.

Example 25.2. The factorial function $n!$ can itself be thought of as a recursively defined sequence. Setting $a_0 = 1$, and $a_{n+1} = a_n(n+1)$, the number $n!$ is n-th number in this sequence, $n! = a_n$.

Example 25.3. Another example of a recursively defined sequence are the Fibonacci numbers, where $x_1 = x_2 = 1$, and $x_k = x_{k-1} + x_{k-2}$ for $k \geq 2$.

The statement we would like to make about the sequence (25.1) is that it "converges to zero", or

$$\frac{a^n}{n!} \to 0 \qquad (n \to \infty) . \tag{25.4}$$

The next definition makes this notion precise.

Definition 25.2 (Convergence)**.** A sequence $\{a_n\}_{n=1}^{\infty}$ **converges** to a limit l, written

$$a_n \to l \qquad (n \to \infty) \qquad \text{or} \qquad \lim_{n\to\infty} a_n = l \tag{25.5}$$

if for every $\epsilon > 0$, there is a natural number N such that

$$|x_n - l| < \epsilon \qquad \text{whenever} \qquad n > N\,. \tag{25.6}$$

We say a sequence **converges** if it converges to a limit l for some $l \in \mathbb{R}$. Otherwise, the sequence is said to **diverge**.

Exercise 25.1. Verify that $a_k = 1/k$ converges in this sense.

Example 25.4. Examples of sequences that do *not* converge are $a_k = k$ which is "going to infinity", and the sequence $a_k = (-1)^k$ which "jumps back and forth between -1, and 1".

Example 25.5. Let us show that

$$\lim_{n\to\infty}\left(\sqrt{n+1} - \sqrt{n}\right) = 0\,. \tag{25.7}$$

Here we use a simple algebraic trick to rewrite this difference:

$$\sqrt{n+1} - \sqrt{n} = \frac{1}{\sqrt{n+1} + \sqrt{n}} \tag{25.8}$$

Given $\epsilon > 0$, we have that for $n \geq 1/\epsilon^2$ that

$$0 \leq \sqrt{n+1} - \sqrt{n} \leq \frac{1}{2\sqrt{n}} \leq \epsilon/2 < \epsilon\,. \tag{25.9}$$

Alternatively, we could have taken the point of view of Definition 25.1, and view $a_n = \sqrt{n} = f(n)$ as the values of a function f, first with domain $\mathbb{N}$, but then as the values of the function $f(x) = \sqrt{x}$ with domain $x > 0$, *evaluated* on the natural numbers $x = n$. This would allow us *to apply the mean value theorem*, to get immediately

$$\sqrt{n+1} - \sqrt{n} = f(n+1) - f(n) = f'(x) = \frac{1}{2\sqrt{x}} \leq \frac{1}{2\sqrt{n}} \to 0 \qquad (n \to \infty) \tag{25.10}$$

for some $x \in (n, n+1)$.

Example 25.6. Another example of a diverging sequence is $x_n = n^2$. Yet we write

$$\lim_{n\to\infty} x_n = \infty$$

if for every $C > 0$ there is an integer N such that $x_n > C$ whenever $n > N$.

Example 25.7. Another typical example is

$$\lim_{n\to\infty} \frac{3n^3 + 7n^2 + 1}{4n^3 - 8n + 63} = \frac{3}{4} \tag{25.11}$$

This is clear because the n^3 terms are the "leading order" terms for large n, and this can be turned into a proof by "dividing through" by n^3:

$$a_n = \frac{3n^3 + 7n^2 + 1}{4n^3 - 8n + 63} = \frac{3 + 7/n + 1/n^3}{4 - 8/n^2 + 63/n^3} \tag{25.12}$$

We could now proceed by first finding an explicit expression for $a_n - 3/4$, and then estimate $|a_n - 3/4|$, with the aim of verifying $a_n \to 3/4$ directly using Definition 25.2. It is easier, however, to apply at this stage the following limit laws.

For the evaluation of limits as in the last example the following facts are useful:

Theorem 25.1 (Limit Laws). *Suppose the infinite sequences $\{a_n\}$ and $\{b_n\}$ both have limits as $n \to \infty$, then*

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \tag{25.13}$$

$$\lim_{n\to\infty}(a_n \cdot b_n) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n\,. \tag{25.14}$$

Moreover, if $\lim_{n\to\infty} b_n \neq 0$, *then*

$$\lim_{n\to\infty}(a_n/b_n) = \lim_{n\to\infty} a_n / \lim_{n\to\infty} b_n\,. \tag{25.15}$$

Remark 25.1. The formulation of the last equality of limits actually requires some more care. As it stands we are considering the sequence $c_n = a_n/b_n$ which may not even be defined for all $n \in \mathbb{N}$. However, since $\lim_{n\to\infty} b_n \neq 0$, we know that for some N sufficiently large, $b_n \neq 0$ whenever $n \geq 0$. Moreover, *redefinining* all b_n, whenever $b_n = 0$ and $n \leq N$, obviously does not effect the statement about the limits.

The proof of these limit laws is so similar to the corresponding statements for limits of functions, that they will not be repeated here. Nonetheless, let us explore the similarity between the definition of limits of functions and sequences a little further.

Note for example that if f is a function that satisfies

$$\lim_{x\to\infty} f(x) = l \tag{25.16}$$

and we set $a_n = f(n)$, then

$$\lim_{n\to\infty} a_n = l\,. \tag{25.17}$$

This observation is often very useful:

Example 25.8. Let $0 < a < 1$, then

$$\lim_{n\to\infty} a^n = 0\,. \tag{25.18}$$

To prove this note that $a^x = e^{x \log a}$, and $\log a$ is negative, hence $\lim_{x\to\infty} e^{x \log a} = 0$.

Exercise 25.2. Show that for any $|a| < 1$, $\lim_{n\to\infty} a^n = 0$. Also show that, if $a > 1$ then $\lim_{n\to\infty} a^n = \infty$.

We can now finally return to (25.4). In other words, let us show that for any $a \in \mathbb{R}$, $\lim_{n\to\infty} a^n/n! = 0$. We can write for $n > N > 2a$,

$$|\frac{a^n}{n!}| \leq \frac{a^N}{N!}\frac{a\cdots a}{(N+1)\cdots n} \leq \frac{a^N}{N!}(\frac{1}{2})^{n-N} \to 0 \qquad (n\to\infty) \tag{25.19}$$

where in the last step we are using that $2^{-n} \to 0$ as $n \to \infty$ as we have shown in (25.18).

Finally the observation made in (25.16) and (25.17) still does not give us convergence of sequences like

$$a_n = \sin\Big(13 + \frac{1}{n^2}\Big) \tag{25.20}$$

$$b_n = \cos\Big(\sin(1 + (-1)^n\frac{1}{n})\Big) \tag{25.21}$$

which clearly should converge to $\sin(13)$, and $\cos\sin(1)$, respectively. (If this is not "clear" draw a few points on the line.) The theorem that allows us to conclude that is the following:

Theorem 25.2. *Let $c \in \mathbb{R}$, and f be function defined on an interval I that contains c, except perhaps at c itself, and suppose*

$$\lim_{x\to c} f(x) = l\,. \tag{25.22}$$

Suppose $\{a_n\}$ is a sequence such that each $a_n \in I$, $a_n \neq c$, and $\lim_{n\to\infty} a_n = c$. Then the sequence $\{f(a_n)\}$ converges, and

$$\lim_{n\to\infty} f(a_n) = l\,. \tag{25.23}$$

Conversely, if this is true for every *sequence $\{a_n\}$ satisfying these conditions, then* (25.22) *holds.*

Proof. If $\lim_{x\to c} f(x) = l$, then for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$|f(x) - l| < \epsilon \tag{25.24}$$

whenever $0 < |x - c| < \delta$. Now by assumption we can choose $N > 0$ such that

$$n > N \Longrightarrow |a_n - c| < \delta\,, \tag{25.25}$$

which then implies that $|f(a_n) - l| < \epsilon$, showing that $f(a_n) \to l$.

Conversely, if (25.22) were *not* true, then there exists $\epsilon > 0$, so that for *every* $\delta > 0$, there exists x with $|x - c| < \delta$, and $|f(x) - l| \geq \epsilon$. However, this can be used to define a sequence a_n of points with the property that say $|a_n - c| < 1/n$, but $|f(a_n) - l| \geq 1$. Then $a_n \to c$, but $f(a_n)$ does not converge to l, in contradiction ot (25.23).

□

Example 25.9.

$$\lim_{n\to\infty}\Big(1 + \frac{a}{n}\Big)^n = e^a\,. \tag{25.26}$$

Problems

1. Verify each of the following limits.
   a) $\lim_{n\to\infty} \frac{n}{n+1} = 1$
   b) $\lim_{n\to\infty} \frac{n+3}{n^3+4} = 0$
   c) $\lim_{n\to\infty} \frac{n!}{n^n} = 0$
   d) $\lim_{n\to\infty} \sqrt[n]{a} = 1, \qquad a > 0$
   e) $\lim_{n\to\infty} \sqrt[n]{n} = 1$
   f) $\lim_{n\to\infty} \sqrt[n]{n^2+n} = 1$

2. Find the following limits.
   a) $\lim_{n\to\infty}\left(\frac{n}{n+1} - \frac{n+1}{n}\right)$
   b) $\lim_{n\to\infty}\left(n - \sqrt{n+a}\sqrt{n+b}\right)$
   c) $\lim_{n\to\infty} \frac{2^n+(-1)^n}{2^{n+1}+(-1)^{n+1}}$
   d) $\lim_{n\to\infty} nc^n$, $|c| < 1$

3. a) Prove that if $0 < a < 2$, then $a < \sqrt{2a} < 2$.
   b) Prove that the sequence

$$\sqrt{2}, \quad \sqrt{2\sqrt{2}}, \quad \sqrt{2\sqrt{2\sqrt{2}}}, \ldots \tag{25.27}$$

   converges.
   c) Find the limit.

4. Let $0 < a_1 < b_1$ and define

$$a_{n+1} = \sqrt{a_n b_n}\,, \qquad b_{n+1} = \frac{a_n + b_n}{2}\,. \tag{25.28}$$

   a) Prove that the sequences $\{a_n\}$ and $\{b_n\}$ each converge.
   b) Prove that they have the same limit.

5. Find a sequence $\{a_n\}$ of points in $(0,1)$ such that $\lim_{n\to\infty} a_n$ is not in $(0,1)$.

6. Suppose that f is continuous and that the sequence

$$x,\ f(x),\ f(f(x)),\ f(f(f(x))),\ \ldots \tag{25.29}$$

   converges to l. Prove that l is a **fixed point** for f, i.e. $f(l) = l$.

Bibliography

Apostel, Tom M. *Calculus*. Second Edition. Vol. I. One-variable Calculus, with an Introduction to Linear Algebra. John Wiley & Sons, 1967.

Folland, Gerald. B. *Advanced Calculus*. Prentice Hall, 2002.

Osserman, Robert. *Two-Dimensional Calculus*. Ed. by Salomon Bochner and W.C. Lister. Harbrace College Mathematics Series. Harcourt, Brace and World, Inc., 1968.

Spivak, Michael. *Calculus*. Fourth Edition. Publish or Perish, Inc., 2008.

Walter, Wolfgang. *Gewöhnliche Differentialgleichungen*. Heidelberger Taschenbucher. Springer, 1972.