Calculus one[1]

A course instructed by Amir Yehudayoff, Department of Mathematics, Technion-IIT

[1]An apology: this text probably contains errors.

Contents

Chapter 1

Preliminaries

1.1 Formalities

This is first course at Technion, so some introduction to system is needed.

There are 4 hours with me, and 2 hours with TA per week.

About 6 homework exercises.

Mathnet: An electronic system of exercises. One every week.

A midterm exam and a final exam.

The final grade is determined by a combination of all the above. Details in Moodle, where all other material appears as well. Including Hebrew notes.

1.2 Goals

There are several goals to this course.

- Require basic understanding and ability to work with limits, derivatives, and integrals.

  These notions are at the basis of the scientific revolution, and understanding them is crucial to understating the world from a modern scientific perspective.

- Develop abstract thinking, and the ability to model a real life system in mathematical terms. Useful in understanding and establish predictions on the real world.

- Learn part of the language of mathematics.

1.3 Numbers

Simplest numbers are natural number $\mathbb{N} = \{0, 1, 2, \ldots\}$. If we want to have subtraction we get integers $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$. If we want to have division we get rationals $\mathbb{Q} = \{a/b : a, b \in \mathbb{Z}, b \neq 0\}$.

Mostly focus on real numbers $\mathbb{R}$, which corresponds to number of line.

A subset A of $\mathbb{R}$ is denoted $A \subset \mathbb{R}$. Can explicitly describe set e.g. $\{x \in \mathbb{R} : x > 0\}$.

Infinity is denoted ∞. We think of it as a point at end of line to the right. There is also $-\infty$, which is to the left. A ray is $[a, \infty) = \{x : a \leq x\}$.

We know

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

Obviously the left 2 containments are not equality. Is the third?

Claim 1. $\sqrt{2} \notin \mathbb{Q}$.

This was a kept secret in the era of Pythagoras and his friends, and legend has it that Hippasus discovered it and was executed to keep the secret.

Proof. Assume towards a contradiction that $\sqrt{2} = \frac{a}{b}$ for $a, b \in \mathbb{N}$ and $b \neq 0$. Among all options, choose a, b so that $a + b$ is minimum. Thus, $2 = \frac{a^2}{b^2}$ or

$$a^2 = 2b^2.$$

This implies that a is even. So, $a = 2c$, and

$$4c^2 = 2b^2$$

which means that b is even. So, $b = 2d$ and

$$\sqrt{2} = \frac{c}{d}$$

but $c + d < a + b$, a contradiction. □

The real numbers are the line. How do the other sets look on the line? Draw. The rationals are dense.

Claim 2. *If $a < b$ in $\mathbb{R}$ then there is $q \in \mathbb{Q}$ so that $a < q < b$.*

Proof. Let n be large enough so that $n > \frac{1}{b-a} > 0$. Let k be the smallest integer so that $\frac{k}{n} > a$ or $k > an$. Specifically, $(k-1) \leq an$ and so

$$a < \frac{k}{n} = \frac{k-1}{n} + \frac{1}{n} < a + (b - a) = b.$$

□

We may conclude that the irrationals $\mathbb{R} \setminus \mathbb{Q}$ are also dense in $\mathbb{R}$.

1.4 Absolute value

Define a function from $\mathbb{R}$ to the non negative reals by $|x| = x$ if $x \geq 0$ and $|x| = -x$ otherwise. Draw graph. In words, $|x|$ is the distance of x from 0, and in general $|x - y|$ is the distance between x, y.

Claim 3. *For all $x, y \in \mathbb{R}$:*

1. $|x + y| \leq |x| + |y|$.
2. $||x| - |y|| \leq |x - y|$.

Proof. For the first item, if the sign of x, y is the same, then there is equality, and if the sign are different on l.h.s. there is cancellation.

The second item follows similarly, or w.l.o.g. assume $|x| \leq |y|$ and

$$|x| = |x - y + y| \leq |x - y| + |y|.$$

□

1.5 Inequlities

An inequality is an expression of the form $A < B$ or $A \leq B$, where A, B are real numbers or expression taking real values. The former is a strong inequality, and the latter weak. It signifies that A is smaller than B, and that B is larger than A.

Rules: Assume the inequality $A < B$ holds.

- Given a number $c > 0$ then the inequality $cA < cB$ also holds.
- The inequality $-A > -B$ also holds.
- Given a second inequality $C < D$, the inequality $A + C < B + D$ also holds.

An example: Assume $A < B < C$ and $D < E < F$. What can we say about $B - E$? Well,

$$-F < -E < -D$$

and so

$$A - G < B - E < C - D.$$

1.6 Sums

Given a sequence/list of number $a_1, a_2, \ldots, a_n$, denote

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n.$$

The sum of an arithmetic sequence $a_i = a + i \cdot d$ is

$$\sum_{i=0}^{n} a_i = (2a + nd)\frac{n+1}{2}.$$

First term plus last term times number of summands over two. Proof: by induction.

The sums of a geometric sequence with $q \neq 1$

$$\sum_{i=0}^{n} a \cdot q^n = a\frac{q^{n+1} - 1}{q - 1}.$$

Proof: multiply both sides by $q - 1$ and see what happens.

1.7 Subsets of real line

There are several types of subsets of $\mathbb{R}$ that we are interested in.

Intervals and rays. A closed interval $[a, b] = \{x : a \leq x \leq b\}$ and an open interval $(a, b) = \{x : a < x < b\}$. There are half open interval too.

Bounded sets. A set $S \subset \mathbb{R}$ is called bounded from above if there is $x \in \mathbb{R}$ so that $s \leq x$ for all $s \in S$. The number x is called an upper bound on S.

Similarly, bounded from below.

A set is called bounded if both hold.

Examples: intervals are bounded, rays are bounded only from one side.

Every bounded set have many upper bounds. E.g. $[0, 1]$ has upper bounds $1, 3, 6$. Of particular importance is the smallest possible upper bound. This number is called the supremum: $\sup(S)$ is the smallest x that is an upper bound on S. E.g. $\sup([0, 1]) = 1$ and $\sup([0, 1)) = 1$. Similarly, define infimum as the largest lower bound on S.

Sometime the supremum of S is also inside S, like in the case of $[0, 1]$. In this case it is called the maximum of S. Some sets do not have a maximum, like $[0, 1)$. Similarly, define minimum.

Note: every set that is bounded from above has a supremum, but only some of these sets also have a maximum. Similarly, for infimum and minimum.

Neighbourhoods. Another special type of sets is a neighbourhood of a points $x \in \mathbb{R}$. Given $\delta > 0$, the δ-neighbourhood of x is the open interval $(x-\delta, x+\delta)$. In other words, it is the set of all points of distance smaller than δ from x. Hence the name. A third way to write

$$\{y \in \mathbb{R} : |y - x| < \delta\}.$$

In general, a neighbourhood of x (without specifying δ) is an open interval containing x, that is, a set of form (a, b) with $a < x < b$.

1.8 Functions

Given two sets A, B, a function $f : A \to B$ maps every element of A to an element of B, that is, for all $a \in A$ we have $f(a) \in B$.

There are several ways to define or describe a function:

Formulas. One way is by formulas $f(x) = x^2$.

Known names. There are some known functions, like $\sin, \cos, \tan$. We shall discuss them in more detail later on.

By words. For example, $f(x)$ is the largest integer that is smaller than x. This function is usually denoted $\lfloor x \rfloor$.

With several domains. For example,

$$f(x) = \begin{cases} 1 & x \geq 0, \\ 0 & x < 0. \end{cases}$$

We can have 3, 4 or infinitely many domains (e.g. $f(x) = n$ if $x \in [n, n+1)$ which we have seen above).

Comments: The set A is called the domain of f, and the set B the codomain.

Sometimes we need to deduce the domain. For example, the function $f(x) = \sqrt{x}$ is defined over $A = \{x : x \geq 0\}$.

If $f(x) = y$ then x is called the source of y and y is called the image of x.

More definitions:

- *Image.* The set

$$f(A) = \{f(a) : a \in A\}$$

is the image of f. For example, the image of $f(x) = x^2$ is the positive reals. It is sometime denoted as $im(f)$.

- *Graph.* The graph of the function is the set of points $\{(x, f(x)) : x \in A\}$. This is the usual notion that you draw. It is very helpful in visualising f.

- *Onto.* The function $f : A \to B$ is onto if $f(A) = B$.

- *One-to-one.* The function f is one-to-one if $f(x) \neq f(x')$ for all $x \neq x'$ in A.

- *Invertible.* A function $f : A \to B$ is called invertible if there is a function $g : B \to A$ so that $g(f(a)) = a$ for all $a \in A$ and $f(g(b)) = b$ for all $b \in B$.

The function g is called the inverse of f and is denoted by f^{-1}. Not to be confused with $1/f$.

Exercise: f is invertible iff f is one-to-one and onto.

- *Monotone.* A function f is (strongly) monotone increasing if $f(x) < f(x')$ for all $x < y$ in A. There is weak monotonicity (with $\leq$), and also monotone decreasing.

Exercise: if f is strongly monotone then f is one-to-one.

Question: does the opposite hold?

Answer: No – 2 separate regions.

- *Even or odd.* A function f is called even if $f(x) = f(-x)$. Draw. It is called odd if $f(x) = -f(-x)$. Draw.

- *Bounded.* A function $f : A \to B$ is bounded, if its image $f(A)$ is bounded. Similarly, bounded from above and below.

1.9 Elementary functions

Polynomials. A function of the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = \sum_{i=0}^{n} a_i x^i$$

is called a polynomial. If $a_n \neq 0$ then the degree of f is n. Degree 0 is a constant, degree 1 is a linear function, degree 2 quadratic, and so forth.

A point $x \in \mathbb{R}$ is called a root of f if $f(x) = 0$. Finding roots of polynomials, and of general functions, is of importance in applications.

Theorem 4. *A polynomial of degree n has at most n roots.*

The image of a polynomial of degree n is: If n is even, a half line. If n is odd, all of $\mathbb{R}$. Draw it.

Rational functions. If $p(x), q(x)$ are polynomials so that $q(x) \neq 0$ then

$$f(x) = \frac{p(x)}{q(x)}$$

is a rational function (like that rationals are rations of integers). The domain of f is

$$\{x : q(x) \neq 0\}.$$

If $deg(q) = n$ then there are at most n such points.

Exponentials. For $a > 0$, the function $f : \mathbb{R} \to \mathbb{R}$

$$f(x) = a^x$$

is called an exponential.

Known rules:

$$a^{x+y} = a^x a^y, \ (a^x)^y = a^{xy}, \ a^{-x} = \frac{1}{a^x}$$

for all $x, y \in \mathbb{R}$. This is well known for integers, and is also true for all reals (will not prove here).

Always $f(0) = 1$. If $a > 1$ then $f(x)$ is increasing, and if $a < 1$ it is decreasing. Two graphs. The image of f is $\{x \in \mathbb{R} : x > 0\}$.

Trigonometric. There are several such functions. Most notably $\sin(x), \cos(x), \tan(x)$. In mathematics, usually angles are measured in radians, which is a number in $[0, 2\pi)$. Like an angle in $[0, 360)$ in degrees. Draw a circle of radius one, and angles of

$$2\pi\{1/8, 1/3, 1/4\}.$$

Focus on $x = 2\pi/10$. What is $\cos, \sin$ in the drawing? $\cos(x)$ is length of line close to angle, and $\sin(x)$ is length of line in front of angle. Pythagoras's theorem:

$$\sin^2(x) + \cos^2(x) = 1.$$

Symmetry:

$$\sin(x) = \cos((\pi/2) - x).$$

There are many more useful equalities.

This drawing defines the functions for $x \in [0, 2\pi)$. What about other $x \in \mathbb{R}$? The same, just go around two or more time, or go in the other direction (negative). So,

$$\sin(x) = \sin(x + 2\pi), \cos(x) = \cos(x + 2\pi),$$

for all x. These functions are periodic. Draw. The image is $[-1, 1]$.

Tangent is

$$\tan(x) = \frac{\sin(x)}{\cos(x)}.$$

It is also periodic with period π. It is enough to draw in $[-\pi/2, \pi/2]$. The domain is $\{x : \cos(x) \neq 0\}$ which is the set of all $x \neq \frac{\pi}{2} + \pi z$ for some $z \in \mathbb{Z}$. It goes from $-\infty$ at $-\pi/2$ to $+\infty$ at $\pi/2$ through 0 at 0.

1.10 Operations on functions

Arithmetic. Given $f, g : A \to B$ we can define several more functions from A to B, by adding, subtracting, multiplying and dividing. What is domain of division?

Composition. Given $f : A \to B$ and $g : B \to C$ we can define a new function $g \circ f : A \to C$ by

$$(g \circ f)(a) = g(f(a)).$$

"Just follow the arrows in the diagram."

Example: $f(x) = 2x + 3$ and $g(x) = \frac{1}{x}$. Note that g is not defined over $f(\mathbb{R})$ but only over a subset of it (the set $\{x : x \neq 0\}$). The function

$$(g \circ f)(x) = g(2x + 3) = \frac{1}{2x + 3}.$$

Its domain is $\{x : x \neq -3/2\}$.

Conclusion: when composing functions we need to keep track of domains (the "weakest link" determines the domain).

Getting more elementary functions. In general, if f is elementary, also f^{-1} is elementary, and arithmetic operations on elementary functions give elementary functions. Let us see some examples.

Roots. The function $f(x) = x^2$ is not one-to-one so is not invertible. But if we restrict it to the domain $A = \{x : x \geq 0\}$ then it becomes one-to-one, and onto $B = \{x : x \geq 0\}$. It therefore become invertible. Its inverse is $f^{-1}(x) = \sqrt{x}$. Draw it. Discuss connection of graph of f to graph of f^{-1}.

Logarithms. The inverse of an exponential is called a logarithm. If $f(x) = a^x$ for $a \neq 1$ positive then f is invertible ($f(\mathbb{R})$ is the positive reals). The inverse $f^{-1}(x)$ is defined over $\{x : x > 0\}$ and is denoted $\log_a(x)$. That is,

$$a^x = y \;\Leftrightarrow\; x = \log_a(y).$$

From the rules of exponentials, we get rules for logs

$$\log_a(xy) = \log_a(x) + \log_a(y),\; \log_a(x^y) = y \log_a(x),\; \log_a(1/x) = -\log_a(x).$$

Draw it.

If $a > 1$ it is increasing and if $a < 1$ it is decreasing (like exponentials).

Trigonometric. There trigonometric functions also have inverses (when restricted to the correct domains). Sometimes they are denoted with "arc" for example $\arccos(x) =$

$\cos^{-1}(x)$, which is defined from $[-1,1]$ to say $[0,\pi]$.

Summary. There are many elementary functions. We can get more by operations we saw (arithmetic and inverse).

Chapter 2

Limits

A limit is one of the fundamental topics that we shall discuss in this course. We will be using the language of mathematics, but it will be helpful to keep in mind that we are just setting the ground to understanding more complicated notions that are helpful in analysing and understanding complex real life systems.

As in any mathematical discussion, our discussion consists of definitions, theorems, and proofs. Definitions provide exact and formal meaning to words and symbols. Theorems are statements concerning the definition that are always true. Proofs use the definitions and the rules of logic to demonstrate the the theorems are indeed true.

2.1 At infinity

Let us start with a formal definition. Later we shall provide intuition and examples.

Definition 5 (Limit at infinity). *Let $f : \mathbb{R} \to \mathbb{R}$. The f has a limit at $+\infty$ and its value is $L \in \mathbb{R}$ if for every $\epsilon > 0$, there is $x_0 = x_0(\epsilon) \in \mathbb{R}$ so that for all $x > x_0$,*

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

The limit is denoted as

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

This definition also makes sense when f is defined in a neighbourhood of infinity, that is, on a ray of the form $[a, \infty)$ for some $a \in \mathbb{R}$.

Roughly, it says that if x is large then $f(x)$ is close to L. Draw. It is important that L is a real number (and not say ∞).

Examples:

1.
$$\lim_{x\to\infty} 1 = 1.$$

2.
$$\lim_{x\to\infty} \frac{1}{x} = 0.$$

Let us prove from definition. Let $\epsilon > 0$. Set $x_0 = 1/\epsilon$. For every $x > x_0$,

$$\left|\frac{1}{x} - 0\right| = \frac{1}{x} < \frac{1}{x_0} = \epsilon.$$

3. Recall $\tan^{-1} : \mathbb{R} \to (-\pi/2, \pi/2)$. What is

$$\lim_{x\to\infty} \tan^{-1}(x) =?.$$

Draw. tan of what is a large number? Of numbers that are close to $\pi/2$. Shall not prove.

4.
$$\lim_{x\to\infty} 2^{-x} = 0.$$

Draw.

5.
$$\lim_{x\to\infty} \frac{\sin(x)}{x} =?$$

Draw. Comes closer and closer to zero, but from both sides, so it is 0. There are two enveloping curves of $1/x$ and $-1/x$. Let us prove from definition. Let $\epsilon > 0$. Set $x_0 = 1/\epsilon$. For every $x > x_0$, since $|\sin(x)| \le 1$,

$$\left|\frac{\sin(x)}{x} - 0\right| = \left|\frac{\sin(x)}{x}\right| \le \left|\frac{1}{x}\right| < \frac{1}{x_0} = \epsilon.$$

<u>Non examples:</u>

1. $\sin(x)$ does not have a limit at ∞. There are infinitely many x's so that $\sin(x) = 1$, and also so that $\sin(x) = -1$.

2. x does not have a limit in $\mathbb{R}$.

We have seen the definition, and several examples and non examples. Let us consider a few properties of limits.

2.2 Arithmetic

What is

$$\lim_{x\to\infty} 1 + \frac{1}{x} + 2^{-x}?$$

It is more convenient to consider each term separately and then compose the answers, using simple arithmetic. The guess would be

$$1 + 0 + 0 = 1.$$

To do so, we need to know when it is okay.

Theorem 6. *Let $f_1, f_2 : \mathbb{R} \to \mathbb{R}$. Assume that (specifically the two limits exist and are finite)*

$$\lim_{x\to\infty} f_1(x) = L_1, \ \lim_{x\to\infty} f_2(x) = L_2.$$

Then,

$$\lim_{x\to\infty} f_1(x) + f_2(x) = L_1 + L_2, \lim_{x\to\infty} f_1(x) - f_2(x) = L_1 - L_2, \lim_{x\to\infty} f_1(x) \cdot f_2(x) = L_1 \cdot L_2.$$

If $L_2 \neq 0$ then

$$\lim_{x\to\infty} \frac{f_1(x)}{f_2(x)} = \frac{L_1}{L_2}.$$

We shall not prove the theorem here, but keep it in mind since it will be useful for you. When you apply a theorem, remember to verify that you are allowed to.

Example:

$$\lim_{x\to\infty} \frac{x}{x+1} =?$$

Why? We can use the definition of a limit, or use known limits. The second one is usually easier. To do so, write

$$\frac{x}{x+1} = 1 - \frac{1}{x+1},$$

and recall that the limit of $1/(x + 1)$ when $x \to \infty$ is 0.

2.3 Basic properties

We shall present several basic properties of limits. They have two other objectives. One is the deepen the understanding of the notion of limit. Another is learn how to mathematically investigate a concept.

Claim 7. *If $f : \mathbb{R} \to \mathbb{R}$ has a limit at infinity, then there $x_0, b \in \mathbb{R}$ so that for all $x > x_0$,*

$$|f(x)| \leq b.$$

In other words, if f has a limit at infinity then the image of f on $[x_0, \infty)$ is bounded for some x_0.

Proof. Denote by L the limit of f when $x \to \infty$, which exists by assumption. Choose $\epsilon = 1$. Since f has limit, there is x_0 so that $|f(x) - L| < 1$ for all $x > x_0$. Finally, choose $b = |L| + 1$. Indeed, for all $x > x_0$,

$$|f(x)| = |f(x) - L + L| \leq |f(x) - L| + |L| \leq 1 + |L| = b.$$

□

Claim 8. *If $\lim_{x\to\infty} f(x) = L > 0$ then there is $x_0 \in \mathbb{R}$ so that for all $x > x_0$,*

$$f(x) > 0.$$

Proof. Choose $\epsilon = L/2$. There is x_0 so that for all $x > x_0$,

$$|f(x) - L| \leq L/2.$$

That is, for $x > x_0$,

$$-L/2 \leq f(x) - L \leq L/2$$

and so (just left hand side)

$$0 < L/2 \leq f(x).$$

□

Exercise 9. *The limit of f at infinity is 0 if and only if the limit of $|f|$ at infinity is 0.*

Exercise 10. *Let $f, g : \mathbb{R} \to \mathbb{R}$. Assume that there is $x_0 \in R$ so that for all $x > x_0$ we have $f(x) \geq g(x)$. Also assume that f, g have a limit at infinity. Then,*

$$\lim_{x\to\infty} f(x) \geq \lim_{x\to\infty} g(x).$$

Comment: even if $f(x) > g(x)$ for all x, in the limits only $\geq$ holds. For example, $g(x) = 0$ and $f(x) = 1/x$.

Claim 11 (Sandwich)**.** *Let $f, g, h : \mathbb{R} \to \mathbb{R}$. Assume that there is $x_0 \in \mathbb{R}$ so that for all $x > x_0$ we have*

$$f(x) \leq g(x) \leq h(x).$$

Also assume

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} h(x) = L.$$

Then,

$$\lim_{x\to\infty} g(x) = L$$

as well.

This claim may be useful in computing a limit. For example, for $x > 0$,

$$0 \le \frac{1}{x+\sqrt{x}} \le \frac{1}{x}.$$

Proof. By previous exercise,

$$L = \lim_{x\to\infty} f(x) \le \lim_{x\to\infty} g(x) \le \lim_{x\to\infty} h(x) = L.$$

□

Another variant of the arithmetic rules:

Claim 12. *Let f, g be two functions that are defined on a ray $[a, \infty)$. Assume $\lim_{x\to\infty} f(x) = 0$ and that g is bounded. Then,*

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

This following holds even when g does not have a limit.

Proof. Follows from sandwich, since there is $M \in \mathbb{R}$ so that $-M \le g(x) \le M$ for all x, so

$$-M|f(x)| \le g(x) \le M|f(x)|.$$

Also use that $|f|$ tends to 0. □

Comment: The discussion above can be made on limits at $-\infty$ by considering $-f$ instead of f.

2.4 Limit at a point

So far we have discussed limits at infinity. We shall move to discussing limits at a point $x \in \mathbb{R}$.

Definition 13. *The function $f : A \to B$ has a limit L at*[1] *$a \in A$ if for every $\epsilon > 0$, there is $\delta > 0$ so that for all $x \in A$ so that*[2] *$0 < |x - a| < \delta$ we have $|f(x) - L| < \epsilon$.*

A neighbourhood of infinity $[x_0, \infty)$ is replaced by a neighbourhood of a point. The number x_0 is replaced by δ. Draw.

Comment: There is an important distinction between the case f is defined at a and not. We shall not discuss it in detail.

Examples:

1. $$\lim_{x \to 1} x = 1.$$

2. $$\lim_{x \to 0} \frac{x}{x^2 + 1} = \frac{0}{1} = 0.$$

3. $$\lim_{x \to 0} \frac{x}{x} = 1.$$

   The function is not defined at 0.

Non examples:

1. $\lim_{x \to 0} \frac{1}{x}$ is not defined.

2. Let f be the step function (i.e. 0 on negative number and 1 otherwise). Then $\lim_{x \to 0} f(x)$ is not defined. Intuitively, if we approach from right we get 1 and from left we get 0. We shall discuss it in more detail soon.

Comment: The claims we stated for limits at infinity holds also at a point (e.g. arithmetic, boundedness, positivity, bounded times tending to zero).

One more example: The limit of $\sin(1/x)$ does not exist at 0. Draw. However, $\sin(1/x)$ is bounded so that limit of $x \sin(1/x)$ at 0 exists and it is 0.

Puncturing: Sometimes f is not even defined at a, but the limit is still defined. This happens when f is defined in a punctured neighbourhood of a. Recall that a neighbourhood I of a is an open interval containing a. A punctured neighbourhood J of a is defined as $J = I \setminus \{a\}$. We just remove a from I.

[1] We shall always assume that A contains a punctured open neighbourhood of a, but we shall not mention it too often.

[2] Note that $x \neq a$ is also assumed.

2.5 One sided limits

The limit of f at x is L means that when we approach x the value of f becomes closer and closer to L. What if we approach x just from the left or right? This may give a different behaviour. For example, the step function does not have a limit at 0. But we can still make sense of approaching 0 from say the left.

Definition 14. *The function $f : A \to B$ has a limit from the left L at $a \in A$ if for every $\epsilon > 0$, there is $\delta > 0$ so that for all $x \in A$ so that $a - \delta < x < a$ we have $|f(x) - L| < \epsilon$. This limit is denoted*

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

That is y is close to x but to its left.

Example: Let f be the step function from above. Observe $f(0) = 1$. But for all $x < 0$ we have $f(x) = 0$, so

$$\lim_{x \to 0^-} f(x) = 0.$$

From right: We can similarly define a limit from the right. If f is the step function then

$$\lim_{x \to 0^+} f(x) = 1.$$

Claim 15. *Let $f : A \to B$ and $a \in A$. The limit of f at a exists if and only if the two one sides limits exist and they are equal.*

Namely, if one of the one sided limits does not exists or they exist but not equal then the limit does not exist. This is a way for proving that a certain limit does not exist.

Intuitively, if the two one sided limits exists, but the limit does not exist then the graph of the function has a "jump" at the point.

As in the case of limit, there is a one sided analog of the claims we have stated.

2.6 Generalized limits

We saw some examples where there is no limit, but we have an intuition that the limit is say ∞. Such a case is called a generalized limit.

Definition 16. *The function $f : A \to B$ has a generlized limits ∞ at a if for every $M > 0$, there is $\delta > 0$ so that for all $x \in A$ so that $x \neq a$ and $|x - a| < \delta$ we have $f(x) > M$.*

Example: $1/x$ at 0, or x at ∞.

Important comment: The arithmetic and other claims do not hold in general for generalised limits.

Can similarly define a generalised limit to be $-\infty$.

2.7 Summary

We defined limits at $\pm\infty$, at a point. One sides limits too. All of these have many properties: arithmetic, comparsion, sandwich,... We have defined generalised limits, and these are more subtle.

Chapter 3

Sequences

So far we have discussed functions and their limits. We now move to talk about sequences of numbers, which are an ordered list of numbers.

Definition 17. *A sequence* $(a_n)_{n=1}^{\infty}$ *is a list* $a_1, a_2, a_3, \ldots$ *of real numbers. It is sometimes denoted as* $\{a_n\}$ *as well.*

As functions, they can be written in several ways.

Formula. $a_n = n^2$ means $1, 4, 9, 16, \ldots$

Recursion. $a_0 = 0$, $a_1 = 1$, $a_{n+2} = a_{n+1} + a_n$ for all $n \geq 0$. Which is the Fibonacci sequence $0, 1, 1, 2, 3, 5, 8, \ldots$

3.1 Limits

Functions have limits at points (or infinity). Sequences just have a limit, which corresponds to infinity for functions.

Definition 18. *A sequence* (a_n) *has a limit* $L \in \mathbb{R}$ *at infinity if for every* $\epsilon > 0$ *there is* $x_0 \in \mathbb{R}$ *so that for all* $n > N$,

$$|a_n - L| < \epsilon.$$

It is denote $\lim_{n\to\infty} a_n$.

Examples: 1, $1/n$, 2^{-n}.

Finbonacci a_n, and $a_{n+1}/a_n \to (1+\sqrt{5})/2$. Shall later understand better.

If a sequence has a limit we say it converges. Otherwise, we say it diverges.

All statement about limits of functions at infinity hold for sequences as well:

Claim 19. *If* (a_n) *is given as* $a_n = f(n)$ *for some* $f : \mathbb{R} \to \mathbb{R}$ *and* $\lim_{x\to\infty} f(x) = L$ *then* $\lim_{n\to\infty} f(n) = L$ *as well.*

Comment: if changing (a_n) in finitely many places then we do not change the limit nor its existence.

The arithmetic of limits holds in this case too. For example, if $a_n \to a$ and $b_n \to b \neq 0$ then $a_n/b_n \to a/b$.

Example: $\lim_{n\to\infty} \frac{n^3+2n+4}{5n^3+10n^2-1} = 1/5$.

The comparison principle holds in this case too. If $a_n \geq b_n$ starting from some N and both converge then $\lim a_n \geq \lim b_n$. Sandwich principle holds too.

There are several more statements.

We mention a few more statement, mostly without a proof.

Claim 20. *If $a_n \to a$ then $|a_n| \to |a|$.*

Proof. If $a = 0$ then we have seen it. If $a > 0$ then we saw that $a_n > 0$ for all $n > n_0$. So $|a_n| = a_n$ for all $n > n_0$. If $a < 0$ then $|a_n| = -a_n$ which means $|a_n| = -a_n \to -a = |a|$. □

Claim 21. *If $a_n \to a > 0$ then $\sqrt{a_n} \to \sqrt{a}$.*

Claim 22. $a_n = n^{1/n} \to 1$.

Examples: $10^{1/10} = 1.25...$, $100^{1/100} = 1.047...$, $1000^{1/1000} = 1.006....$

Proof. Let us use sandwich. Write $a = 1 + b_n$. That is, $n^{1/n} = 1 + b_n$ or

$$n = (1 + b_n)^n = 1 + \binom{n}{1} b_n + \binom{n}{2} b_n^2 + \dots .$$

We want to show that b_n is small. Since by above $n > \binom{n}{2} b_n^2$,

$$0 \leq b_n < \sqrt{\frac{2}{n-1}} \to 0.$$

□

Corollary 23. *For all $c > 0$, we have $c^{1/n} \to 1$.*

Proof. Sandwich: $1 \leq c^{1/n} \leq n^{1/n}$ for all $n \geq c$. □

Generalized limits.

Definition 24. *We say that the (generalized) limit of a_n is ∞ if for every $M \in \mathbb{R}$ there is $x_0 \in R$ so that $a_n > M$ for all $n > x_0$.*

Claim 25. *If $a_n > 0$ for all n then $a_n \to 0$ iff $1/a_n \to \infty$.*

Some positivity assumption should be made: $\sin(n)/n$ tends to 0 but $n/\sin(n)$ does not tend to ∞.

3.2 Montone bounded sequences

Definition 26. *A sequence is monotone (weakly) increasing if* $a_{n+1} \geq a_n$ *for all* n*. Decreasing if* $a_{n+1} \leq a_n$*. If one of the two holds, we call it monotone.*

Definition 27. *A sequence* (a_n) *is bounded if there is* $M \in \mathbb{R}$ *so that* $|a_n| \leq M$ *for all* $n \geq 1$*.*

A key property we shall use is the following:

Theorem 28. *If* (a_n) *is monotone and bounded then it converges.*

For example, if there is $M \in \mathbb{R}$ so that

$$a_1 \leq a_2 \leq a_3 \leq \ldots \leq M$$

then (a_n) converges.

Explanation. We shall not prove it here but the intuition is that a_n keeps going up, but it is bounded so it can not go up too much, so it must be stuck at some point.

What is the limit? It is the supremum of the set $\{a_n\}$. □

This property is very useful – it actually gives us a way to compute a limit.

Example: Define $a_1 = 1$ and $a_{n+1} = \sqrt{1 + a_n}$. So it is $1, \sqrt{2}, \sqrt{1+\sqrt{2}}, \ldots$. Does it converge? What is the limit? Perhaps surprisingly it is $(1 + \sqrt{5})/2$, again.

We will see that it is bounded and monotone, and hence converges. Then we shall see the limit.

Claim 29. $a_n \leq a_{n+1}$ *for all* $n \geq 1$*.*

Proof. By induction. Base: $n = 1$ is true. Step: By induction,

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

□

Claim 30. $a_n \leq 2$ *for all* $n \geq 1$*.*

Proof. By induction. Base holds. Step:

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

□

Corollary 31. *The sequence* (a_n) *converges. Denote by* a *its limit. Thus*

$$a = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \sqrt{a_n + 1}.$$

By previous properties, $a = \sqrt{a+1}$ *or*

$$a^2 - a - 1 = 0 \Rightarrow a = \frac{1 \pm \sqrt{5}}{2}.$$

Note that $a > 0$ *since the sequence is positive. So we found another sequence with the golden number as its limit.*

Do not use arithmetics without first verifying that the limit exists.

3.2.1 e

Another important example of a sequence is $a_n = (1 + 1/n)^n$. We will show as before that it has a limit.

Claim 32. (a_n) *is monotone increasing and bounded.*

Proof. Bounded:

$$a_n = \sum_{k=0}^{n} (1/n)^n \binom{n}{k} \leq \sum_{k=0}^{n} 1/k! \leq 1 + 1 + 1/2 + 1/4 + 1/8 + \ldots = 3.$$

Monotone: Shall use the arithmetic-mean-geometric-mean inequality: For all $m \in \mathbb{N}$ and $b_1, \ldots, b_m$,

$$(b_1 b_2 \ldots b_m)^{1/m} \leq (b_1 + \ldots + b_m)/m.$$

For example, $(2 \cdot 3)^{1/2} = \sqrt{6} < (2+3)/2 = 2.5$. Now,

$$(1 \cdot (1 + 1/n)^n)^{1/(n+1)} \leq \frac{1}{n+1}(1 + n(1 + 1/n)) = 1 + 1/(n+1).$$

Raising both sides to power $n+1$ completes the proof. □

So we know this sequence converges. What is a limit? It has a special symbol

$$\lim_{n\to\infty} (1 + 1/n)^n = e \approx 2.71828182845904523536 0.$$

This is a very useful number which appears in many places in nature. Perhaps we shall understand this better later on.

Comments: For every $\lambda \in \mathbb{R}$,

$$\lim_{n\to\infty}(1+\lambda/n)^n = e^\lambda$$

and similarly

$$\lim_{x\to 0}(1+\lambda x)^{1/x} = e^\lambda.$$

We shall not prove these now.

3.2.2 A converging sum

Let

$$a_n = \sum_{i=1}^{n} \frac{1}{i^2}.$$

Write a few elements for example. This is clearly monotone and positive. Why is it bounded? Well, we can use a telescopic sum:

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

So this infinite sum converges. What is limit? This is a nontrivial question, and it turns out to be $\pi^2/6$.

3.3 Subsequences

Let us see some example of subsequences.

The even numbers are a subsequence of the integers.

$a_n = 2^{-n}$ is a subsequence of $b_n = 1/n$.

There are 2 main reasons to consider subsequences.

1. If we find two subsequences that converge to different limits then we know that the sequence does not converge. For example, $0, 1, 0, 1, 0, 1, 0, \ldots$.

Claim 33. *If (a_n) converges to L then every subsequence of (a_n) converges to L, and vice versa.*

2. It may be that some subsequence converges even though the sequence does not converge. For example, $0, 1, 0, 1, 0, 1, 0, \ldots$.

The general phenomenon is relevant to bounded sequences.

Theorem 34 (Bolzano-Weirstrass)**.** *Let* (a_n) *be a bounded sequence. Then it has a subsequence that converges to a limit.*

Explanation. We shall not formally prove, but give a brief explanation. It will help if you draw the following. So, there is $M \in \mathbb{R}$ so that $|a_n| \leq M$ for all n, or $a_n \in [-M, M]$ for all n. So either there are infinitely many n so that $a_n \in [-M, 0]$ or in $[0, M]$ (or in both). Assume that the latter holds. So, either there are infinitely many n so that $a_n \in [0, M/2]$ or in $[M/2, M]$. We can keep on going and find for every ϵ an interval of length ϵ in which there are infinitely elements of the sequence. This gives a converging subsequence – all elements of it become closer and closer. □

Chapter 4

Continuity

We are now ready to introduce the first application of the relative abstract notion of "limit." Our goal will be to define the notion of a continuous function. This roughly corresponds to functions whose graph is a continuous line. But our goal is to provide a formal definition. This may be thought of as one of the first motivations for considering the notion of a limit.

How would you define this notion? A function f is continuous if the value $f(x)$ is close to $f(a)$ if x is close to a.

Definition 35. *A function $f : \mathbb{R} \to \mathbb{R}$ is continuous at the point $a \in \mathbb{R}$ if*[1]

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

Examples:

The function $f(x) = x$ is continuous in every $a \in \mathbb{R}$ since $\lim_{x \to a} x = a$.

The function $\sin(x)$ is continuous:

$$|\sin(x) - \sin(a)| = 2\left|\sin\frac{x-a}{2}\cos\frac{x+a}{2}\right| \leq 2\left|\sin\frac{x-a}{2}\right| \leq 2\left|\frac{x-a}{2}\right| = |x-a|.$$

A non example. A step function at 0.

4.1 One sided

As in case of limits, we can define one sided continuity.

[1]We had some discussion related to the meaning of "limit" when f is defined at a. Here we take only $x \neq a$ in definition of limit, that is, for every $\epsilon > 0$ there is $\delta > 0$ so that if $0 < |x-a| < \delta$ then $|f(x) - f(a)| < \epsilon$.

Definition 36. *The function f is right continuous at a if*

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

Similarly, define left continuous.

For example, the 0/1 function defined as $f(x) = 1$ iff $x \geq 0$ is right continuous but not left.

Claim 37. *f is continuous iff it is continuous from both sides.*

4.2 Properties

As for limits, we can generate new functions from two given functions.

Proposition 38 (arithmetic)**.** *Let f, g be two functions that are continuous at a. Then $f + g, f \cdot g$ are continuous at a. If $g(a) \neq 0$ then f/g is continuous at a as well.*

What about composition $f \circ g$? They should both be continuous, but where? g at a, and f at $g(a)$.

Proposition 39 (composition)**.** *Let g be continuous at a and f be continuous at $g(a)$ then $f \circ g$ is continuous at a.*

What about inverses? If a line is continuous, then the line of inverse it too...

Claim 40 (inverse)**.** *Let f be strictly monotone and continuous at every point a domain A. Then f^{-1} is also continuous.*

We shall not prove these claims now, but the proofs follow from the known properties of limits.

As a conclusion we learn that

Theorem 41. *An elementary function f is continuous in all the points in which it is defined.*

4.3 Classification

There are 3 types of points of discontinuity. We go from simpler to most complicated.

Removable (slika). The function f has a removable discontinuity at a if the limit $\lim_{x \to a} f(x) = L$ exists but $f(a) \neq L$ (here we insist that $x \neq a$ in the definition of limit).

For example, f that is 0 everywhere except at 0. (We had some discussion of limit at 0 of this function. Here we say that the limit exists.)

Jump (kfitza). The function f has a jump discontinuity at a if the two sided limits $\lim_{x\to a^+} f(x)$ and $\lim_{x\to a^-} f(x)$ exist but are not equal.

For example, f is the step function at 0.

Essential (ikarit). The function f has an essential discontinuity at a if at least one of the sided limits $\lim_{x\to a^+} f(x)$ and $\lim_{x\to a^-} f(x)$ does not exist.

For example, $\sin(1/x)$ at 0.

4.4 Intermediate point theorem

Here we describe a simple and very useful property of continuous function. The property is very intuitive, but we shall not provide a formal proof. Roughly, it says that a line from A to B must cross the middle between A, B. Draw.

Theorem 42. *Let $f : [a, b]$ be a continuous function on $[a, b]$. For every y that is between $f(a)$ and $f(b)$, there is $x \in [a, b]$ so that $f(x) = y$.*

Explanation. Assume for example $f(a) = -1, f(b) = 1$ and $z = 0$. We can find a sequence of intervals $[a, b], [a_1, b_1], [a_2, b_2]$ so that

- $|b_n - a_n| \to 0$.
- $f(b_n) > 0$ and $f(a_n) < 0$.

The second property using continuity say that $|f(b_n) - f(a_n)| \to 0$. The first property implies that both go to 0. □

There are many applications of this theorem. Here is one.

Claim 43. *If p is a polynomial of odd degree then there it has a root, that is, there is x so that $p(x) = 0$.*

Proof. Let c be the leading coefficient of p. Assume that $c > 0$, the proof in the other case is similar. We know that if $a > 0$ is very large then $p(a) > 0$ and $p(-a) < 0$. So there is $x \in [-a, a]$ so that $p(x) = 0$. □

4.5 Maximum and minimum

Another important property of continuous functions is that they attain a maximum and minimum on closed intervals.

Theorem 44 (Weirstrass). *Let $f : [a, b] \to \mathbb{R}$ be a continuous function on $[a, b]$. Then f is bounded and attains a maximum and a minimum.*

We shall not prove the theorem here. But we comment that it is important that the interval is closed. For example, $1/x$ is continuous in $(0, 1]$ but not bounded.

Here is a application of this theorem.

Claim 45. *If $f : [a, b] \to \mathbb{R}$ is continuos then the image $f([a, b])$ of f is an interval.*

The image is not necessarily $[f(a), f(b)]$!

Explanation. Maximum and minimum are attained. Everything between them is attained due to intermediate point theorem. □

Chapter 5

Derivatives

We have defined and studied one basic aspect of functions: continuity. We now move to studying a more complicated notion. The introduction of this notion by Newton and Leibniz is one of the reasons for the scientific revolution.

The basic idea is that say we are looking at a particle moving in space, and we want to measure its speed. Let us say for simplicity the particle is a car on a road. How can we measure the speed? We can let it go from Tel Aviv to Haifa and estimate

$$\text{speed} = \frac{\text{distance from Tel Aviv to Haifa}}{\text{time it took the car}}.$$

But this is only an average speed. How can we measure the speed of a car at a given moment? We take shorter and shorter parts of the road, and do the same calculation. In other words, we take a limit.

Definition 46. *Let $f : \mathbb{R} \to \mathbb{R}$. The derivative of f at $a \in \mathbb{R}$ is*

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

if the limit exists.

The derivative of x^2 at 1 is

$$\lim_{x \to 1} \frac{x^2 - 1^2}{x - 1} = \lim_{x \to 1} x + 1 = 2.$$

In general, the derivative of x^n at a is (the first step is useful in many cases)

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

This is a formula you all know.

There is a different notation: Think of $f(x)-f(a)$ as a small change in f, which is denoted df (d from difference). Think of $x-a$ as a small change in x. The notation is

$$f'=\frac{df}{dx}.$$

This is not division.

Geometric interpretation. Draw f in the neighbourhood of a. The height is $f(x)-f(a)$ and the width is $x-a$. The limit is the slope of the tangent line to f at a. A tangent is not necessarily on one side of the graph: x^3 at 0.

More examples. Derivative of sin:

$$\begin{aligned}\lim_{h\to 0}\frac{\sin(a+h)-\sin(a)}{h} &= \lim_{h\to 0}\frac{2\sin(\frac{a+h-a}{2})\cos(\frac{a+h+a}{2})}{h}\\ &= \lim_{h\to 0}\frac{2\sin(\frac{h}{2})}{h}\cos(\frac{2a+h}{2})=1\cdot\cos(a).\end{aligned}$$

5.0.1 Differentiability and continuity

Claim 47. *If f has a derivative at a then f is continuous at a.*

Proof.

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

□

The other direction is not true. For example, $|x|$ at 0:

$$\lim_{h\to 0^+}\frac{|h|}{h}=1$$

and

$$\lim_{h\to 0^-} \frac{|h|}{h} = -1.$$

5.1 Rules

Theorem 48. *Assume f, g are differentiable at a.*

scalar. *For every $c \in \mathbb{R}$, we have $(cf)'(a) = c(f'(a))$.*

additivity. *$(f+g)'(a) = f'(a) + g'(a)$.*

Leibniz. *$(f \cdot g)'(a) = f'(a)g(a) + f(a)g'(a)$.*

Explanation. The first 2 items follow from definition of limit. Item 3:

$$\begin{aligned}\lim_{x\to a} \frac{f(x)g(x) - f(a)g(a)}{x-a} &= \lim_{x\to a} \frac{f(x)g(x) - f(a)g(x) + f(a)g(x) - f(a)g(a)}{x-a} \\ &= \lim_{x\to a} \frac{f(x)-f(a)}{x-a} g(x) + f(a)\frac{g(x)-g(a)}{x-a} \\ &= f'(a)g(a) + f(a)g'(a).\end{aligned}$$

□

Example: $(x^3)' = (x \cdot x^2)' = x \cdot 2x + x^2 = 3x^2$.

Claim 49. *If g is differentiable at a and f is differentiable at $g(a)$ then*

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a).$$

Explanation. Shall not provide full proof, but explanation:

$$\frac{f(g(a)) - f(g(x))}{x-a} = \frac{f(g(a)) - f(g(x))}{g(x) - g(a)} \cdot \frac{g(a) - g(x)}{x-a}.$$

□

Example: $\sin(x^2)$ at 3. The derivative is $\cos(x^2) \cdot 2x$ and at 3 it is $6\cos(9)$.

Corollary 50. *If $g(a) \neq 0$ then*

$$(f/g)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}.$$

Proof.

$$\begin{aligned}\left(\frac{f}{g}\right)'(a) &= \left(f\cdot\frac{1}{g}\right)'(a)\\ &= f'(a)\left(\frac{1}{g}\right)(a) + f(a)\left(\frac{1}{g}\right)'(a)\\ &= \frac{f'(a)}{g(a)} - f(a)\frac{1}{g^2(a)}g'(a)\\ &= \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}.\end{aligned}$$

□

Here we used that for all $n \geq 1$, $(1/x^n)' = (x^{-n})' = nx^{-n-1}$ for all x. We shall prove below.

It also holds that $\ln'(x) = 1/x$ for $x > 0$. Explanation:

$$\frac{\ln(a+h) - \ln(a)}{h} = \ln\left(\frac{a+h}{a}\right)^h = \ln\left(1+\frac{h}{a}\right)^h \to_{h\to 0} \ln e^{1/a} = 1/a.$$

How to use ln for $1/x^n$? Let $b > 0$ and let $f(x) = x^b$ for $x > 0$. Write

$$g(x) := \ln(f(x)) = b\ln(x).$$

On one hand

$$g'(x) = b\frac{1}{x}$$

and on other hand

$$g'(x) = \frac{1}{f(x)}f'(x) = \frac{1}{x^b}f'(x).$$

So,

$$f'(x) = bx^{b-1}.$$

Last example:

Claim 51. *If f is differentiable at a and invertible at a neighborhood of a, and $g = f^{-1}$ is its inverse then*

$$g'(x) = \frac{1}{f'(g(x))}.$$

Recall that inverse is just flipping, which shows that angle of tangent is inverted.

Explanation. $(g\circ f)(x) = x$ so by composition rule $f'(g(x))g'(x) = 1$. (We did not prove that g is differentiable.) □

Example: $f(x) = \sin^{-1}(x)$. Thus $\sin(f(x)) = x$ and $\cos(f(x))f'(x) = 1$ or

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

5.2 Higher order

If we take one derivative, why not take more?

Definition 52. *If f has a derivative f' and f' is differentiable then $f^{(2)} = f''$. Similarly, define $f^{(n)}$ as the n'th derivative of f, if it exists.*

These also have rules: If f, g have n derivatives then

$$(f+g)^{(n)} = f^{(n)} + g^{(n)}$$

and

$$(fg)^{(2)} = (f^{(1)}g + fg^{(1)})^{(1)} = f^{(2)}g + f^{(1)}g^{(1)} + f^{(1)}g^{(1)} + fg^{(2)} = f^{(2)}g + 2f^{(1)}g^{(1)} + fg^{(2)}.$$

This is similar to the binomial theorem, and indeed

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

5.3 Summary

So far we provided the definition of derivatives and described the rules that derivatives follow. You can think of it as the instructions that we get when buying an electric appliance. What we can and can not do with it. We shall now start the deeper and perhaps more interesting part of this study. How to apply derivatives, according to the rules we discussed.

5.4 Finding extreme points

As we discussed, finding extreme points of function is a really important and useful task. Derivatives are very helpful in doing so.

Definition 53. *A point a is a local maximum of f if there is a neighborhood I of a so that for all $x \in I$ we have $f(x) \leq f(a)$. Can similarly define a local minimum. If a is either minimum or maximum then it is an extreme point (not necessarily global ones).*

Theorem 54 (Fermat). *If f is differntiable in an open interval I and $x \in I$ is an extreme point then $f'(a) = 0$.*

Proof. Consider the case that a is a local maximum. Then for every x that is close enough to a we have

$$f(x) \leq f(a).$$

So, if $x > a$ then $(f(x) - f(a))/(x - a) \leq 0$ and if $x < a$ then $(f(x) - f(a))/(x - a) \geq 0$. So the right limit is non positive and the left limit is non negative. But both are equal since the limit exists. So it must be equal to 0. □

Example: $f(x) = x^3 - x + 5$ in $(-10, 10)$. Then $f'(x) = 3x^2 + 1$. So the extreme points are $\pm\frac{1}{\sqrt{3}}$. Now $f(10)$ is large positive, and $f(-10)$ is large negative. And $f(-1/\sqrt{3}) > 0$. So it should be a local maximum. The other point is a local minimum. There is a root for f between -10 and $-1/\sqrt{3}$.

5.5 Rolle's theorem

This is a variant of the intermediate value theorem, with respect to derivatives. It says that if we draw the graph of a smooth function starting and ending at the same height then we there must be a point in which the tangent is parallel to axis. This is a global to local deduction.

Theorem 55 (Rolle). *If f is continuous in $[a, b]$ and differentiable in (a, b) and $f(a) = f(b)$ then there is $x \in (a, b)$ so that $f'(x) = 0$.*

Proof. We know that f gets a minimum and a maximum. If both are equal then (?) the function is constant and f' is zero. If not, then there is an extreme point in (a, b) and there the derivative is zero (Fermat). □

Application/exercise: If $c > 0$ then the number of solutions to

$$x^3 + cx + d = 0$$

is exactly 1. We saw that there is at least one (since degree is odd). Assume $a < b$ are two solutions (towards a contradiction). Then by Rolle's theorem since the function is differentiable everywhere there is $x \in (a, b)$ so that $3x^2 + c = 0$ but this is impossible.

5.6 Lagrange theorem

Rolle's theorem only applies to cases when $f(a) = f(b)$. What is the general phenomenon behind it?

Theorem 56 (Largrange)**.** *Let f be continuous in $[a, b]$ and differentiable in (a, b). Then, there is $x \in (a, b)$ so that*

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

Geometrically: there is a point in the interval in which the tangent is parallel to the line through $(a, f(a))$ and $(b, f(b))$.

One interpretation is related to speeding. If someone drive 100 km in 30 minutes, then although his speed was never measured, it is still certain that at some point in time his speed was 200 km/h.

In high level, the theorem translates global information ($f(a), f(b)$ for far away a, b) to local information (the derivative at x).

We shall not prove, but one can use Rolle's theorem to prove this one.

We can apply Lagrange theorem to learn global properties of f from local ones. Imagine, for example, a queue in which every person is higher than the person in front of her. Then e.g. the last person is higher than the first person.

Corollary 57. *Let f be continuous in $[a, b]$ and differentiable in (a, b).*

- *If $f'(x) = 0$ for all $x \in (a, b)$ then f is constant.*
- *If $f'(x) \geq 0$ for all $x \in (a, b)$ then f is increasing.*
- *If $f'(x) > 0$ for all $x \in (a, b)$ then f is strongly increasing.*
- *Similarly for ≤ 0 and < 0.*

Proof. Let us prove the third item, for example. Let $a' < b'$ be two points in (a, b). By Lagrange's theorem, there is $x \in (a', b')$ so that

$$f(b') - f(a') = f'(x)(b' - a').$$

Since $f'(x) > 0$ and $b' > a'$, we know $f(b') - f(a') > 0$. □

Comment: For items 1,2 the other direction holds as well. That is, if f is constant then $f' = 0$, and if f is increasing and differentiable then $f' \geq 0$. The other direction in item 3 does not necessarily hold: the function $x \mapsto x^3$ is strongly monotone but the derivative at 0 is 0.

5.7 Applications

Inequalities.

Claim 58. *For every $x \geq 0$ we have*

$$\ln(1+x) \geq \frac{x}{1+x}.$$

Proof. Denote

$$f(x) = \ln(1+x) - \frac{x}{1+x}.$$

Thus

$$f'(x) = \frac{1}{1+x} \cdot 1 - \frac{1+x-x}{(1+x)^2} = \frac{1+x-1}{(1+x)^2} = \frac{x}{(1+x)^2} \geq 0.$$

Lagrange theorem implies that f is increasing so $f(x) \geq f(0) = 0$. □

Qualitative behavior of functions. One important ability to have is to describe the qualitative behavior of a system: Is the market about to go to depression? Is the health of a person improving? ...

The general idea is simple. If we want to understand the regions in which f increases or decreases then we just need to understand the positivity or negativity of f'.

Before stating the general property we are going to use, let us consider a simple example that will help to clarify it. Consider the function $f(x) = x^3$. We know it is strongly increasing on $\mathbb{R}$. However, $f'(x) = 3x^2$ has a root at 0. So the roots of f' are not necessarily local maximum or minimum in f. How can we recognize this? The idea is to use another derivative. In this example, not only that $f'(0) = 0$ but also $f''(0) = 0$.

Claim 59. *Let f be differentiable twice in (a,b) and $x \in (a,b)$. If $f'(x) = 0$ and $f''(x) \neq 0$ then x is a local minimum or maximum.*

Explanation. Assume $f''(x) > 0$ (the other case is similar). Therefore, we know that if z is close to x we have

$$\frac{f'(z)}{z-x} = \frac{f'(z) - f'(x)}{z-x} > 0.$$

So f' is positive to the right of x and negative to the left of x. So f is increasing to the right of x and decreasing to the left of x. This means that x is a local minimum. □

Example: $f(x) = x^3 - 3x + 4$. Thus $f'(x) = 3x^2 - 3$ and $f''(x) = 6x$. The roots of f' are ± 1. And $f''(1) > 0$ and $f''(-1) < 0$. This means that 1 is a local minimum and -1 is a local maximum. Also $f(1) = 1 - 3 + 4 = 2$ and $f(-1) = -1 + 3 + 4 = 6$. We can finally provide a rough sketch of f.

Finding special points. Let $A = (-1, 0)$ in $\mathbb{R}^2$. Consider the line $y = x + 2$. Its graphs is set of points of form $(x, x+2)$. Find a point B on the line that is closest to A. For a point $(x, x+2)$ the distance from A is $\sqrt{(x+1)^2 + (x+2)^2}$. This is minimal iff the square of it is. That is, we want to find x the minimizes $f(x) = (x+1)^2 + (x+2)^2$. The derivative is $f'(x) = 2(x+1) + 2(x+2) = 4x + 6$ and so the root of f' is $-3/2$.

The point that minimized distance is $(-1.5, 0.5)$. Observe that $B - A = (-0.5, 0.5)$ is orthogonal to say $(1, 3) - (0, 2) = (1, 1)$ which makes sense.

5.8 L'hospital rule

In many cases we get an expression of the form $\lim_{x\to a} \frac{f(x)}{g(x)}$ where both f, g tend to 0 or ∞ at a. For example, $\lim_{x\to 0} \frac{1-\cos^3(x)}{\sin(x)}$. How do we deal with them?

Theorem 60 (L'hospital)**.** *Let f, g be two functions that are differentiable in a neighborhood of a. Assume that*

1. $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$.
2. $g'(x) \neq 0$ *in a (punctured) neighborhood of a.*
3. $\lim_{x\to a} \frac{f'(x)}{g'(x)} = L$.

Then,

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

How to use for example above? Well, $f(x) = 1-\cos^3(x)$ gives $f'(x) = 3\cos^2(x)\sin(x)$, and $g(x) = \sin(x)$ gives $g'(x) = \cos(x)$. Also,

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} 3\cos(x)\sin(x) = 0.$$

So, $\lim_{x\to 0} \frac{1-\cos^3(x)}{\sin(x)} = 0$.

Rough explanation.

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{\frac{f(x)-f(a)}{x-a}}{\frac{g(x)-g(a)}{x-a}} \to L.$$

□

Comments:
- Use only in situations like $0/0$ or ∞/∞. In other cases it does not work.
- You can also use for limits at ∞ or for one sided limits.
- If $g'(a) = f'(a) = 0$ then we use rule again and consider $f''(x)/g''(x)$.

- There are many ways to apply rule. For example, also for "$\infty - \infty$".

$$\begin{aligned}\lim_{x\to 0} \frac{1}{\sin(x)} - \frac{1}{x} &= \lim_{x\to 0} \frac{x-\sin(x)}{x\sin(x)} \\ &= \lim_{x\to 0} \frac{1-\cos(x)}{\sin(x)+x\cos(x)} \\ &= \lim_{x\to 0} \frac{\sin(x)}{\cos(x)+\cos(x)-x\sin(x)} = \frac{0}{2} = 0.\end{aligned}$$

Orders of magnitude. In general for x large we have

$$\cdots \ll \sqrt{\log(x)} \ll \log(x) \ll x \ll x^2 \ll \cdots \ll 2^x \ll e^x \ll \cdots .$$

Let us see some examples:

$$\lim_{x\to\infty} \frac{\log(x)}{x} = \lim_{x\to\infty} \frac{1/x}{1} = 0.$$

and

$$\lim_{x\to\infty} \frac{x^3}{e^x} = \lim_{x\to\infty} \frac{3x^2}{e^x} = \lim_{x\to\infty} \frac{6}{e^x} = 0.$$

Generally, if $\frac{f(x)}{g(x)} \to 0$ when $x \to \infty$ then we say that f is greater order of magnitude than g.

5.9 Convex functions

Convex functions form an important family of functions. There are several equivalent definitions of convexity, each provides a different perspective.

Definition 61 (Convexity-geometric)**.** *The function f is convex in an interval I if for every $a < b$ in I the line segment between $(a, f(a))$ and $(b, f(b))$ is above the graph of the function.*

Definition 62 (Convexity-analytic)**.** *The function f is convex in an interval I if for every $a < b$ in I and for every $p \in [0, 1]$ we have*

$$f(pa + (1-p)b) \le pf(a) + (1-p)f(b).$$

Discuss the connection between the two definitions by drawing an example.

If there is a strict inequality or the line does not touch the graph (except at endpoints) then we say that f is strongly convex.

Claim 63. *If f is convex in an open interval I then f is continuous in I.*

Shall not prove. However, there are convex functions that are not differentiable. For example, $|x|$.

Convexity is a global property – it is related to far away points. We can also give local definitions.

Claim 64 (Equivalent definitions)**.** *Assume f is convex in an open interval I.*

1. *If f is differentiable in I then f is convex iff for every $x \in I$ the tangent to f at x is under the graph of the function.*

2. *If f is differentiable in I then f is convex iff the function f' is increasing in I.*

3. *If f is twice differentiable in I then f is convex iff for all $x \in I$ we have $f''(x) \geq 0$.*

We shall not prove.

The last item is easiest to verify, as long as f is twice differentiable. We can now see that x, x^2, e^x are convex on $\mathbb{R}$. The function x^3 is not convex on $\mathbb{R}$ but is on $(0, \infty)$.

An important convex function is $-\ln(x) = \ln(1/x)$ for $x > 0$. Indeed, the second derivative is $1/x^2 > 0$. Application: the arithmetic mean geometric mean inequality.

Claim 65. *For all $a, b > 0$, we have $\sqrt{ab} \leq (a+b)/2$.*

Proof. The function $\ln(1/x)$ is decreasing, so we need to show

$$(1/2)(\ln(1/a) + \ln(1/b)) = \ln(1/\sqrt{ab}) \geq \ln(1/((a/2) + (b/2))).$$

This is exactly convexity of this function (2nd derivative is $1/x^2 > 0$). □

Using induction, we can generalize the definition of convexity, and prove for example the general arithmetic mean geometric mean inequality.

Claim 66. *If f is convex in an open interval I and $x_1, x_2, \ldots, x_n \in I$ and $p_1, p_2, \ldots, p_n > 0$ so that $\sum_{i=1}^n p_i = 1$ then*

$$f\left(\sum_{i=1}^n p_i x_i\right) \leq \sum_{i=1}^n p_i f(x_i).$$

This can be proved by induction, but we shall not.

5.10 Graphs

It is very helpful to draw, and graphical representation may dramatically help in understanding complex systems. One example is Chernoff Faces in which high dimensional data is represented as a drawing of some face, and the expression of the face can be related to the state of the system.

Let us consider a few examples.

$\mathbf{f(x) = x^3}$. The derivative is 0 only at 0 and is always ≥ 0. So the function is increasing. To the right of 0 we have $f'' \geq 0$ which mean that f is convex. To the left of 0 we have $f'' \leq 0$ which means that f is concave ($-f$ is convex). The point 0 is suspicious, since $f'(0) = 0$. But it is not a maximum or a minimum. It is a transition point between convexity and concavity. Such a point is called an inflection point ("pitul").

Definition 67. *A point a is an inflection point of f if f is convex to one side of a and concave in the other.*

Comments:

- This is not equivalent to $f''(a) = 0$. For example, x^4 at 0.

- Sometimes a is an inflection point even though f is not defined at a. For example, $1/x$ at 0.

$\mathbf{f(x) = \frac{x}{1+x^2}}$. Let us go over the list of things to do. It is helpful to have a drawing the you update as you go along.

1. Domain. x and $1 + x^2$ are defined over all $\mathbb{R}$ and $1 + x^2$ is always non zero. So the domain of f is $\mathbb{R}$.

2. Continuity. Again, it is a ration of 2 continuos function so it is continuous.

3. Even/odd. It is odd $f(-x) = \frac{-x}{1+(-x)^2} = -f(x)$.

4. Roots. The only root is 0.

5. Derivatives, extreme points, monotonicity...

$$f'(x) = \frac{1 + x^2 - x \cdot 2x}{(1 + x^2)^2} = \frac{1 - x^2}{(1 + x^2)^2}.$$

It is zero at ± 1. It is positive in $(-1, 1)$. It is negative in $(-\infty, -1) \cup (1, \infty)$. So: It is decreasing to left of -1. It has a local minimum at -1. It is increasing between $-1, 1$. It has a local maximum at 1 and it is increasing after 1.

Since the only root of f is 0, we learn that -1 is a global minimum and 1 is a global maximum.

6. Behavior at $\pm\infty$. The limit at ∞ and $-\infty$ of f is 0.

The values $f(1) = 1/2$ and $f(-1) = -1/2$. So f attains a global maximum at 1 and a global minimum at -1.

7. Convexity, inflection points, 2nd derivatives,... Since f is convex at -1 and concave at 1, there is an inflection point between them. Similarly, f is convex x much larger than 1 so there is an inflection point to right of 1, and also to left of -1 due to symmetry.

To find inflection points, need to understand the monotonicity of f'. One way is to compute

$$\begin{aligned} f''(x) &= \frac{-2x(1+x^2)^2 - 2(1+x^2)2x(1-x^2)}{(1+x^2)^4} \\ &= \frac{-2x(1+x^2)((1+x^2)+2(1-x^2))}{(1+x^2)^4} \\ &= \frac{-2x(1+x^2)(-x^2+3)}{(1+x^2)^4}. \end{aligned}$$

So the roots of f'' are $0, \pm\sqrt{3}$. These are the 3 inflection points of f, as we foresaw.

8. Asymptotes *(related to behavior at* $\pm\infty$*).* In the example from above,

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

and also at $-\infty$. So the asymptotic behavior at $\pm\infty$ is clear.

Sometimes the behavior is more complex. The line $ax+b$ is an asymptote of f at ∞ if f gets closer and closer to it. Draw.

Definition 68. *The line $ax+b$ is an asymptote of f at ∞ if*

$$\lim_{x\to\infty} f(x) - ax - b = 0.$$

Example: The function $\frac{2x^2+\sin(x)}{x-3}$ has an asymptote $2x$ at infinity.

A function may have a right vertical asymptote at a point a if the generalized limit $\lim_{x\to a^+} f(x) = \infty$. The function above has a right vertical asymptote at 3 (and also a left one).

How to compute a, b? Well,

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

and after finding a,

$$b = \lim_{x\to\infty} f(x) - ax.$$

In the example,

$$a = \lim_{x\to\infty} \frac{2x^2+\sin(x)}{x(x-3)} = 2,$$

and

$$b = \lim_{x\to\infty} \frac{2x^2+\sin(x)}{x-3} - 2x = \lim_{x\to\infty} \frac{2x^2+\sin(x)-2x(x-3)}{x-3} = 0.$$

To conclude, when studying a function in this course, there is a list of 7-8 things to do or check.

5.11 Approximations

We have used derivatives in several ways. As a tool for finding extreme points, and understanding a given function better in general. As a tool for proving that a function is convex. We now use it to approximate quantities of interest. This is of great importance, and in some way this is what calculators do.

Assume that we want to approximate a function f around a point a. For example, the function $\sin(x)$ around 0; what is $\sin(0.1)$? If the function is continuous, then we know that $f(x) \approx f(a)$ if x is close to a. If the function is differentiable, we can get a better approximation. We know that if x is close to a then

$$f'(a) \approx \frac{f(x) - f(a)}{x - a}.$$

In other words, we can approximate f in the vicinity of a by the tangent to the graph of f at a:

$$f(x) \approx f(a) + f'(a)(x - a).$$

Why is this a better approximation? Well,

$$|f(x) - f(a)| \approx |f'(a)| \cdot |x - a|.$$

But

$$|f(x) - f(a) - f'(a)(x - a)| = |x - a| \cdot \left| \frac{f(x) - f(a)}{x - a} - f'(a) \right| \ll |x - a|,$$

which is a much better approximation. For example, the first approximation gives $\sin(0.1) \approx 0$ and the second gives $\sin(0.1) \approx 0 + \cos(0) \cdot 0.1 = 0.1$. A more accurate answer using a calculator is $\sin(0.1) \approx 0.09983$.

Theorem 69 (Taylor)**.** *A function f is differentiable at a iff there are functions G, E so that*

$$f(x) = f(a) + G(a)(x - a) + E(x),$$

where

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

In addition, $f'(a) = G(a)$.

We shall not prove theorem, but it shows that approximating $f(x)$ as $f(a)+f'(a)(x-a)$ is a good approximation if x is close to a. Think of $E(x)$ as the error of the estimate.

If f is continuous, we get a good approximation to $f(x)$. If it is differentiable, we get a better approximation. Having more and more data on f, we get better and better approximations.

Definition 70. *Let f be a function that is differentiable n times at a. The degree n Taylor polynomial for f at a is defined as*

$$\begin{aligned} T_n(x) &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k. \end{aligned}$$

This polynomial is a good approximation for f. The 0th approximation is $f(a)$. The 1st approximation is $f(a)+f'(a)(x-a)$. The 2nd is $f(a)+f'(a)(x-a)+f''(a)(x-a)^2/2$. And so forth.

Claim 71. *If f is differentiable n times at a then for all $1 \le k \le n$*

$$T^{(k)}(a) = f^{(k)}(a).$$

Proof. The kth derivative of T_n removes all powers less than k. The kth power is multiplied by $k!$. The powers that are larger than k are zeroed since we substitute a. Overall,

$$T^{(k)}(a) = k! \cdot \frac{f^{(k)}(a)}{k!}.$$

□

The Taylor polynomial is usually a good approximating to f, and as the degree grows, the approximation becomes better and better.

Examples.

$\mathbf{f(x) = e^x}$ **at** 0**.** $f^{(k)}(0) = e^x|_{x=0} = 1$. So,

$$T_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}.$$

$\mathbf{f(x) = \sin(x)}$ **at** 0**.** $f^{(1)}(x) = \cos(x)$, $f^{(2)}(x) = -\sin(x)$, $f^{(3)}(x) = -\cos(x)$, $f^{(4)}(x) = \sin(x)$, and so forth. At 0, all the even derivates are 0. And,

$$T_n(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots.$$

So we can approximate $f(x)$ in this way. But how can we know the size of the error? Let

$$E_n(x) = f(x) - T_n(x).$$

It is the error that T_n yields.

Theorem 72 (Taylor)**.** *Let f be differentiable $n+1$ times in a neighborhood I of a. For every $x \in I$, there is c between a and x so that*

$$E_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Comments:

- It is an extension of Lagrange's theorem.

- We do not know what c is, just that it is between a and x. This is still useful: We can calculate $e^{1/2}$ up to error $1/1000$. Let $f = e^x$. We want $f(1/2)$. We saw

$$T_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}.$$

By the theorem, there is $0 < c < 1$ so that

$$E_n(1) = \frac{f^{(n+1)}(c)}{(n+1)!}(1/2)^{n+1} = \frac{e^c}{(n+1)!}(1/2)^{n+1}.$$

So, if $n = 5$ since $e < 3$

$$|E_n(1)| < \frac{3}{5!}(1/2)^5 < 1/1000,$$

so we approximate

$$e \approx 1 + \frac{(1/2)}{1} + \frac{(1/2)^2}{2} + \frac{(1/2)^3}{6} + \frac{(1/2)^4}{24} = 1.6484375,$$

and the error is less than $1/1000$ (calculator says 1.6487212707).

This is a numerical approximation. Instead of computing the value exactly (which we can not for example with e) we approximate it well enough, using the theory we built so far. This is what calculators basically do. Given an error parameter $\epsilon > 0$, we need to find n so that $|E_n(x)| < \epsilon$ and then we estimate $f(x)$ by $T_n(x)$. You can now estimate $e^{1/2}$, $\sqrt{1.1}$,...

- This procedure does not always work. For example, the function

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

This is a continuous function, but all the derivatives at 0 are 0, so $T_n(x) = 0$ for all n.

We can still describe general conditions for which this procedure actually works.

Theorem 73. *Assume that f is differentiable n times for all natural n. Let $a \in \mathbb{R}$. Assume that there is a constant $C > 0$ so that for all $x \in \mathbb{R}$ and n,*

$$|f^{(n)}(x)| < C^n.$$

Then,

$$\lim_{n \to \infty} |E_n(x)| = 0.$$

Proof. By Taylor's theorem there is $c \in \mathbb{R}$ so that

$$|E_n(x)| = \left| \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} \right| \leq \frac{C^n |x-a|^n}{(n+1)!} \to 0,$$

when $n \to \infty$. □

Conclusion: $\sum_{k=0}^{n} \frac{1}{k!} \to e$, when $n \to \infty$.

We thus concluded our discussion of derivatives.

Chapter 6

Integrals

6.1 Indefinite integrals

We start with a more syntactic definition of integral and later on we give it semantics/meaning.

Definition 74. *An indefinite integral of f is a function F so that $F' = f$.*

Example: $f(x) = 2x$ then x^2 and $x^2 + 1$ are both indefinite integrals of f. In general, $F(x) = x^2 + c$ for all c. So, F is not unique.

Notation: F is denoted $\int f(x)dx$. This notation will be explained later on.

More examples:

- $\int e^x dx = e^x + c$
- $\int x^n dx = \frac{1}{n+1} x^{n+1} + c$ if $n \neq -1$
- $\int 1/x dx = \ln|x| + c$. Check: $\frac{d}{dx} \ln(x) = 1/x$ for $x > 0$ and $\frac{d}{dx} \ln(-x) = (1/(-x)) \cdot (-1) = 1/x$ for $x < 0$.
- $\int \frac{1}{1+x^2} dx = \tan^{-1}(x) + c$.
- All the examples we already saw with derivatives (just in reverse);

$$\int f'(x)dx = f(x) + c.$$

Properties of derivatives imply properties of integration.

Linearity of derivative implies

Theorem 75 (Linearity)**.** *If f, g have integrals then for all $a, b \in \mathbb{R}$*

$$\int af(x) + bg(x)dx = a \int f(x)dx + b \int g(x)dx.$$

For example,

$$\int x + e^x dx = x^2/2 + e^x + c.$$

Derivative of product implies: If u, v are differentiable function

$$(u \cdot v)' = u' \cdot v + u \cdot v'.$$

Integrate by side to get

$$u(x) \cdot v(x) = \int u'(x)v(x)dx + \int u(x)v'(x)dx$$

or

Theorem 76 (Integration by parts)**.** *If u, v are integrable then*

$$\int u'(x)v(x)dx = u(x) \cdot v(x) - \int u(x)v'(x)dx.$$

How to use? This is useful if we want to integrate $f(x) \cdot g(x)$ and we know $\int f(x)dx := u(x)$. Because then

$$\int f(x)g(x)dx = u(x)g(x) - \int u(x)g'(x)dx.$$

Examples:

- $\int x \ln(x)dx$: Since $\int xdx = x^2/2 + c$,

$$\int x \ln(x)dx = (x^2/2)\ln(x) - \int (x^2/2)(1/x)dx = x^2 \ln(x)/2 - x^2/4 + c.$$

- $\int \ln(x)dx$:

$$\int 1 \cdot \ln(x)dx = x\ln(x) - \int x(1/x)dx = x\ln(x) - x + c.$$

- There are many more, you will see in exercises.

Derivative of composition implies: assume $F' = f$ then

$$(F(g(x)))' = f(g(x)) \cdot g'(x).$$

Theorem 77 (Substitution)**.** *Assume f, g are integrable and $F = \int f(x)dx$ then*

$$\int f(g(x))g'(x)dx = F(g(x)) + c.$$

Example: $\int e^{x^2} 2x dx =?$ Set $f(x) = e^x$ and $g(x) = x^2$, and get $f(g(x)) = e^{x^2}$ and $\int f(x)dx = e^x$, and

$$\int e^{x^2} 2x dx = e^{x^2} + c.$$

How do we use? We will not provide full proof here, but here is the method. Say we want to compute some $\int h(x)dx$. And we want to change variable to $y = y(x)$. (Above we chose $y = x^2$.) Then we write $\frac{dy}{dx} = y'$, and so $dy = y'dx$. (We have not formally justified.) Thus,

$$\int h(x)dx = \int h(y)\frac{1}{y'}dy.$$

In the example above:

$$\int e^{x^2} 2x dx =_{y=x^2} \int e^y 2x \frac{1}{2x} dy = e^y + c = e^{x^2} + c.$$

Another example:

$$\int \frac{f'(x)}{f(x)} dx = \ln |f(x)| + c.$$

In many cases all three methods are used.

In general, there is no way to compute an integral of an elementary function. The integral $\int e^{x^2} dx$ for example is not elementary. It is not so easy to prove this.

Rational functions. There are other general methods that we shall talk about. We will just give one example.

$$\int \frac{-2x+4}{(x^2+1)(x-1)^2} dx =?$$

The idea is to decompose the rational function to parts. What are A, B, C, D so that

$$\frac{-2x+4}{(x^2+1)(x-1)^2} = \frac{Ax+B}{x^2+1} + \frac{C}{x-1} + \frac{D}{(x-1)^2}?$$

This is the first step in general. The rule is that the powers in numerator are smaller than power in denominator. To find A, B, C, D calculate

$$\begin{aligned} &\frac{Ax+B}{x^2+1} + \frac{C}{x-1} + \frac{D}{(x-1)^2} \\ &= \frac{(Ax+B)(x^2-2x+1) + C(x^2+1)(x-1) + D(x^2+1)}{(x^2+1)(x-1)^2} \\ &= \frac{x^3(A+C) + x^2(-2A+B-C+D) + x(A-2B+C) + B-C+D}{(x^2+1)(x-1)^2}. \end{aligned}$$

So, $A + C = 0$ or

$$A = -C.$$

And $-2 = A + C - 2B$ so

$$B = 1.$$

And

$$0 = -2A + B - C + D = -A + 1 + D$$

and

$$4 = B - C + D = 1 + A + D.$$

Sum the two: $4 = 2 + 2D$ and

$$D = 1, A = 2, C = -2.$$

Finally,

$$\begin{aligned}
&\int \frac{-2x+4}{(x^2+1)(x-1)^2}dx \\
&= \int \frac{2x}{x^2+1} + \frac{1}{x^2+1} - \frac{2}{x-1} + \frac{1}{(x-1)^2}dx \\
&= \ln|x^2+1| + \tan^{-1}(x) - 2\ln|x-1| - \frac{1}{x-1} + c.
\end{aligned}$$

Square roots. Sometimes when there are square roots it is helpful to use trigonometric functions. For example.

$$\begin{aligned}
\int \sqrt{1-x^2}dx &=_{x=\sin y, dx=\cos y dy} \int \cos^2(y)dy =_{\cos(2y)=2\cos^2(y)-1} \frac{1}{2}\int \cos(2y)+1dy \\
&= \frac{1}{4}\sin(2y) + \frac{y}{2} \\
&= \frac{1}{4}\sin(2\sin^{-1}(x)) + \frac{\sin^{-1}(x)}{2} \\
&=_{\sin(2t)=2\sin(t)\cos(t)} \frac{1}{2}x\sqrt{1-x^2} + \frac{\sin^{-1}(x)}{2}.
\end{aligned}$$

There are more examples...

6.2 Definite integrals

We have discussed integrals on a syntactic level. We shall now provide semantics/meaning to them.

Here the issue we will address is how to compute the area or volume of a given shape. This is a useful ability to have, for example, you may wonder how much water does a given bottle contain, or what is the area of a circle?

The idea is very simple. First, we learn how to measure the area of simple forms, like rectangles and triangles. Then, we approximate every shape by small enough simple shapes.

We focus on the area between the x-axis and a graph of a function f in the interval $[a, b]$. This area is denoted by

$$\int_a^b f(x)dx.$$

Area above the x-axis is positive and below is negative. Here and below, it is helpful to draw examples.

One step functions. The simplest functions to consider are step functions. Let $f(x) = c$ for $x \in [a, b]$ and $f(x) = 0$ otherwise. The area $\int_a^b f(x)dx$ is the area of a rectangle, so it is $c(b - a)$.

Steps functions. Now consider a collection of step functions. Let $a_0 < a_1 < a_2 < \ldots < a_n$. Let f be a function that equals c_i in $[a_{i-1}, a_i]$ for $i \in \{1, 2, \ldots, n\}$. Draw it. The area $\int_{a_0}^{a_n} f(x)dx$ is the area of several rectangles and equals $\sum_{i=1}^n c_i(a_i - a_{i-1})$.

General functions. Let f be a general function. We shall approximate f by a steps function.

Partition $[a, b]$ to many small parts:

$$a = a_0 < a_1 < \ldots < a_n = b.$$

When is such a partition good? If all parts are small. Define the parameter of this choice of (a_i) to be

$$\max\{a_i - a_{i-1} : i \in \{1, 2, \ldots, n\}\}.$$

The smaller this parameter is, the more accurate the partition is.

Between a_{i-1} and a_i the function f is not constant. We need to choose a point in $[a_{i-1}, a_i]$ to evaluate f in. Denote this point by t_i.

Given a partition of $[a, b]$ and a choice of evaluation points, the area under f should be close to

$$\sum_{i=1}^n f(t_i)(a_i - a_{i-1}).$$

Draw it. This is called a Riemann sum.

Let us see a simple example. Consider $f(x) = x$ in $[0, 1]$. What is the area? 1/2. Let $a_i = i/n$ for large n define a partition of $[0, 1]$. Choose $t_i = i/n$ as well, for simplicity.

The Riemann sum is

$$\sum_{i=1}^{n} \frac{i}{n}\frac{1}{n} = \frac{1}{n^2}\frac{n(n-1)}{2} \to 1/2,$$

when $n \to \infty$. The larger n is, the smaller the partition is, the better the approximation to area is.

Of special interest to us are functions for which this method works. This roughly means that as long as the parameter of the partition is small, we get a good approximation of area.

Definition 78. *The function f is integrable in $[a,b]$ if there is $I \in \mathbb{R}$ so that for every $\epsilon > 0$ there is $\delta > 0$ so that for every $a = a_0 < a_1 < \ldots < a_n = b$ with parameter at most $\delta$and for every choice of $t_i \in [a_{i-1}, a_i]$,*

$$\left| I - \sum_{i=1}^{n} f(t_i)(a_i - a_{i-1}) \right| < \epsilon.$$

We denote

$$\int_a^b f(x)dx = I$$

and call it the integral of f between a, b.

The notation $\int$ in this context is clearer than the "antiderivative" from before. Later we shall explain connection. It is obtained from S which corresponds to "sum" by elongating it, and the "dx" indicates what we are summing over.

We have defined the notion of integrability. It essentially says that we do not need to make very clever choices in order to approximate the area under f using the method above. But are there integrable functions?

Theorem 79. *Let f be defined and bounded on $[a,b]$. If f is either continuous or monotone then f is integrable.*

We shall not prove. Note that if $[a,b]$ can be partition to finitely many areas in which f is integrable then f is also integrable.

There are functions that are not integrable, but they are relatively rare and we shall not see an example.

Properties:

1. Changing the value of f on finitely many points does not change its integral.

2. Linearity: if f, g are integrable on $[a,b]$ then

$$\int_a^b f(x) + g(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx$$

and for every constant $c \in \mathbb{R}$

$$\int_a^b cf(x)dx = c\int_a^b f(x)dx.$$

3. Respect order: If $f(x) \leq g(x)$ for all $x \in [a,b]$ and f, g are integrable then

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

Draw. Specifically, if $m \leq f(x) \leq M$ for all $x \in [a,b]$ then

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a).$$

Note: if $f(x) \geq 0$ and is strictly positive at a single point then the integral can still be zero. However, if f is continuous then it can not. If $f(x_0) > 0$ then there is $\delta > 0$ so that $f(x) > f(x_0)/2$ for all x so that $|x - x_0| < \delta$ so

$$\int_a^b f(x)dx > 2\delta f(x_0)/2 > 0;$$

draw.

4. Triangle inequality: If f is integrable in $[a,b]$ then $|f|$ is also integrable and

$$\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx.$$

5. Partition: If f is integrable in $[a,b]$ and $c \in [a,b]$ then

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

We further denote

$$\int_b^a f(x)dx = -\int_a^b f(x)dx.$$

Then, for every $c_1, c_2, c_3 \in [a,b]$,

$$\int_{c_1}^{c_2} f(x)dx = \int_{c_1}^{c_3} f(x)dx + \int_{c_3}^{c_2} f(x)dx.$$

6. Continuity: If f is integrable in $[a,b]$ and $c \in [a,b]$ then the function $F(t) =$

$\int_a^t f(x)dx$ for $t \in [a,b]$ is continuous. Intuitively, if we change t slightly, the area changes slightly.

Proof for bounded functions: Let M be so that $|f(x)| \leq M$ for all $x \in [a,b]$. Let $\epsilon > 0$. Write

$$|F(t+\delta) - F(t)| = \left| \int_t^{t+\delta} f(x)dx \right| \leq M\delta < \epsilon$$

if $\delta < M/\epsilon$.

6.3 The fundamental theorem of calculus

We now explain the connection between indefinite and definite integrals.

Theorem 80. *Assume f is continuous in $[a,b]$ and that $F(t) = \int_a^t f(x)dx$. Then, F is differentiable and $F'(x) = f(x)$ for all $x \in (a,b)$.*

Proof. Let $x \in (a,b)$. Let us start with limit from right. What is

$$\lim_{z \to x^+} \frac{F(z) - F(x)}{z - x} = \lim_{z \to x^+} \frac{\int_x^z f(y)dy}{z - x}?$$

Use sandwich. For every $z \geq x$, denote

$$M_z = \sup\{f(y) : x \leq y \leq z\}, \; m_z = \inf\{f(y) : x \leq y \leq z\}.$$

The continuity of f implies that

$$\lim_{z \to x^+} M_z = \lim_{z \to x^+} m_z = f(x).$$

Now,

$$m_z \leq \frac{F(z) - F(x)}{z - x} \leq M_z,$$

so

$$\lim_{z \to x^+} m_z = \lim_{z \to x^+} \frac{\int_x^z m_z dy}{z - x} \leq \lim_{z \to x^+} \frac{\int_x^z f(y)dy}{z - x} \lim_{z \to x^+} \frac{\int_x^z M_z dy}{z - x} = \lim_{z \to x^+} M_z,$$

and indeed the limit from right is correct. A similar argument works for limit from left, and completes the proof. □

Comments:

- Continuity of f is essential. For example, the integral of the function that is -1 on $[-1, 0)$ and 1 on $[0, 1]$ has a "corner" at 0.

- This shows that if f is continuous then $\int_a^t f(x)dx$ is an indefinite integral of it.

This connection allows to compute areas.

Theorem 81 (Newton-Leibnitz). *If f is continuous in $[a, b]$ and G is so that $G' = f$ then*

$$\int_a^b f(x)dx = G(b) - G(a).$$

Proof. By fundamental theorem, the derivative of $F(t) = \int_a^t f(x)dx$ is f. So, $(F-G)' = 0$ which implies $F = G + c$. So,

$$0 = F(a) = G(a) + c$$

and

$$\int_a^b f(x)dx = F(b) = G(b) + c = G(b) - G(a).$$

□

This is very useful. It gives us the ability to compute areas (so far we had no general way to do so, except than computing the limit). For example, we can calculate area of circle, which is 1/4 of

$$\int_0^1 \sqrt{1-x^2}dx = \frac{1}{2}x\sqrt{1-x^2} + \frac{\sin^{-1}(x)}{2}\Big|_0^1 = \pi/2.$$

Comments:

- Another intuitive perspective (not formal). Assume $F' = f$ and we want to compute the area under f in $[a, b]$. Draw. Choose a partition and evaluation points:

$$\sum_i f(t_i)(a_i - a_{i-1}) \approx \sum_i F(a_i) - F(a_{i-1}) = F(b) - F(a),$$

since it is a telescopic sum.

- Composition: If $H(t) = \int_{a(t)}^{b(t)} f(x)dx$ with a, b differentiable and $F' = f$ then

$$H(t) = F(b(t)) - F(a(t))$$

and by chain rule

$$F'(t) = f(b(t))b'(t) - f(a(t))a'(t).$$

Taking such derivatives is actually easy, for example,

$$\frac{d}{dt}\int_{\cos(t)}^{t^2} e^{x^2}dx = e^{t^4}2t + e^{\cos^2(t)}\sin(t).$$

- Newton-Leibnitz also holds when f is integrable and $G' = f$ except finitely many points.

6.4 Applications

Computing the area between two graphs. Say we want to compute the (positive) area between the graphs of x and x^2 between 0 and 2. Draw.

$$\text{area} = \int_0^1 x - x^2 dx + \int_1^2 x^2 - x dx = \frac{1}{2} - \frac{1}{3} + \frac{8}{3} - \frac{1}{3} - \frac{4}{2} + \frac{1}{2} = 2.$$

Position and speed. Imagine a particle moving on the real line (one dimensional for simplicity). If $x(t)$ denotes the position on the line of a particle at time t, then $v(t) = x'(t)$ is its speed at time t and $x(t) = \int_0^t v(s)ds$.

This gives further motivation for considering "negative area" since a negative position is natural.

Energy. If we want to move a suitcase from height 0 to height h, we need to invest energy, use a force that cancels gravity. The energy or work we need to invest is $\int_0^h G dx = Gh$, where G is the gravity force (this is part of the theory of physics). This is called the potential energy.

Mass of a wire. Consider a wire between $[0, 1]$ with non uniform mass. Denote by $m(x)$ the mass of the part of the wire between 0 and x. Thus, $\rho(x) = m'(x)$ is the mass density at x and $m(x) = \int_0^x \rho(y)dy$.

Length of a line. Say we want to compute the length of a graph of a function f (not the area under it). We need it to be nice enough, so we assume f is differentiable. How can we do it? Let us start with an approximation, as we did for area. Let $a_0 < a_1 <$

$\ldots < a_n$ be a partition of $[a, b]$.

$$\begin{aligned} length &\approx \sum_{i=1}^{n} \text{distance from } (a_{i-1}, f(a_{i-1})) \text{ to } (a_i, f(a_i)) \\ &= \sum_{i=1}^{n} \sqrt{(a_i - a_{i-1})^2 + (f(a_i) - f(a_{i-1}))^2} \\ &= \sum_{i=1}^{n} (a_i - a_{i-1}) \sqrt{1 + \left(\frac{f(a_i) - f(a_{i-1})}{a_i - a_{i-1}} \right)^2}. \end{aligned}$$

What does this converge to (if it does)?

$$\int_a^b \sqrt{1 + (f'(x))^2} dx.$$

We shall not prove. Try to use this formula to calculate the circumference of a unit circle (twice the length of the function $\sqrt{1 - x^2}$ from -1 to 1).

Approximations. As we mentioned, some functions do not have elementary integrals, and for some the integrals are just very hard to compute. In many cases, we can use the approximation methods we learned.

Assume we want to compute $\int_a^b f(x)dx$, but it is too complicated. We know we can approximate f be a polynomial (Taylor a.k.a. MacLaurin) T_n, so we can approximate the integral by $\int_a^b T_n(x)dx$.

6.5 Generalized integrals

So far we have considered integrals on intervals. What about a ray or all of $\mathbb{R}$? The idea is to use a limit to get from interval to a ray: If f is defined in $[a, \infty)$, we say that f is integrable over $[a, \infty)$ if the limit

$$\lim_{b \to \infty} \int_a^b f(x)dx$$

exists, and then we denote it by

$$\int_a^\infty f(x)dx.$$

Example:

$$\int_1^\infty \frac{1}{x^2} dx = \lim_{b \to \infty} \frac{-1}{b} + 1 = 1.$$

This is an infinite body with a finite volume.

Non example: $\int_0^\infty \cos(x)dx$ does not exist. Also $\int_1^\infty 1/x$ does not exist (the area in infinite).

Another problematic case is e.g. when a is a vertical asymptote of f. In this case, we use a limit too. Assume f is defined in $(a, b]$. We say that f is integrable in $[a, b]$ if the limit

$$\lim_{c\to a^+} \int_c^b f(x)dx$$

exists, and then we denote this limit by $\int_a^b f(x)dx$. For example,

$$\int_0^1 (1/\sqrt{x})dx = \lim_{c\to 0^+} 2\sqrt{1} - 2\sqrt{c} = 2.$$

The integral $\int_0^1 (1/x)dx$ however does not converge.

Comment: the two example above show that for every $r \in \mathbb{R}$ the integral $\int_0^\infty x^r dx$ does not converge, since either on $[0, 1]$ or on $[1, \infty]$ it does not. Note that to compute for example $\int_0^\infty 1/x^2$ we actually need to take 2 limits, not just 1 (a limit to 0 and a limit to infinity).

Finally, the integral $\int_{-\infty}^\infty f(x)dx$ is defined if the two integrals of f on $(-\infty, 0]$ and $[0, \infty)$ converge (we can choose any other point instead of 0). These are two different limits: It is not equivalent to that $\lim_{R\to\infty} \int_{-R}^R f(x)dx$ exists, e.g. for $f(x) = 1/\sqrt{x}$ for every R the integral is 0 but the entire integral is not defined.

The generalized integral has the same properties we saw for usual integrals.

6.5.1 Comparison

Sometimes it is difficult to see if $\int_0^\infty f(x)dx$ converges. Then, we can bound $0 \le f(x) \le g(x)$ for all x, and prove that the integral of g converges.

Theorem 82. *Assume f, g are defined in $[a, \infty)$ and that $0 \le f(x) \le g(x)$ for all $x \ge a$. Then*

$$\int_a^\infty f(x)dx \le \int_a^\infty g(x)dx.$$

Specifically, if the integral of f diverges to ∞ then so does that of g, and if the integral of g converges then so does that of f.

Proof. Denote $F(t) = \int_a^t f(x)dx$ and $G(t) = \int_a^t g(x)dx$. These two functions are monotone non decreasing, and $F(t) \le G(t)$ for all t. Thus, the same holds for generalized limits. □

It suffices that $f(x) \le g(x)$ for large enough x.

This is useful to prove convergence or divergence. For example, the integral $\int_0^\infty e^{-x^2}dx$ is defined (although we do not have a nice formula for it) since $0 < e^{-x^2} < e^{-x}$ for $x > 1$, and we know how to compute the integral of e^{-x}.

6.5.2 Absolute convergence

When computing say $\int_{-\infty}^\infty f(x)dx$ we would like to have a condition that guarantees that the integral is "well behaved." (The importance of absolute convergence will become clearer later on.)

Definition 83. *The integral of f on $[a,\infty)$ absolutely converges if $\int_a^\infty |f(x)|dx$ converges.*

This is a stronger condition than convergence. If an integral absolutely converges then it converges—exercise.

The other direction does not necessarily hold. Let us see an example. The following integral converges

$$\int_1^\infty \frac{\sin(x)}{x}dx =_{\text{parts}} -\frac{\cos(x)}{x}\Big|_1^\infty - \int_1^\infty \frac{\cos(x)}{x^2}dx,$$

and the two terms are finite; the second since $|\cos(x)| < 1$ and since the integral of $1/x^2$ converges—using comparison. But it does not absolutely converge:

$$\int_1^\infty \frac{|\sin(x)|}{x}dx \geq \int_1^\infty \frac{\sin^2(x)}{x}dx = \int_1^\infty \frac{(1-\cos(2x))/2}{x}dx = \int_1^\infty \frac{1}{2x}dx - \int_1^\infty \frac{\cos(2x)}{2x}dx,$$

the left term diverges and the second converges.

Summary. In this chapter we have discussed integrals. We have seen indefinite integral (anti derivative) and definite integral (area under a graph). We have seen to connection between the two (Newton-Leibnitz), as a tool to compute areas, and more generally to understand physical systems and other systems. We have also defined the area of an infinite domain.

Chapter 7

Series

In the previous section, we have discussed integrals, which are a sum over a continuous domain. We now move to discuss discrete sums.

We are given a sequence of numbers $a_1, a_2, \ldots$ and we would like to give a formal meaning to their sum (if it makes sense). For example, $1/2+1/4+1/8+\ldots=1$. Draw.

Definition 84. *Let $a_1, a_2, \ldots$ be a sequence of real numbers. Define*

$$S_n = a_1 + a_2 + \ldots + a_n.$$

We call $\sum_{n=1}^{\infty} a_n$ a series. Define

$$\sum_{n=1}^{\infty} a_n = \lim_{n\to\infty} S_n,$$

if the limit exists. If the limit exists we say the series ("tur") converges and write $|\sum_{n=1}^{\infty} a_n| < \infty$, and otherwise we say it diverges.

The letter n just denotes the index of summation. It can also be m, k et cetera ($\sum_{n=1}^{\infty} a_n = \sum_{k=1}^{\infty} a_k$).

There are 2 different questions we can ask. One is "does the series converge?" A more difficult one is "what is the limit?" We mostly focus on the easier question, but provide some simple examples for the harder one.

Geometric series. Let $q \in \mathbb{R}$ be so that $|q| < 1$. Let $a_n = q^{n-1}$ and

$$S_n = 1 + q + q^2 + \ldots + q^{n-1} = \frac{1-q^n}{1-q}.$$

So,

$$\sum_{n=0}^{\infty} q^n = \lim_{n\to\infty} S_n = \frac{1}{1-q}.$$

For example,

$$\sum_{n=0}^{\infty}(1/2)^n = 2,$$

and

$$\sum_{n=0}^{\infty}(-1/2)^n = 2/3.$$

If $|q| \geq 1$ then the series diverges.

Telescopic sum. Let

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

Then,

$$\sum_{n=1}^{\infty} a_n = \lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{k} - \frac{1}{k+1} = \lim_{n\to\infty} 1 - \frac{1}{n+1} = 1.$$

Non example. Let

$$a_n = \frac{1}{\sqrt{n}}.$$

Then,

$$S_n \geq n \cdot \frac{1}{\sqrt{n}} \to \infty,$$

when $n \to \infty$.

We see that although $a_n \to 0$ the series does not converge. For convergence, we need a_n to go fast enough to 0. But it is a necessary condition.

Theorem 85. *If* $\sum_{n=1}^{\infty} a_n$ *converges then* $\lim_{n\to\infty} a_n = 0$.

Proof. Write $a_n = S_n - S_{n-1}$ for $n > 1$. By arithmetic of limits,

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = 0.$$

□

Series have other useful properties, we now list 2 that are similar to properties we saw already (and will not prove but easily follows from properties of limits).

Arithmetic. If $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ converge then

$$\sum_{n=1}^{\infty}(\alpha a_n + \beta b_n) = \alpha\sum_{n=1}^{\infty} a_n + \beta\sum_{n=1}^{\infty} b_n,$$

for all $\alpha, \beta \in \mathbb{R}$.

Comparison. If $a_n \leq b_n$ for all n then

$$\sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} b_n.$$

The comparison test is useful to prove that monotone series converge.

Definition 86. *A positive series is a series $\sum_{n=1}^{\infty} a_n$ with $a_n > 0$ for all n.*

For example $a_n = \sum_{n=1}^{\infty} 1/n^2$. It is not easy to compute this limit, but the comparison test shows that it converges:

$$0 < \sum_{n=1}^{\infty} 1/n^2 \leq 1 + \sum_{n=2}^{\infty} \frac{1}{n(n-1)} < 3,$$

as we saw.

Another example is

$$\sum_{n=1}^{\infty} \sin(1/n^2) \leq C + \sum_{n=1}^{\infty} 2/n^2 < \infty$$

since $\sin(1/n^2) < 2/n^2$ for large n. Note that it is a positive series, so it converges.

Theorem 87. *If $\sum_{n=1}^{\infty} a_n$ is positive then it converges iff the sequence of partial sums $S_n = \sum_{i=1}^{n} a_i$ is bounded.*

Proof. This follows from that a monotone sequence converges iff it is bounded. □

There are 2 simple tests the guarantee convergence, if (a_n) is positive.

1. Root test (Cauchy): If there is $0 < q < 1$ so that $(a_n)^{1/n} \leq q$ for every large enough n then $\sum_{n=1}^{\infty} a_n$ converges. Indeed,

$$0 \leq \sum_{n=1}^{\infty} a_n \leq C + \sum_{n=1}^{\infty} q^n < \infty.$$

2. Ratio test (Delamber): If there is $0 < q < 1$ so that $\frac{a_{n+1}}{a_n} \leq q$ for every large enough n then $\sum_{n=1}^{\infty} a_n$ converges. Indeed, it follows by induction that $a_n < Cq^n$ for large enough n, and we can apply the previous argument.

Examples: 1. Let us consider

$$\sum_{n=1}^{\infty} \frac{2^n + 4^n}{3^n + 5^n}$$

in 3 different ways. The root test:

$$\left(\frac{2^n+4^n}{3^n+5^n}\right)^{1/n} \to 4/5.$$

The ratio test:

$$\frac{(2^{n+1}+4^{n+1})/(3^{n+1}+5^{n+1})}{(2^n+4^n)/(3^n+5^n)} \to 4/5.$$

And straight away to comparison test:

$$\sum_{n=1}^{\infty}\frac{2^n+4^n}{3^n+5^n} \le \sum_{n=1}^{\infty}\frac{2\cdot 4^n}{5^n} \le 2\sum_{n=1}^{\infty}(4/5)^n < \infty.$$

In general, use the comparison test, you need to find an upper bound that you already proved that converges (like a geometric series).

Another comment is that to use the comparison test you do not need to show that $a_n \le b_n$ for all n, it suffices to prove that $a_n \le b_n$ for large n, which specifically holds if $a_n/b_n \to L < \infty$. A similar statement holds for the other methods.

2. For every $A > 0$,

$$\sum_{n=1}^{\infty}\frac{A^n}{n!} < \infty$$

since

$$\frac{A^{n+1}/(n+1)!}{A^n/n!} = A/(n+1) < 1/2$$

for large enough n.

Another very useful test for convergence is the integral test.

Theorem 88. *Assume that f is non decreasing, non negative and integrable on $[0,\infty)$. Then,*

$$\sum_{n=1}^{\infty} f(n) \le \int_0^{\infty} f(x)dx \le \sum_{n=0}^{\infty} f(n).$$

Specifically, if the left integral diverges the sum diverges, and if the right integral converges then sum converges.

Proof. Draw: For every k it holds that

$$\int_0^k f(x)dx \le \sum_{n=1}^{k} f(n) \le \int_1^k f(x)dx.$$

□

Examples: 1.

$$\sum_{n=1}^{\infty} 1/n \geq \int_2^{\infty} 1/x dx = \infty.$$

This is called the harmonic series. It diverges.

The integral test can actually give a hint on the value of the sum. We can also deduce that $\sum_{n=1}^{k} 1/n$ is roughly $\ln(k)$.

2. The sum of $1/n^p$ for $0 < p \leq 1$ therefore also diverges. On the other hand, if $p > 1$ then the series converges.

3. Something we already saw:

$$\sum_{n=1}^{\infty} 1/n^2 \leq 1 + \int_1^{\infty} 1/x^2 dx < \infty.$$

4. Something we did not see:

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln(n))^q}$$

for $q > 0$. We use the integral test: it is at most

$$\int_1^{\infty} \frac{1}{x(\ln(x))^q} dx \underset{y=\ln(x), dy/dx=1/x}{=} \int_1^{\infty} \frac{1}{y^q} dy,$$

which converges iff $q > 1$ (also a lower bound holds). To conclude,

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln(n))^q}$$

converges iff $q > 1$. So, for example,

$$\sum_{n=2}^{\infty} \frac{1}{n \ln(n)} = \infty$$

and

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln(n))^2} < \infty.$$

7.1 Signs matter

When there is a series that is not positive, it is easier to work with it if the series absolutely converges.

Definition 89. *A series* $\sum_{n=1}^{\infty} a_n$ *absolutely converges if* $\sum_{n=1}^{\infty} |a_n|$ *converges. Otherwise, we say it conditionally converges.*

Theorem 90. *If* $\sum_{n=1}^{\infty} a_n$ *absolutely converges then it converges.*

Proof idea. The sum of all the negative parts converges to L_+ and the sum of all the positive parts converges to L_-. The overall series converges to $L_+ + L_-$. □

The other direction does not hold in general. The sequence $\sum_{n=1}^{\infty}(-1)^n/n$ does not absolutely converges but it does converge. For every n, let

$$\begin{aligned} S_n &= (-1)^1/1 + (-1)^2/2 + \ldots + (-1)^n/n \\ &= -1 + 1/2 - 1/3 + 1/4 - 1/5 + 1/6 - \ldots + (-1)^n/n \\ &= -1/2 - 1/12 - 1/30 - \ldots - \frac{1}{2i(2i+1)} - \ldots + (-1)^n/n. \end{aligned}$$

This is a negative sum, except perhaps the last term that tends to 0. The negative sum is bounded from below by $-\sum_{n=1}^{\infty} \frac{1}{n^2} > -\infty$ as we saw, so the sequence indeed converges.

In fact, the following always holds (we shall not prove).

Theorem 91 (Leibnitz)**.** *If* $a_n \to 0$ *as* $n \to \infty$ *and* $0 < a_n \leq a_{n+1}$ *for all* n *then* $\sum_{n=1}^{\infty}(-1)^n a_n$ *converges.*

Comment: Monotonicity is important (if $a_{2k} = 1/k$ and $a_{2k+1} = 1/k^2$ then the alternating sum does not converge).

Example: for all $p > 0$ the sum $\sum_{n=1}^{\infty}(-1)^n/n^p$ converges (as we saw for some p it does not absolutely converges).

Summary. In this chapter, we discuss infinite (but countable) sums. Defined as limit of partials sums. We saw that some sums converge, some diverges, and gave some simple criterions to decide.

Chapter 8

Power series

This is the last chapter in this course. In it, we represent functions in a certain useful way. There is a lot more to it than we shall get to.

Definition 92. *A power series is an expression of the form*

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n.$$

We will sometimes think of it as a function of x, but this does not always makes sense, and we need to understand when it does.

Example:

$$1 + x + x^2 + \ldots = \sum_{n=0}^{\infty} 1 \cdot (x - 0)^n.$$

This makes sense as a function only for $x \in (-1, 1)$. For such x,

$$1 + x + x^2 + \ldots = \frac{1}{1 - x}.$$

We shall mostly consider $x_0 = 0$, but remember that we can choose x_0 as we wish.

A power series is a "polynomial of infinite degree." As mention, it does not always make sense. Therefore, we mark the domain in which it does as follows.

Definition 93. *The domain of convergence of* $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ *is* $D \subseteq \mathbb{R}$ *if for every* $x \in D$ *the sum converges.*

Examples:

1. The domain of $\sum_{n=0}^{\infty} x^n$ is $(-1, 1)$.
2. The domain of $\sum_{n=0}^{\infty} x^n / n!$ is $\mathbb{R}$.
3. The domain of $\sum_{n=0}^{\infty} x^n / n$ is $[-1, 1)$.

4. The domain of $\sum_{n=0}^{\infty} x^{2n}$ is $(-1,1)$; it is equal to $\sum_{n=0}^{\infty}(x^2)^n$ so converges iff $|x^2| < 1$ or $|x| < 1$.

As a rule, the smaller the coefficients are, the larger to domain is.

In general, a power series has a radius of convergence.

Definition 94. *The radius of convergence of $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ is the largest (supremum) number $r > 0$ so that for every x so that $|x - x_0| < r$ the sum converges.*

Theorem 95 (Cauchy-Hadamard)**.** *For every power series $\sum_{n=0}^{\infty} a_n x^n$ there is $0 \leq R \leq \infty$ so that for every x so that $|x| < R$ the series converges at x. (We allow R to be infinity as well.) And for every x so that $|x| > R$ the sum diverges (so if $R = \infty$ the sum always converges). At the two points $\pm R$ the sum may converge or diverge.*

Moreover, if the following limit exists

$$\limsup_{n\to\infty} a_n^{1/n} = L$$

then

$$R = \frac{1}{L}.$$

(If $L = 0$ then $R = \infty$.) Similarly, if the following limit exists

$$\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = L$$

then

$$R = \frac{1}{L}.$$

This gives two ways to compute the radius of convergence of a given series. It is a way to understand when does the infinite sum makes sense. We shall not fully prove, but explain why it works.

In examples 1,2,4 above $R = 1$ and in 3 it is ∞.

First, what is lim sup? Given (a_n), for every n define $b_n = \sup\{a_k : k \geq n\}$. It is a new sequence (which is non increasing). Define

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

Specifically, the limits of a_n exists then it is equal to that of b_n and the lim sup is just lim. (Focus on this case.)

Second, why does it work? Assume $x_0 = 0$ and assume that the limit of $a_n^{1/n}$ exists and equals to L. Thus, for every $K > L$ there is n_0 so that if $n > n_0$,

$$a_n^{1/n} \leq K.$$

So, if $|x| < 1/K$ then the series absolutely converges (and hence converges):

$$\sum_{n=0}^{\infty} |a_n x^n| = \sum_{n=0}^{n_0} |a_n x^n| + \sum_{n=n_0+1}^{\infty} |a_n x^n| \leq \sum_{n=0}^{n_0} |a_n x^n| + \sum_{n=n_0+1}^{\infty} K^n x^n < \infty,$$

since the first sum is finite and the second sum is a geometric series.

Third, examples:

1. The radius of $\sum_{n=0}^{\infty} (2/n) x^n$ is 1 since $\limsup_{n\to\infty} (2/n)^{1/n} = 1$.
2. The radius of $\sum_{n=0}^{\infty} x^n/n!$ is ∞ since $\limsup_{n\to\infty} \frac{n!}{(n+1)!} = 0$.
3. The radius of $\sum_{n=0}^{\infty} (x-1)^n$ is 1.
4. The radius of $\sum_{n=0}^{\infty} n x^n$ is 0; for every x the sum diverges.

Fourth, properties:

Theorem 96 (Continuity)**.** *If the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$ is R then for every $x \in (-R, R)$, the function*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

is continuous at x.

Example is $\sum_{n=0}^{\infty} x^n = 1/(1-x)$.

Theorem 97 (Arithmetic)**.** *If the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$ is at least R and $\sum_{n=0} b_n x^n$ is at least R the for every $x \in (-R, R)$, the function*

$$\sum_{n=0}^{\infty} (c a_n + d b_n) x^n = c \sum_{n=0}^{\infty} a_n x^n + d \sum_{n=0}^{\infty} b_n x^n.$$

Theorem 98 (Derivative and integration)**.** *Assume the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$ is R. For every $x \in (-R, R)$, let*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n.$$

The function f is differentiable in $(-R, R)$ and

$$f'(x) = \sum_{n=0}^{\infty} a_n n x^{n-1}.$$

The function f is integrable in $(-R, R)$ and

$$\int f(x)dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}.$$

In short, we can do integration and derivative inside the sum, but only inside the radius of convergence.

In fact, since we can take on derivative, we can take as many derivatives as we want. Before we saw that $(a_n) \mapsto f$. This actually tells us how to get from f to (a_n). Indeed,

$$f(0) = a_0 + a_1 0 + a_2 0 + \ldots = a_0,$$

$$f'(0) = 0 + a_1 + 2a_2 x + 3a_3 x^2 + \ldots \Big|_{x=0} = a_1$$

and in general

$$f^{(n)}(0) = n(n-1)(n-2)\ldots 1a_n,$$

or

$$a_n = \frac{f^{(n)}(0)}{n!},$$

which is exactly the coefficient of x^k in the Taylor polynomial of f around 0. We can write f as

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n,$$

inside the radius of convergence. This is called the Taylor expansion of f around 0. The understand when it is meaningful, we need to compute the radius of convergence.

Example:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

This holds for all $x \in \mathbb{R}$.

This and similar representations of a function is very useful in science. For example, in physics it gives a way to solve the wave equation.

Another example: Consider

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

for $|x| < 1$. Integrate to get

$$\sum_{n=0}^{\infty} x^{n+1}/(n+1) = -\ln(1-x).$$

Specifically,

$$\sum_{n=0}^{\infty}(1/2)^{n+1}/(n+1) = \frac{1}{2}\frac{1}{1} + \frac{1}{2^2}\frac{1}{2} + \frac{1}{2^3}\frac{1}{3} + \ldots = \ln(2) \approx 0.693147$$

and (with a little faith)

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \ln(1-(-1)) = \ln(2).$$