MTH1030/1035 Lecture Notes 2022

Part 2: Calculus

Burkard Polster

Contents

1 Calculus and limits **9**

2 Sequences and series **17**

2.1 The limit of a sequence again 17

2.2 Infinite series 23

2.2.1 A telescoping sum 25

2.2.2 The harmonic series diverges 26

2.2.3 Integral comparison test 27

2.2.4 Integral test remainder estimate 30

2.3 Some more basic facts about series 33

2.4 Representing functions by infinite series 35

2.5 Taylor series 41

2.6 Biting the bullet: Taylor's formula
(MTH1035 and MTH1040 only) 42

2.6.1 Applying Taylor's formula 44

2.7 A bit of bad news 47

2.8 Further applications of Taylor series 48

2.8.1 The number e 48

2.8.2 New Taylor series from old ones 49

2.8.3 Extending real valued functions to complex valued functions: Euler's formula 50

2.8.4 The Number π: Life beyond the circle 51

2.9 Absolute convergence 52

2.10 The ratio test 54

2.11 Power series 56

2.11.1 Convergence 57

2.11.2 Differentiating and integrating power series 59

2.11.3 Power series in powers of $x - c$ 60

3 Integration **63**

3.1 Some integration techniques 63

3.1.1 Elementary functions 63

3.1.2 Differentiating and integrating 63

3.1.3 Integration by parts 66

3.1.4 Substitution 69

3.2 Proper and improper integrals 73

3.2.1 Geometrical interpretation of integrals in a nutshell 73

3.2.2 Improper integrals 76

4 Differential equations **83**

4.1 Some easy first order stuff: Separation of variables 85

4.2 The integrating factor trick 89

4.3 Second order linear differential equations 93

4.3.1 Homogeneous equations 93

4.3.2 Non-homogeneous equations 96

4.3.3 General and more general 100

Preface

In the linear algebra part of this unit, we develop basic linear algebra in a lot of detail, both in terms of theory and applications. This is not easy but definitely doable since linear algebra has very minimal prerequisites. Calculus is a different story altogether.

To develop the theory behind calculus properly, even to just fully justify all the things you do in school would take up most of the lectures of this unit. This would definitely be a worthwhile project and you'll undertake it in second year if maths is "your thing". However, for the majority of you using the tools of calculus is really all that you'll ever need to know and in this unit we'll content ourselves with introducing you to more calculus tools that extend and complement those that are already at your disposal.

Having said this, one of our main aims will be to justify pretty much everything we do, just not from first principles but rather starting with all those things you already "know", such as: all our favourite functions like $\sin(x)$, e^x, polynomials, etc., are infinitely often differentiable, $(\sin(x))' = \cos(x)$, differentiation and integration rules work as usual, etc.

A lot of the overall structure, definitions, phrasing of theorems and the worked examples in these lecture notes were adapted from my favourite calculus textbook *Calculus, Early transcendentals* by Edwards and Penny. Here and there you'll also find some leftover fragments from the old MTH1030 lecture notes which were prepared by my colleague Dr Leo Brewin.

Melbourne, January 2017

Burkard Polster

Chapter 1

Calculus and limits

The whole of calculus is based on limits: continuity, derivatives, integrals, etc., are all defined via limits. Here is a very brief sketch to show you how we get from the basic notion of the limit of an infinite sequence to all those other familiar concepts.

Limit of a sequence

Let's start with an infinite sequence of numbers $\{a_n\}$. This sequence has a limit L if, no matter how small a number $\epsilon > 0$ you choose, from a certain n onward the absolute difference $|L - a_n|$ between L and a_n is always less than ϵ. We express this by saying

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

For example, for the sequence

$$\left\{\frac{1}{n}\right\} = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$$

we find that $\lim\limits_{n\to\infty} \frac{1}{n} = 0$ because no matter what tiny $\epsilon > 0$ we choose, $|\frac{1}{n} - 0| = \frac{1}{n}$, the absolute difference between the nth term of the sequence and 0 is less than ϵ from a certain n onward.[1] If an infinite sequence has a limit we say that it **converges**. If an infinite sequence does not converge it **diverges**.

It is important to note that divergence of a sequence can mean a number of different things. The term divergence suggests that it should mean that somehow, as you move further along the sequence, you come across larger and larger numbers. This is true for many divergent sequences like

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

[1] In fact this will be the case for all $n > \frac{1}{\epsilon}$.

However, sequences like

$$1, -1, 1, -1, 1, ...$$

also diverge, despite their elements not getting arbitrarily large. The only thing that matters here is that these sequences don't converge, that they don't have a limit.

Limit of a function

Apart from limits of sequences there are also **limits of functions**. There are a number of different but equivalent ways to introduce limits of functions. One such way is via limits of sequences.

Suppose you have a real-valued function $f(x)$ that is defined in an open interval, except possibly at some point a in this interval. Then the function $f(x)$ has the limit L at the point a if for EVERY sequence of numbers $\{a_n\}$ in the interval that avoids our point a and at the same time converges to a, the corresponding sequence $\{f(a_n)\}$ converges to L. We express this by writing

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

If we actually wanted to use this definition to verify that

$$\lim_{x \to 0} x^2 = 0$$

we would have to check that for our converging sequence $\{\frac{1}{n}\}$ with limit 0 the corresponding sequence $\{f(\frac{1}{n})\} = \{\frac{1}{n^2}\}$ also converges to 0 which, of course, it does. However, to be absolutely sure we would have to do the same for ALL those infinitely many other sequences that converge to 0.

This is not what mathematicians usually do in practice when they want to check that a function has a limit at a certain point. For this task other (equivalent) definitions of the limit of a function are more suited.[2]

Continuity, derivative and integral

Thinking of limits of functions in terms of limits of sequences is about a close as you can get to capturing the intuitive notion of continuity of a function at a: "no matter how you sneak up on a, the function sneaks up on $f(a)$".

[2] First and foremost among these equivalent definitions of the limit of a function is the so-called *epsilon-delta definition of limits* which you'll explore in some advanced second year units: $\lim_{x \to a} f(x) = L$ if and only if given any number $\epsilon > 0$, there exists a corresponding number $\delta > 0$ such that $|f(x) - L| < \epsilon$ for all x such that $0 < |x - a| < \delta$.

Formally, we say that our function is **continuous** at a if the limit

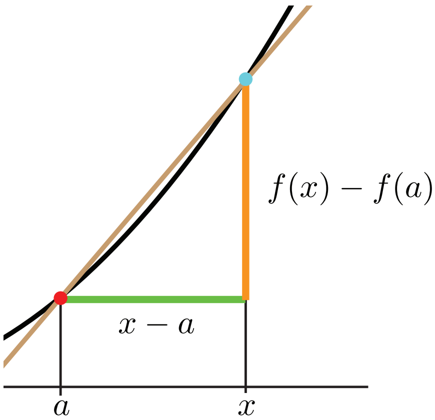
$$\lim_{x \to a} f(x)$$

exists and is equal to $f(a)$. Furthermore, the function is continuous in the whole interval if it is continuous at all points of the interval.

Our function is **differentiable** at a if

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

exists. And if it does exist this limit is called the **derivative** of $f(x)$ at a. You should all be familiar with the following picture that goes with this limit.

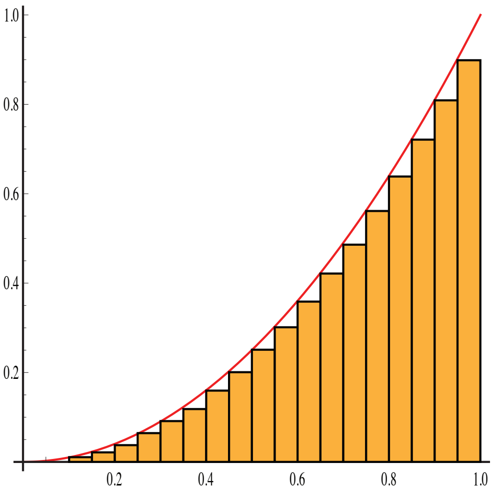


Here is what this picture illustrates: We are interested in the slope of our (nice) function at the red point, that is, the slope of the tangent to the black graph of the function at this point. We can approximate this slope by the slope $(f(x) - f(a))/(x - a)$ of the line through the red point and the nearby light blue point. We get better and better approximations as we move the light blue point towards the red point and in the limit we get the slope we are really interested in.

And what about integrals? They, too, are defined in terms of limits. For example,

$$\int_0^1 x^2 dx = \lim_{k \to \infty} \sum_{n=0}^{k-1} \left(\frac{1}{k}\right) \left(\frac{n}{k}\right)^2,$$

which is the limit of the area of a "staircase" with k steps of width $1/k$ that approximates the graph of x^2. Here is the staircase that corresponds to $k = 20$.



The definition of an integral is a generalization of this basic idea of approximating areas under graphs of functions by staircases.

Our limits

Anyway, as I said, overall we'll just trust Newton and Co. that functions like e^x are really continuous, differentiable, etc. This then immediately gives us lots of limits to play with, e.g.,

$$\lim_{x \to a} e^x = e^a$$

for any real number a—the limit of a continuous function $f(x)$ as x approaches a point a where the function is defined is just the value of the function $f(a)$ at this point. We'll also use the extension of finite limits to infinite ones, as usual. Examples include

$$\lim_{x \to \infty} x^n = \lim_{x \to \infty} e^x = \infty$$

and

$$\lim_{x \to \infty} \frac{1}{x^7} = \lim_{x \to -\infty} e^x = 0.$$

Furthermore, we'll pretend that we've convinced ourselves that the following basic type of limit really exists:

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

With this set of limits at our disposal, plus a few more facts from school maths, plus a lot of brain power we are ready to tackle the rest of this unit ....

A word of warning

So it's all good, right? In fact, you may be under the impression that you have a pretty good understanding of what continuous or differentiable functions look like and how they behave, and you may be wondering whether all this limit stuff is a bit over the top. However, these supposedly well-behaved functions can behave in a very unexpected manner. Here are examples of functions illustrating just a few of the counterintuitive things that can happen.

1. Let

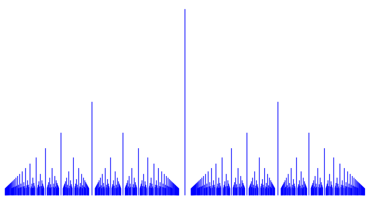
$$f(x) = \begin{cases} x & \text{if } x \text{ is a rational number;} \\ 0 & \text{if } x \text{ is an irrational number.} \end{cases}$$

   This is a function which is defined for all $x \in \mathbf{R}$ but is only continuous at $x = 0$.

2. Let

$$t(x) = \begin{cases} 1 & \text{if } x = 0; \\ \frac{1}{n} & \text{if } x \text{ is a rational number, } x = m/n \text{ in lowest terms, } n > 0; \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

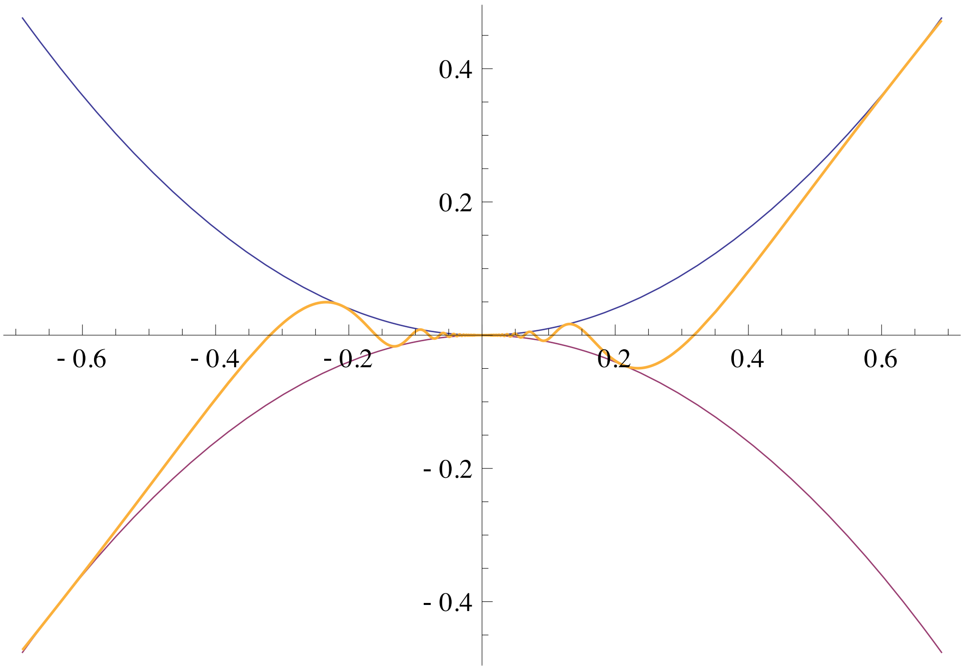
   This function is called **Thomae's function** and is defined for all $x \in \mathbf{R}$. It is continuous at every irrational point but discontinuous at every rational point. Here is a sketch of this function around 0. The spike in the middle stands for $t(0) = 1$, the second highest spikes for $t(1/2) = t(-1/2) = 1/2$, the third highest spikes for $t(1/3) = t(-1/3) = t(2/3) = t(-2/3) = 1/3$, and so on.



3. Let

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x}) & \text{if } x \neq 0; \\ 0 & \text{if } x = 0. \end{cases}$$

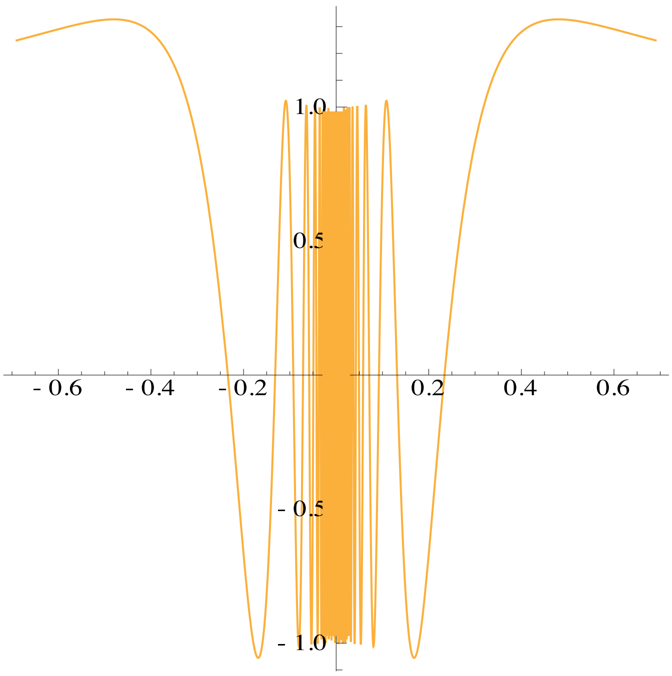
   This is a differentiable function with a discontinuous derivative. The graph of the function is the orange curve in the following picture. It is oscillating between the graphs of x^2 and $-x^2$ drawn in blue and purple. Obviously, the function is differentiable everywhere off 0. And since it is boxed in between x^2 and its negative and both these functions have derivative 0 at $x = 0$ we can conclude that our function must have the same derivative at 0. This means that the function is everywhere differentiable.



However, the derivative of $f(x)$ is

$$f'(x) = \begin{cases} 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) & \text{if } x \neq 0; \\ 0 & \text{if } x = 0, \end{cases}$$

which is discontinuous at $x = 0$. Its graph looks something like this



4. There are even functions like the famous **Weierstrass function** that are continuous everywhere but nowhere differentiable.

5. E.t.c.

For the moment you may pretend that you know what a limit is and I'll make sure that we stay clear of these sort of monsters. However, if you really want to understand what makes calculus tick you'll have to take a unit on *Mathematical Analysis*.

Check out the book *Counterexamples in Analysis* by Gelbaum and Olmsted in the library if you cannot wait to see more calculus monsters. Also check out the textbook *Calculus* by Michael Spivak, one of the best rigorous introductions to calculus.

Chapter 2

Sequences and series

In this chapter we'll be considering infinite sequences and limits of such sequences. Let's start by looking at a few examples of infinite sequences and ways to write them. Here we go,

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots, \frac{1}{n}, \cdots.$$

A more compact way of writing this sequence is

$$\left\{\frac{1}{n}\right\}_{n=1}^{\infty} \text{ or simply } \left\{\frac{1}{n}\right\}.$$

Or, we could just write down the nth element of the sequence

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

2.1 The limit of a sequence again

We've already defined what we mean by the limit of a sequence in the introductory chapter of these lecture notes. But just to keep things self-contained here is this definition and the accompanying discussion again with a few more important details added.

Definition 2.1.1 (Limit of a sequence) *We say that the sequence $\{a_n\}$* ***converges to the real number*** *L, or has the* ***limit*** *L, and we write*

$$\lim_{n\to\infty} a_n = L,$$

provided that a_n can be made as close to L as we like by choosing n to be sufficiently large. That is, given any number $\epsilon > 0$, there exists a natural number N such that

$$|a_n - L| < \epsilon \text{ for all } n \geqq N.$$

If the sequence $\{a_n\}$ does not converge, then we say that it **diverges**.

Let's say somebody asks you to *prove* that the sequence $\{1/n\}_{n=1}^{\infty}$ converges to zero, that is,

$$\lim_{n\to\infty} \frac{1}{n} = 0,$$

then here is how you should argue. Because in this case $L = 0$, we need only convince ourselves that to each positive number ϵ there corresponds a natural number N such that

$$\left|\frac{1}{n}\right| = \frac{1}{n} < \epsilon$$

if $n \geqq N$. Obviously, any fixed natural number $N > 1/\epsilon$ will do, because $n \geqq N$ implies that

$$\frac{1}{n} \leq \frac{1}{N} < \epsilon,$$

which then completes the proof.[1]

It is important to note that divergence of a sequence can mean a number of different things. The term divergence suggests that it should mean that somehow, as you move further along the sequence, you come across larger and larger numbers. This is true for many divergent sequences like

$$1, 2, 3, 4, ...$$

However, sequences like

$$1, -1, 1, -1, 1, ...$$

also diverge, despite their elements not getting arbitrarily large. The only thing that matters here is that these sequences don't converge, that they don't have a limit.

On the other hand, it is important to have a concise way of expressing that a sequence does get arbitrarily large.

Definition 2.1.2 (Divergence to infinity) *A sequence $\{a_n\}$* **diverges to** ∞ *if, given any number $R > 0$, there exists a natural number N such that*

$$R < a_n \text{ for all } n \geq N.$$

We then say that

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

A sequence $\{a_n\}$ **diverges to** $-\infty$ *if the sequence $\{-a_n\}$ diverges to ∞ We then say that*

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

[1]For example, if $\epsilon = 10^{-6}$, then choosing $N > 10^6$ will do the trick.

For example, $\lim_{n\to\infty} n = \lim_{n\to\infty} n^3 = \lim_{n\to\infty} 2^n = \infty$. (Boring, I know.)

We'll state a bunch of theorems about limits that are intuitively obvious. We'll skip the proofs which are all very easy. For each of these theorems just make sure that you understand what it says.

Theorem 2.1.1 (Limit laws for sequences) *Assume that the limits* $\lim_{n\to\infty} a_n = A$ *and* $\lim_{n\to\infty} b_n = B$ *exist. Then*

1. $\lim_{n\to\infty} ca_n = cA$*, for* c *being any real number;*
2. $\lim_{n\to\infty} (a_n + b_n) = A + B$*;*
3. $\lim_{n\to\infty} a_n b_n = AB$*;*
4. *if* $B \neq 0$*, then* $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{A}{B}$.

Note that these rules also hold with obvious provisos if one or both of the limits A and B are equal to $+\infty$ or $-\infty$.

Theorem 2.1.2 (Substitution law for sequences) *If* $\lim_{n\to\infty} a_n = A$ *and if* $f(x)$ *is a function that is continuous at* $x = A$*, then*

$$\lim_{n\to\infty} f(a_n) = f(A).$$

For example, if $f(x) = \cos(x)$ and $a_n = 1/n$, then the limit of the sequence is 0 and

$$\lim_{n\to\infty} f(a_n) = \cos(0) = 1.$$

Theorem 2.1.3 (Squeeze law for sequences) *If* $a_n \leq b_n \leq c_n$ *for all* n *and*

$$\lim_{n\to\infty} a_n = L = \lim_{n\to\infty} c_n,$$

then $\lim_{n\to\infty} b_n = L$ *as well.*

Let's use the squeeze law to show that

$$\lim_{n\to\infty} \frac{\sin(n)}{n} = 0.$$

Since $-1 \leq \sin(x) \leq 1$ we have

$$-\frac{1}{n} \leq \frac{\sin(n)}{n} \leq \frac{1}{n}$$

for every positive integer n. Therefore

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

We conclude that

$$\lim_{n\to\infty}\frac{\sin(n)}{n} = 0.$$

Theorem 2.1.4 (Limit of function equals limit of sequence) *If $f(x)$ is a real-valued function and $a_n = f(n)$ for each positive integer n, then*

$$\lim_{x\to\infty} f(x) = L \textit{ implies that } \lim_{n\to\infty} a_n = L.$$

The converse of the statement in this theorem is generally false. For example, let

$$f(x) = \sin(\pi x)$$

and let $a_n = f(n) = \sin(n\pi) = 0$. Obviously,

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

as well. On the other hand, $\sin(nx)$ oscillates between 1 and -1, and therefore

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

does not exist.

Our final theorem concerns a special class of sequences. But to be able to state this theorem we'll need the following definition.

Definition 2.1.3 (Monotone sequences) *The sequence $\{a_n\}$ is said to be* **increasing** *if*

$$a_1 \leq a_2 \leq a_3 \leq \cdots \leq a_n \leq \cdots$$

and **decreasing** *if*

$$a_1 \geqq a_2 \geqq a_3 \geqq \cdots \geqq a_n \geqq \cdots$$

The sequence $\{a_n\}$ is **monotonic** *if it is either decreasing or increasing. The sequence $\{a_n\}$ is* **bounded** *if there is a number M such that $|a_n| \leq M$ for all n.*[2]

[2]Note in some calculus books a sequence is called increasing if $a_1 < a_2 < a_3 < \cdots < a_n < \cdots$ and non-decreasing if $a_1 \leq a_2 \leq a_3 \leq \cdots \leq a_n \leq \cdots$. People like us who call sequences $a_1 \leq a_2 \leq a_3 \leq \cdots \leq a_n \leq \cdots$ increasing usually call sequences with $a_1 < a_2 < a_3 < \cdots < a_n < \cdots$ strictly increasing.

Note that an increasing sequence is automatically bounded from below by its first element. Therefore to show that an increasing sequence is bounded it suffices to find an **upper bound**, that is, a number that is greater than all elements of the sequence. Similarly, to show that a decreasing sequence is bounded it suffices to find a lower bound.

Theorem 2.1.5 (Bounded monotonic sequences) *Every bounded monotonic infinite sequence converges, that is, has a finite limit.*

This theorem is about the existence of a limit but does not specify how to find it. However, it's not hard to see what this limit will be. Let's say we are dealing with an increasing sequence that is bounded above, that is, there is a number M that is greater than all elements of the sequence. If the sequence is eventually constant, that is, all elements of the sequence are equal to a certain number from some point onward, then this number is the limit of the sequence. Let's say our increasing sequence is not of this type. This means that the elements of the sequence are getting bigger and bigger. Then the limit is just the minimum of all the upper bounds of the sequence.

For example, the sequence

$$-1, -\frac{1}{2}, -\frac{1}{3}, -\frac{1}{4}, \ldots$$

is an increasing sequence. It has lots of upper bounds: 0, 5, 13, 1000, and so on. In fact all the positive real numbers are upper bounds. However, the minimum of these upper bounds is 0, which is also the limit of our sequence.

Intuitively the existence of a minimum of upper bounds is fairly obvious. However, the existence of this minimum is something very subtle and requires that the real numbers are *complete*, that is, have no holes. In fact, a common way of constructing the real numbers from scratch uses this sort of completeness as one of its axioms.

Let's consider the following infinite sequence:

$$0, 0.2, 0.23, 0.235, 0.2357, 0.235711, 0.23571113, \cdots.$$

So, you make larger and larger decimal numbers by attaching subsequent primes. Obviously, this sequence has an upper bound.[3] This means that by the last theorem this sequence has a limit. In fact, this limit can be seen to be an irrational number since the decimal number it corresponds to does not have a repeating tail.

Let's ponder the following sequence of numbers.

$$\sqrt{6},\ \sqrt{6+\sqrt{6}},\ \sqrt{6+\sqrt{6+\sqrt{6}}},\ \sqrt{6+\sqrt{6+\sqrt{6+\sqrt{6}}}}, \cdots$$

[3]The choice of the upper bound does not matter. We could as well have chosen 25 or 666.

To turn the nth element of this sequence into the $(n+1)$st element we add 6 and find the square root of the new number, that is, $a_{n+1} = \sqrt{6 + a_n}$. This implies that if this sequence has a limit L, then $\sqrt{6+L} = L$. Solving this equation gives that the only candidates for L are -2 and 3. Obviously -2 is not possible. So we are aiming for 3.

If we turn the first couple of elements of our sequence into decimal numbers using a calculator we get, 2.4494..., 2.906..., 2.984..., etc. So, 3 seems like a good candidate for a limit. In fact, our numerical experiment suggests that we are dealing with an increasing sequence. This means that to show that 3 is really the limit it suffices, by our last theorem, to show that the sequence is really increasing and has 3 as an upper limit.

To show that 3 is an upper limit note that $a_n < 3$ implies that $a_{n+1} = \sqrt{6 + a_n} < \sqrt{6+3} = 3$. Then since $a_1 < 3$, it follows by induction that $a_n < 3$ for all n.

It remains to show that it is an increasing sequence. But

$$(a_{n+1})^2 - (a_n)^2 = (6 + a_n) - (a_n)^2 = (2 + a_n)(3 - a_n) > 0$$

because $a_n < 3$. Because all terms of the sequence are positive, it therefore follows that $a_{n+1} > a_n$ for all $n \geqq 1$, as desired.

The last two examples are really about making sense of the infinite expressions

$$0.23571113...$$

and

$$\cdots + \sqrt{6 + \sqrt{6 + \sqrt{6 + \sqrt{6}}}}.$$

What you do is to interpret an expression like this as an infinite sequence. If this sequence has a limit, then we say that the infinite expression **as a number** is equal to this limit. Other examples that can be dealt with in a similar fashion are

$$\sqrt{2\sqrt{2\sqrt{2\sqrt{2\cdots}}}},$$

$$1 + \frac{3 + \frac{1 + \frac{3+\cdots}{2+\cdots}}{3 + \frac{1+\cdots}{3+\cdots}}}{2 + \frac{2 + \frac{2+\cdots}{1+\cdots}}{1 + \frac{3+\cdots}{2+\cdots}}}$$

and the intriguing identity

$$\sqrt{q + p\sqrt{q + p\sqrt{q + p\sqrt{q + \cdots}}}} = p + \frac{q}{p + \frac{q}{p + \frac{q}{p + \cdots}}}$$

This also introduces us to our next topic.

2.2 Infinite series

Definition 2.2.1 *An* **infinite series** *is an expression of the form*

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots + a_n + \cdots$$

where $\{a_n\}$ is an infinite sequence of real numbers. The number a_n is called the **nth term** *of the series. Note that sometimes, when it is more convenient, we'll start the summation not at $n = 1$ but instead at $n = 0$ or another positive integer.*

As with the infinite expressions that we discussed in the previous section, we make sense of an infinite sum by interpreting it as a sequence, the sequence of **partial sums**,

$$S_1, S_2, S_3, \dots S_n, \dots,$$

where

$$S_1 = a_1,$$

$$S_2 = a_1 + a_2,$$

$$S_3 = a_1 + a_2 + a_3,$$

and so on. We define the sum of an infinite series to be the limit of its sequence of partial sums, provided that this limit exists.

Definition 2.2.2 (The sum of an infinite series) *We say that the infinite series $\sum_{n=1}^{\infty} a_n$ converges (or is convergent) with sum S provided that the limit of its sequence of partial sums,*

$$S = \lim_{n\to\infty} S_n$$

exists (and is finite). Otherwise we say that the series diverges (or is divergent). If a series diverges, then it has no sum.

In other words, the sum of an infinite series is a limit of finite sums,

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

provided that this limit exists.

Geometric series

Let's kick off by revisiting one type of infinite series that you are already familiar with from school.

Definition 2.2.3 *A* **geometric series** *is an infinite series of the form*

$$\sum_{n=0}^{\infty} ar^n,$$

where a and r are constants and $a \neq 0$.

An example of a geometric series is

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots + \frac{1}{2^n} + \cdots .$$

In this example, $a = 1$ and $r = \frac{1}{2}$.

The nth partial sum of the geometric series is

$$S_n = a(1 + r + r^2 + \ldots + r^n).$$

The standard trick to calculate this partial sum goes as follows.

Multiply by r, then

$$rS_n = a(r + r^2 + \ldots + r^n + r^{n+1}).$$

Subtracting the second equation from the first the red sum cancels and we get

$$S_n(1-r) = a(1 - r^{n+1})$$

and therefore, if $r \neq 1$,

$$S_n = a\frac{1 - r^{n+1}}{1 - r}.$$

To figure out whether or not the geometric series converges we have to establish for which combinations of a and r the sequence of partial sums converges. However, clearly,

$$\lim_{n\to\infty} a\frac{1 - r^{n+1}}{1 - r} = \frac{a}{1-r}$$

for $|r| < 1$ and the series diverges for $|r| > 1$. It is also easy to see that geometric series with $r = 1$ or $r = -1$ diverge, too.

So, for example,

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

We summarize this result in the following theorem.

Theorem 2.2.1 (The sum of a geometric series) *If $|r| < 1$, then the geometric series converges, and its sum is*

$$S = \sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}.$$

If $|r| \geqq 1$, then the geometric series diverges.

Here is a nice application. Using this result we can find the rational number represented by a given infinite repeating decimal. Here are two examples that show what needs to be done

$$\begin{aligned} 0.99999\ldots &= \frac{9}{10} + \frac{9}{100} + \frac{9}{1000} + \cdots = \frac{9}{10}\left(1 + \frac{1}{10} + \frac{1}{100} + \cdots\right) \\ &= \frac{9}{10}\sum_{n=0}^{\infty}\left(\frac{1}{10}\right)^n = \frac{9}{10}\frac{1}{1-\frac{1}{10}} = \frac{9}{10}\cdot\frac{10}{9} = 1. \end{aligned}$$

And here is how you take care of a more complicated case.

$$0.97537373737\ldots = \frac{975}{1000} + \frac{37}{100000}\sum_{n=0}^{\infty}\left(\frac{1}{100}\right)^n = \frac{975}{1000} + \frac{37}{100000}\frac{1}{1-\frac{1}{100}} = \frac{48281}{49500}.$$

In fact, we can use this technique to prove that every repeating decimal is equal to a rational number. On the other hand, it is not hard to show that every rational number corresponds to a repeating decimal.[4] Together these results amount to a characterization of the rational numbers in terms of their decimal expansions and a characterization of the irrational numbers as those numbers whose decimal expansions are non-repeating.

2.2.1 A telescoping sum

There are only very few interesting classes of infinite series for which it is possible to find simple formulas for the partial sums. The geometric series represent one such type. The so-called **telescoping sums** are another. Here is the simplest example of a telescoping sum.

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

[4] The standard proof uses long division.

The nth term of this series can be written as follows

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

It follows that the sum of the first n terms of the given series is

$$\begin{aligned} S_n &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{4} - \frac{1}{5}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= 1 + \left(\frac{1}{2} - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{3}\right) + \left(\frac{1}{4} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n}\right) - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1} = \frac{n}{n+1}. \end{aligned}$$

Note that we just collapsed the red middle part of the "telescope". Hence

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \lim_{n\to\infty} \frac{n}{n+1} = 1.$$

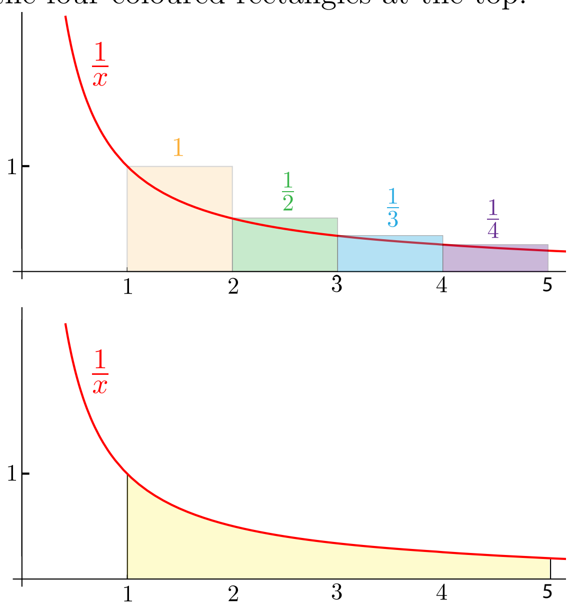
We'll look at a few more examples of telescoping sums in the tutorials.

2.2.2 The harmonic series diverges

The **harmonic series** is the infinite series

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

The following diagram shows that the fourth partial sum S_4 of the harmonic series is equal to the combined area of the four coloured rectangles at the top.



Then it is clear that

$$\int_1^5 \frac{1}{x}\, dx = \text{ yellow area under the graph } < \text{ area of the coloured rectangles } = S_4.$$

In general we find

$$\int_1^{n+1} \frac{1}{x}\, dx < S_n,$$

or

$$[\ln(x)]_1^{n+1} = \ln(n+1) < S_n.$$

However, since $\lim_{n\to\infty} \ln(n+1) = \infty$ we conclude that

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges to infinity as well.

2.2.3 Integral comparison test

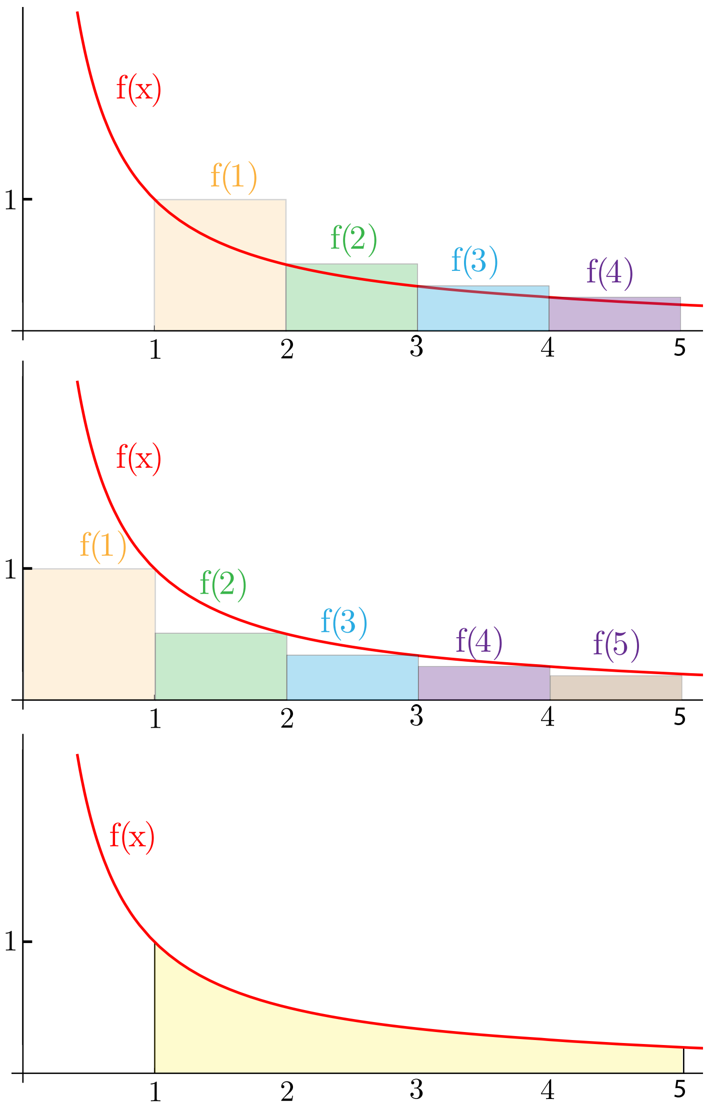
The method that we just used to show that the harmonic series diverges generalizes and can be used to show that certain series diverge and others converge. This is done by comparing the partial sums of qualifying series with integrals.

For us to be able to use this method the nth term of our series a_n has to be given in terms of a non-negative decreasing function $f(n)$ that can be integrated.[5] So let

$$a_n = f(n).$$

Here is our picture again using the general setup. Note that we've added a third diagram in the middle. It results from the first diagram by shifting all the rectangle by one unit to the left. This has the effect that now all of the rectangles end up *below* the graph of the function $f(x)$ whereas in the first diagram they were sticking out *above* the graph of $f(x)$.

[5]Actually it is true that every decreasing function is automatically integrable, but since "we" don't know this I've added this as an extra requirement.



Now let's translate the pictures into formulas.

The first diagram combined with the last diagram tells us that, as in the case of the harmonic series,

$$\text{yellow area under the graph } = \int_1^{n+1} f(x)\, dx < \text{ area of the coloured rectangles } = S_n,$$

On the other hand, the second diagram combined with the last diagram tells us that

$$S_{n+1} - f(1) = \text{ area of rectangles 2,3,4,5 } < \text{ yellow area under the graph } = \int_1^{n+1} f(x)\, dx.$$

Combining the two inequalities we get

$$S_{n+1} - f(1) < \int_1^{n+1} f(x)\, dx < S_n.$$

There are only two possibilities for

$$\lim_{n\to\infty}\int_1^{n+1} f(x)\ dx = \int_1^{\infty} f(x)\ dx.$$

Either the limit exists and is a positive number, or it is ∞. The same is true for the (increasing) sequence of partial sums. Therefore,

$$\lim_{n\to\infty} S_{n+1} - f(1) < \int_1^{\infty} f(x)\ dx < \lim_{n\to\infty} S_n.$$

Finally, these inequalities translate into the following theorem.

Theorem 2.2.2 (Integral test) *Let $f(x)$ be a non-negative decreasing function in the interval $[1,\infty)$. Then*

$$\sum_{n=1}^{\infty} f(n)$$

converges (diverges to infinity) if and only if the improper integral[6]

$$\int_1^{\infty} f(x)\ dx$$

does.

Let's put this theorem to work. The harmonic series is the case $p = 1$ of the so-called **p-series**

$$\sum_{n=1}^{\infty}\frac{1}{n^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} + \cdots,$$

where $p > 0$. Let's figure out for which values of p this series converges and for which it diverges.

We already know that the series diverges for $p = 1$. Since $p > 0$, the function $f(x) = 1/x^p$ is decreasing for positive x and therefore satisfies the condition of our integral test. Now

$$\int_1^{\infty}\frac{1}{x^p}\ dx = \lim_{b\to\infty}\int_1^{b}\frac{1}{x^p}\ dx = \lim_{b\to\infty}\left[-\frac{1}{(p-1)x^{p-1}}\right]_1^b = \lim_{b\to\infty}\frac{1}{p-1}\left(1-\frac{1}{b^{p-1}}\right)$$

$$= \frac{1}{p-1}\left(1-\frac{1}{\lim\limits_{b\to\infty} b^{p-1}}\right)$$

We can see at a glance that for $p > 1$ both the integral and the series converge and that for $0 < p < 1$ both diverge. And just because p-series are so important let's say this again.

[6]The integral is called "improper" because its upper limit is ∞. We'll consider improper integrals in detail in Chapter 3.2.2. Also, at this stage just take my word for it that you can actually always integrate a decreasing function like the one this theorem is all about.

Theorem 2.2.3 (p-series) *A p-series converges if $p > 1$ and diverges if $0 < p < 1$.*

So, for example,

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

converges $(p = 2 > 1)$, whereas

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} = 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} + \cdots$$

diverges $(p = \frac{1}{2} \leq 1)$.

We can define a complex-valued function by replacing the parameter p in the p-series $\sum 1/n^p$ with complex variable z. The resulting function is one of the most important functions in mathematics, the so-called **Riemann zeta function**

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z} = 1 + \frac{1}{2^z} + \frac{1}{3^z} + \frac{1}{4^z} + \cdots .$$

In fact, the most famous unsolved conjecture in mathematics is the **Riemann Hypothesis**: It says that the only complex zeros of the Riemann zeta function have real part $\frac{1}{2}$. I have to admit that this does not sound very impressive at first glance. However, if this hypothesis turned out to be true it would have all sorts of important implications, especially in terms of the distribution of the prime numbers.

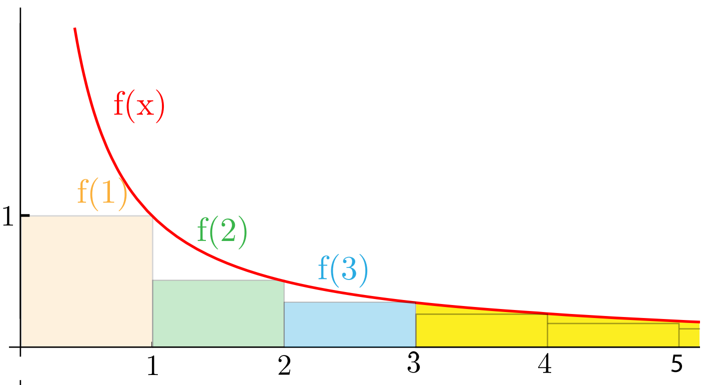
2.2.4 Integral test remainder estimate

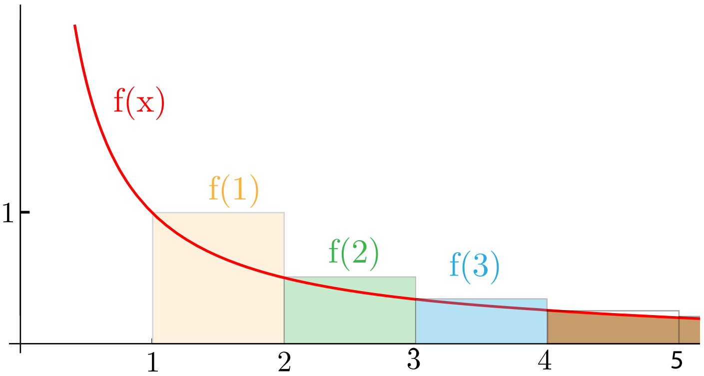
There is some more useful information that we can extract from the considerations in the previous section.

We just showed that

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

converges to a finite number S (the sum). Then what is this number approximately? Let's have a look at the following diagram.





Clearly,

$$S_3 + \text{the brown area in 2nd diagram} < S < S_3 + \text{the yellow area in 1st diagram}$$

Now expressing the yellow and brown areas as integrals we get

$$S_3 + \int_{3+1}^{\infty} \frac{1}{x^2}\, dx < S < S_3 + \int_{3}^{\infty} \frac{1}{x^2}\, dx.$$

After evaluating these (simple) integrals this chain of inequalities turns into

$$S_3 + \frac{1}{4} < S < S_3 + \frac{1}{3}.$$

Replacing S_3 by S_n we get

$$S_n + \frac{1}{n+1} < S < S_n + \frac{1}{n}.$$

Using a calculator we add up the first 50 terms of the series and find

$$S_{50} \approx 1.625132734.$$

Then our string of inequalities tells us that

$$1.64474057 < S < 1.64513274.$$

It turns out that the sum of our series is $\pi^2/6$, that is,

$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \cdots$$

This beautiful identity relates π to the natural numbers and is one of the gems of calculus. We'll prove that $\pi^2/6$ is really the sum of our series in one of the following chapters. In the meantime our considerations suggest a way in which we can get better and better approximations of π by adding more and more terms of our series.

For example, multiplying

$$1.64474057 < \pi^2/6 < 1.64513274.$$

by 6 and then finding the square roots of all numbers in sight gives

$$3.1414078 < \pi < 3.1417823.$$

It is also useful to transform the string of inequalities

$$S_n + \frac{1}{n+1} < S < S_n + \frac{1}{n}.$$

by subtracting S_n. This gives

$$\frac{1}{n+1} < S - S_n < \frac{1}{n}.$$

So, in the case of $n = 50$ this says that the difference between the sum and the 50-th partial sum is a number between 0.0196 and 0.02. By following the above steps you should be able to find approximations to other sums of series to which the integral test can be applied.

Just for the sake of completeness, here is the general form of the last string of inequalities written up in the form of a theorem.

Theorem 2.2.4 (The integral test remainder estimate) *Suppose that the infinite series $\sum_{n=1}^{\infty} a_n$ and the improper integral $\int_1^{\infty} f(x)dx$ satisfy the hypotheses of the integral test, and suppose in addition that both converge. Then*

$$\int_{n+1}^{\infty} f(x)\ dx \le S - S_n \le \int_n^{\infty} f(x)\ dx.$$

Euler showed that if p is an even number then the sum of the p-series is a rational multiple of π^n. For example,

$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \cdots$$

$$\frac{\pi^4}{90} = 1 + \frac{1}{2^4} + \frac{1}{3^4} + \frac{1}{4^4} + \frac{1}{5^4} + \cdots.$$

$$\frac{\pi^6}{945} = 1 + \frac{1}{2^6} + \frac{1}{3^6} + \frac{1}{4^6} + \frac{1}{5^6} + \cdots.$$

$$\frac{\pi^8}{9450} = 1 + \frac{1}{2^8} + \frac{1}{3^8} + \frac{1}{4^8} + \frac{1}{5^8} + \cdots.$$

Very pretty/mysterious/amazing. Strangely enough we hardly know anything about the nature of the sums of the p-series for odd p.

2.3 Some more basic facts about series

The following is often useful when it comes to showing that a given series does *not* converge.

Theorem 2.3.1 (The nth-term test for divergence) *If*

$$\lim_{n\to\infty} a_n \neq 0,$$

then the infinite series $\sum a_n$ diverges.

In fact, if someone gives you a series and asks you whether it converges or not, the first thing you should always do is to check whether the nth term of the series goes to 0. If it doesn't the series diverges.[7]

It is VERY important to note that just because the nth term of a series converges to 0 does not imply that the series converges. For example, the nth term of the harmonic series does converge to 0. On the other hand, we have already shown that the harmonic series diverges.

Example 2.3.1 *The series*

$$\sum_{n=1}^{\infty}(-1)^{n-1}n^2 = 1 - 4 + 9 - 16 + 25 - \cdots$$

diverges because $\lim_{n\to\infty} a_n$ does not exist, whereas the series

$$\sum_{n=1}^{\infty}\frac{n}{3n+1} = \frac{1}{4} + \frac{2}{7} + \frac{3}{10} + \frac{4}{13} + \cdots$$

diverges because

$$\lim_{n\to\infty}\frac{n}{3n+1} = \frac{1}{3} \neq 0.$$

The following theorem is really completely obvious. Still, it is very important to highlight it.

[7] *Proof for Theorem 2.3.1.* If, as usual S_n stands for the nth partial sum of the series,

$$a_n = S_n - S_{n-1}.$$

But, if the series converges to S, then

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty}(S_n - S_{n-1}) = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = S - S = 0. \quad \blacksquare$$

Theorem 2.3.2 (It's the tail that counts) *Infinite series that are eventually the same either both converge or both diverge.*

This just says that we can add, delete or change a *finite* number of terms in a infinite series series without altering its convergence or divergence. Of course, the *sum* of a convergent series will usually be changed by such modifications.

Theorem 2.3.3 (Termwise addition and multiplication) *If* $\sum a_n = A$ *and* $\sum b_n = B$ *and* c *is a constant, then*

$$\sum (a_n + b_n) = A + B$$

and

$$\sum c a_n = cA.$$

Note that these rules also hold with obvious provisos if one or both of the sums A and B are equal to $+\infty$ or $-\infty$.

Earlier we compared infinite series to integrals. Now we'll compare series to each other.

Definition 2.3.1 *A* **positive-term series** $\sum b_n$ **dominates** *another series* $\sum a_n$ *if, from some* n *onward,* $a_n \leq b_n$.

Theorem 2.3.4 (Comparison test) *Suppose that* $\sum a_n$ *and* $\sum b_n$ *are* **positive-term series** *and that the second sum dominates the first one.*

1. *If* $\sum b_n$ *converges, then* $\sum a_n$ *converges as well.*
2. *If* $\sum a_n$ *diverges, then* $\sum b_n$ *diverges as well.*

Example 2.3.2 *We have*

$$\frac{1}{n(n+1)(n+2)} < \frac{1}{n^3}$$

for all $n \geqq 1$. *Therefore the series*

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

is dominated by the series $\sum 1/n^3$. *Since the second series is a convergent p-series we can apply the first part of our comparison test to conclude that* $\sum 1/[n(n+1)(n+2)]$ *converges as well.*

Example 2.3.3 *Because*

$$\frac{1}{\sqrt{2n}} < \frac{1}{\sqrt{2n-1}}$$

for all $n \geqq 1$, the positive-term series

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{2n-1}} = 1 + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{7}} + \cdots$$

dominates the series

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{2n}} = \frac{1}{\sqrt{2}} \sum_{n=1}^{\infty} \frac{1}{n^{1/2}}.$$

But $\sum 1/n^{1/2}$ is a divergent p-series, and a constant nonzero multiple of a divergent series diverges. Therefore using part 2 of our comparison test we conclude that the series $\sum 1/\sqrt{2n-1}$ diverges as well.

Proof of the last Theorem. If $\sum b_n$ converges, then its sum is an upper bound of $\sum a_n$. But the partial sums of a positive-term series $\sum a_n$ form an increasing sequence. By Theorem 2.1.5 this implies that $\sum a_n$ converges.

If $\sum a_n$ is a positive-term series it either converges or diverges to infinity. If it diverges it is clear that $\sum b_n$ will also diverge to infinity. ∎

2.4 Representing functions by infinite series

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots .$$

This is just the sum of the geometric series with $a = 1$ and $r = x$. Since the geometric series converges for $|x| < 1$, this means that the function $f(x) = 1/(1-x)$ can be written as an infinite series depending on the variable x in the interval $(-1, 1)$.

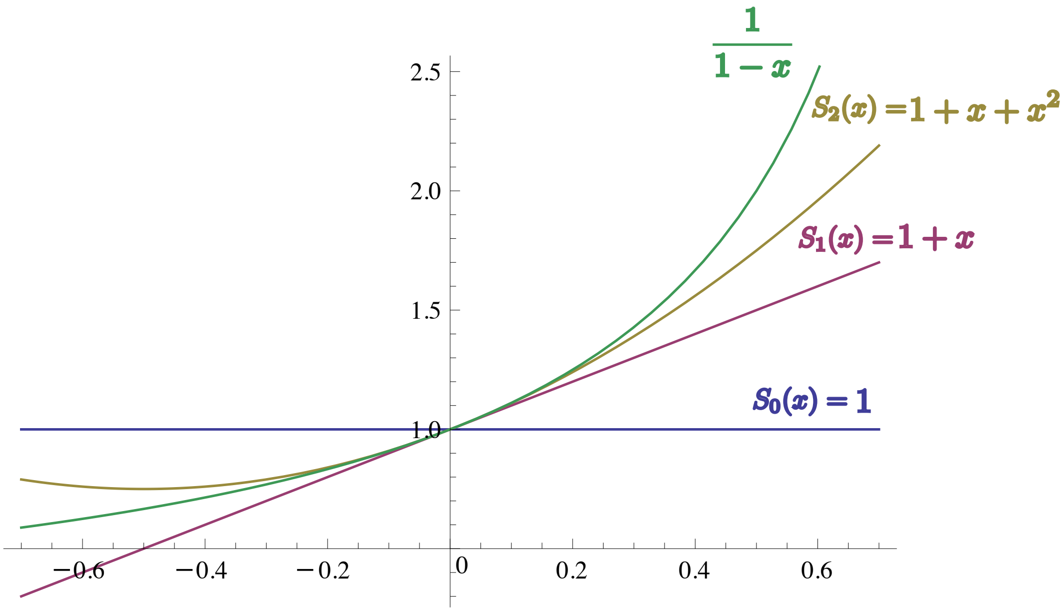
The nth partial sum

$$S_n(x) = 1 + x + x^2 + x^3 + \cdots + x^n$$

of the infinite series is an nth-degree **polynomial** that approximates the function $f(x) = 1/(1-x)$. Note that, unlike in previous sections, in the following we'll call the first element of the sequence of partial sums $S_0 = 1$ and refer to it as the 0th partial sum. The convergence of the infinite series for $|x| < 1$ suggests that the approximation

$$\frac{1}{1-x} \approx 1 + x + x^2 + x^3 + \cdots + x^n$$

should get arbitrarily close if we choose n sufficiently large. The following diagram shows the graphs of $1/(1-x)$ and the three approximations $S_0(x) = 1, S_1(x) = 1+x$, and $S_2(x) = 1+x+x^2$. It appears that, as can be expected, the approximations are more accurate when n is larger and when x is closer to zero.



The definitions of many of our favourite functions do not tell us how to compute their values precisely, except at a few isolated points. For example, by definition, $\ln(1) = 0$, but it is not clear how to calculate $\ln(2)$. Even such a simple function as $\sqrt{x}$ is not computable (precisely and in a finite number of steps) unless x happens to be the square of a rational number.

However any value of a polynomial such as

$$5 + 56x - 3x^2 + 67x^3 + 7x^4$$

is easy to calculate as only addition and multiplication are required. One goal of this section is to use the fact that values of polynomials are so easy compute to help us calculate approximate values of basic functions such as $\ln(x)$, e^x, or $\sin(x)$.

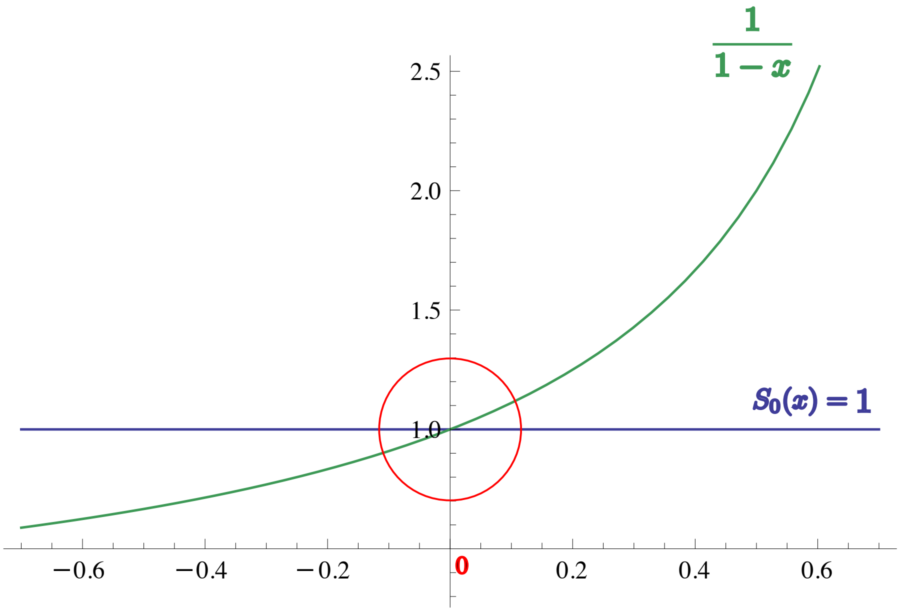
To get an idea as to what we should be looking for, let's have another close look at the partial sums of our motivating example

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots .$$

For the following discussion let's set $f(x) = 1/(1-x)$. Just looking at the diagram suggests a pattern. Let's see how long it takes you to spot it:

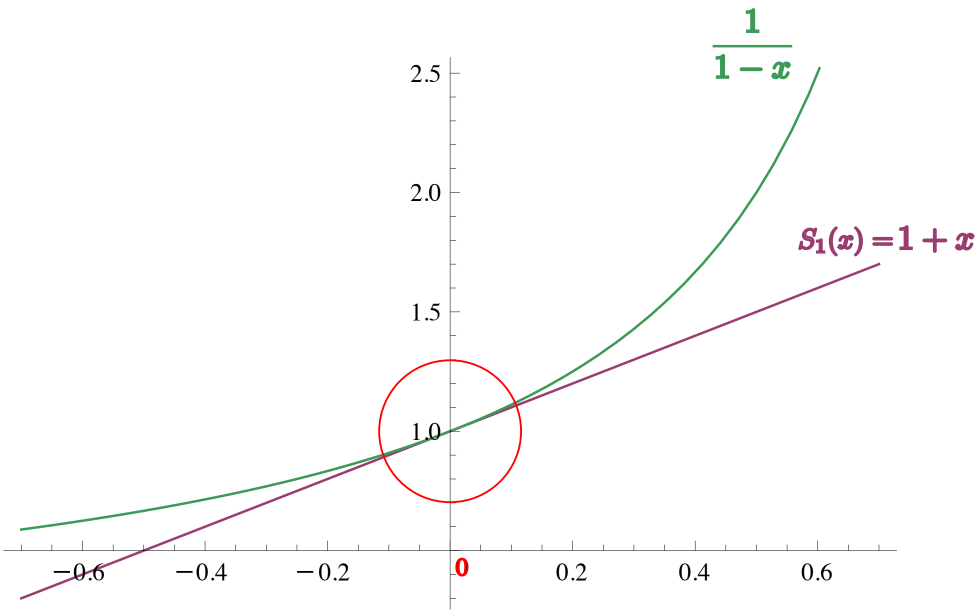
The value of the 0th partial sum at 0 is equal to that of $f(x)$, that is,

$$S_0(0) = f(0).$$



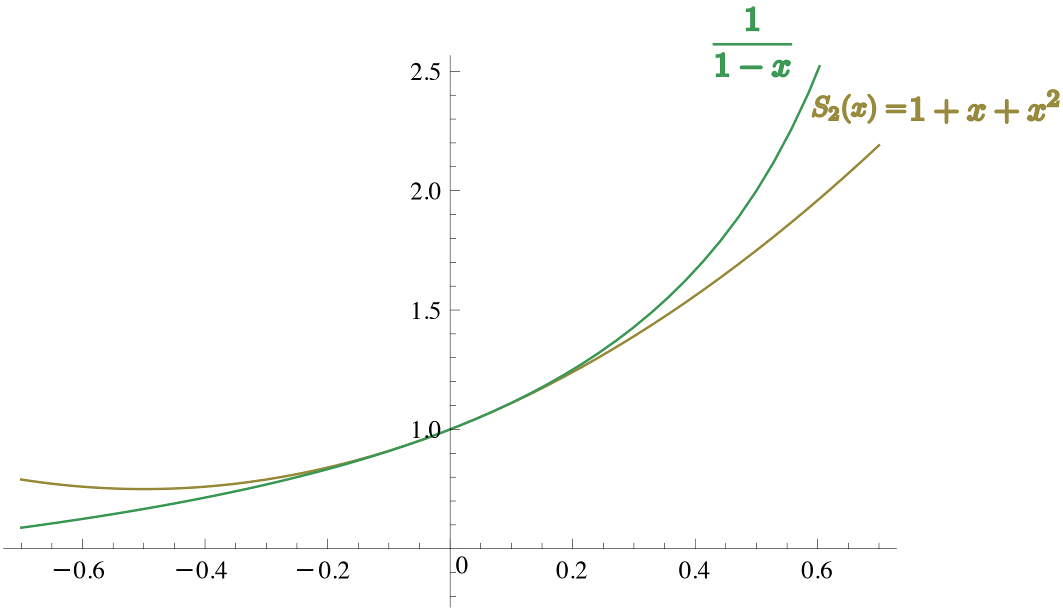
The value of the first partial sum $1 + x$ at 0 is equal to that of $f(x)$ AND the value of the first derivative of the first partial sum at 0 is equal to that of $f(x)$, that is,

$$S_1(0) = f(0) \text{ AND } S_1'(0) = f'(0).$$



For the 2nd partial sum $1 + x + x^2$ we find that that its value, its first derivative and its second derivative at 0 coincide with those of $f(x)$.

$$S_2(0) = f(0) \text{ AND } S_2'(0) = f'(0) \text{ AND } S_2''(0) = f''(0).$$



In general, we find that the nth partial sum of

$$1 + x + x^2 + x^3 + \cdots$$

is a polynomial of degree n whose value as well as its 1st, 2nd, 3rd, ..., nth derivatives at 0 coincide with those of $f(x) = 1/(1-x)$.

Now it turns out that, given any other function $f(x)$ that is infinitely often differentiable at 0, we can find an infinite series depending on the variable x whose partial sums have the same property.

Theorem 2.4.1 (The Maclaurin series of a function) *Let $f(x)$ be an infinitely often differentiable function at 0. Then the* **Maclaurin series** *of the function is*

$$f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots .$$

The Maclaurin series has the property that for all n the value of its nth partial sum $S_n(x)$ at $x = 0$ and the value of the first n derivatives of $S_n(x)$ at $x = 0$ agree with the values of $f(x)$ and its first n derivatives there.

Proof. Let's just check for

$$S_2(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2$$

that its value as well as the values of its first and second derivative coincide with those of $f(x)$ at $x = 0$. So,

$$S_2(0) = f(0) + f'(0)0 + \frac{f''(0)}{2!}0^2 = f(0)$$

Tick, its value at 0 is the same as that of $f(x)$. Now,

$$S_2'(x) = f'(0) + \frac{f''(0)}{1!}x,$$

and so

$$S_2'(0) = f'(0) + \frac{f''(0)}{1!}0 = f'(0).$$

Tick. And, finally,

$$S_2''(x) = f''(0),$$

hence tick again: $S_2''(0) = f''(0)$ At this point it should be clear that the nth partial sum $S_n(x)$ would get n ticks in the same way. ■

For example, in the case of $f(x) = e^x$ we know that all the derivatives of $f(x)$ are equal to e^x. Hence $f^{(n)}(0) = e^0 = 1$ and so the Maclaurin series of e^x is

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots .$$

Does this mean that we can conclude that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots$$

and that therefore for fixed x the partial sums of this series are suited to approximate e^x arbitrarily well?

It turns out that the above identity is true for all $x \in \mathbf{R}$. However, at this point of our discussion the only thing that is certain is that this identity holds true for $x = 0$. What could go wrong? Well, it could be that the series does not converge for any x other than 0, or it could converge only for some values of x and, who knows, maybe even for those values that it does converge the sums have nothing to do with e^x.

Anyway, for the moment just take my word for it that e^x really is equal to its Maclaurin series. In fact, let's just just derive the Maclaurin series of $\sin(x)$ and $\cos(x)$ that also "work".

Let's start with $f(x) = \cos x$. We first calculate the derivatives

$$f(x) = \cos(x), \text{ hence } f(0) = 1$$

$$f'(x) = -\sin(x), \text{ hence } f'(0) = 0$$

$$f''(x) = -\cos(x), \text{ hence } f''(0) = -1$$

$$f^{(3)}(x) = \sin(x), \text{ hence } f^{(3)}(0) = 0$$

$$f^{(4)}(x) = \cos(x), \text{ hence } f^{(4)}(0) = 1$$

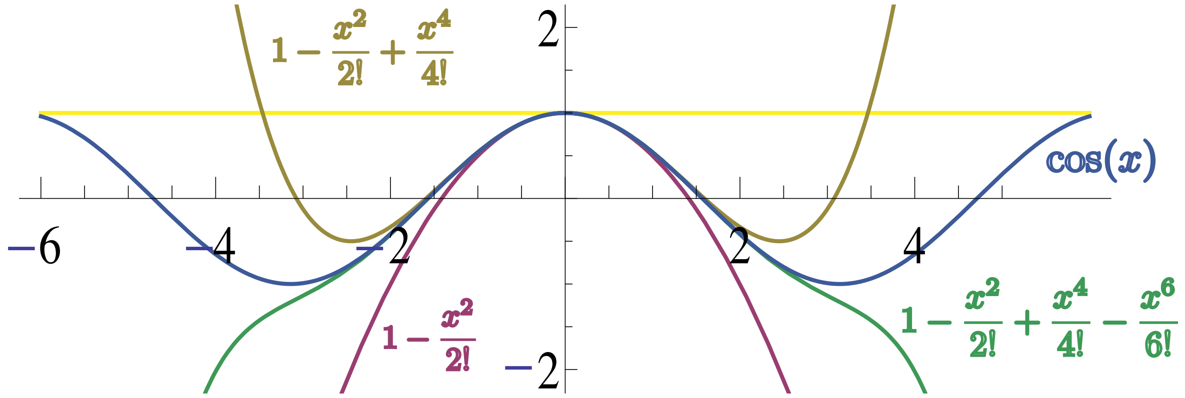
Things start repeating at this point and we conclude that the Maclaurin series of $\cos(x)$ is

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots .$$

Similarly, we find that the Maclaurin series of $\sin(x)$ is

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots .$$

Here is a diagram of $\cos(x)$ and the first couple of partial sums of its Maclaurin series. Definitely, looks like a pretty nice fit, doesn't it?



Of course, this pictures does not prove anything, but it certainly looks like we are on to something. Here is a summary of what we just discussed.

Theorem 2.4.2 (The Maclaurin series of $e^x, \sin(x), \cos(x)$) *We have*

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots ,$$

where $x \in \mathbf{R}$.

We'll indicate how this result is proved in the next section.

2.5 Taylor series

In our discussion of the Maclaurin series the value $x = 0$ plays a central role. Of course, there is nothing really special about this value as far as a function is concerned. In fact, any other value can take its role in this discussion and we also have the following theorem.

Theorem 2.5.1 (Taylor series) *Let $f(x)$ be an infinitely often differentiable function at a. Then the* **Taylor series** *of the function at a is*

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$$

The Taylor series at a has the property that for all n the value of its nth partial sum $S_n(x)$ at $x = a$ and the value of the first n derivatives of $S_n(x)$ at $x = a$ agree with the values of $f(x)$ and its first n derivatives there. The nth partial sum of the Taylor series at a is also called the **nth degree Taylor polynomial of $f(x)$ at** *a.*

The Maclaurin series of a function is simply the Taylor series of a function at 0.

As an example, let's find the Taylor series of $f(x) = \ln(x)$ at 1. The first few derivatives of $f(x)$ are

$$f(x) = \ln(x), \text{ hence } f(1) = 0$$

$$f'(x) = \frac{1}{x}, \text{ hence } f'(1) = 1$$

$$f''(x) = -\frac{1}{x^2}, \text{ hence } f''(1) = -1$$

$$f^{(3)}(x) = \frac{2}{x^3}, \text{ hence } f^{(3)}(1) = 2!$$

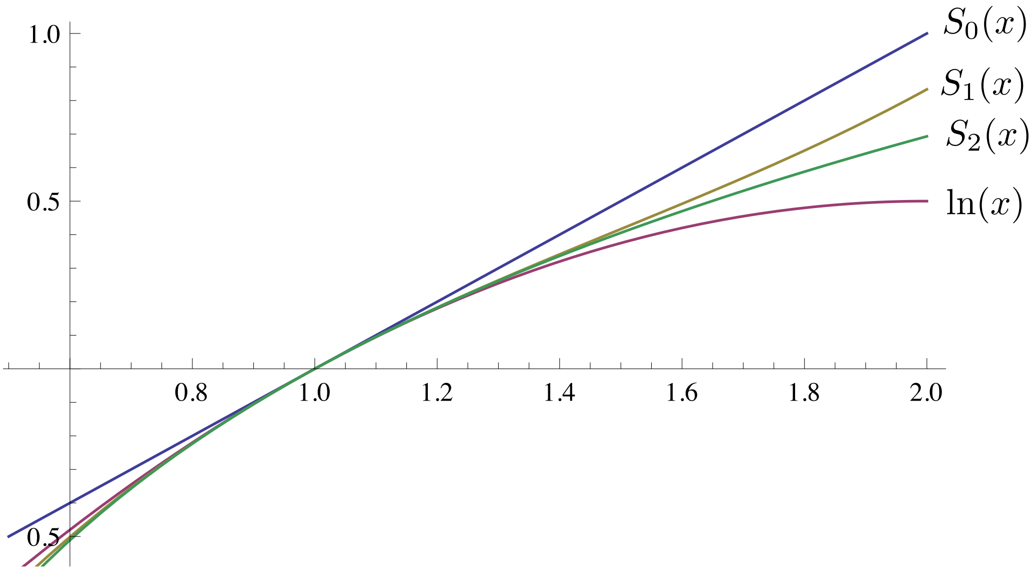
$$f^{(4)}(x) = -\frac{3!}{x^4}, \text{ hence } f^{(4)}(1) = -3!$$

$$f^{(5)}(x) = \frac{4!}{x^5}, \text{ hence } f^{(5)}(1) = 4!$$

The pattern is clear. This means that the Taylor series of $\ln(x)$ at 1 is

$$(x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 + \cdots + \frac{(-1)^{n-1}}{n}(x-1)^n + \cdots$$

Here is a diagram of $\ln(x)$ and the first couple of partial sums of its Taylor series at 1.



Definitely, looks like a pretty nice fit, doesn't it? In fact, in the next section we'll find that $\ln(x)$ equals this Taylor series for all $0 < x \leq 2$. If we accept this as a fact and substitute $x = 2$, we arrive at the following remarkable identity.

$$\ln(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} + \cdots .$$

2.6 Biting the bullet: Taylor's formula (MTH1035 and MTH1040 only)

One of our main messages so far is the following: Starting with a function $f(x)$ that is sufficiently "nice" we often find that the polynomial

$$S_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

is a good approximation to $f(x)$ near a. That's a good starting point, but to be of any practical use we need some tools that allows us to measure how good an approximation we are talking about. The main tool for doing this is called **Taylor's formula**. Now, before we start on this very important topic I should warn you that things are about to get quite technical and not for the faint of heart. So, buckle your mathematical seatbelts and let's give this our best shot.

How good the polynomial $S_n(x)$ approximates the function $f(x)$ is measured by the difference

$$R_n(x) = f(x) - S_n(x).$$

This difference $R_n(x)$ is called the **nth-degree remainder** of $f(x)$ at $x = a$. It captures the error we make if the value $f(x)$ is replaced by $S_n(x)$.

This means that

$$f(x) = S_n(x) + R_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x).$$

Let's evaluate all this at a specific point b. So,

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(a)}{2!}(b-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(b-a)^n + R_n(b).$$

Theorem 2.6.1 (Taylor's formula) *Let $f(x)$ be a function that is $n+1$ times differentiable in an interval containing a and b. Then*

$$R_n(b) = \frac{f^{(n+1)}(z)}{(n+1)!}(b-a)^{n+1}$$

for some number z between a and b.

We won't prove this theorem, but we'll try to get a feel for what it does by pondering the simplest case and applying this result a few times. In particular, we'll show how this result can be used to prove that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots .$$

Taylor's theorem: the simplest case = the mean value theorem

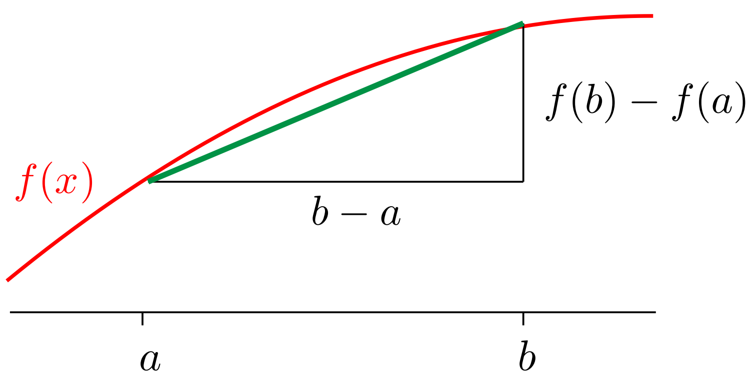
Let's first have a look at the simplest case $n = 0$. Then this theorem asserts that

$$f(b) = f(a) + f'(z)(b-a)$$

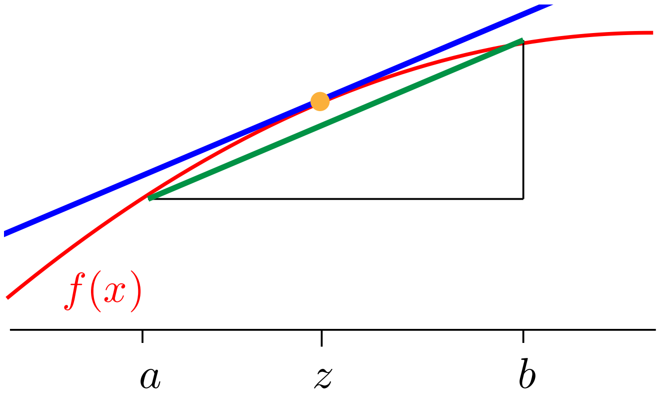
for some z in the interval between a and b. To get a feel for why this should be the case we first rewrite this equation as

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

The left side of this equation is just the slope of the green line segment in the following diagram



And now the theorem says that somewhere in the interval between a and b we can find a point z such that at this point $f'(z)$, the slope of the tangent to the red curve is equal to the slope of the green segment. In this diagram it is clear that and where this happens.



This simplest case of Taylor's formula has a special name. It is called the **mean value theorem** and is one of the most fundamental theorems of calculus. After pondering the above diagrams it may seem pretty obvious to you, but remember, there are a lot of subtleties hiding in words like "continuous" or "differentiable" that we are glossing over in this unit and that need to be addressed very carefully when we are trying to build a sound foundation for calculus. The mean value theorem is one of the cornerstones for such a theory.

Anyway, in a unit solely dedicated to calculus you would first come across the mean value theorem and then much later prove Taylor's formula. Given that this is the natural order of introducing these theorems, Taylor's formula is usually considered to be a generalization of the mean value theorem.

2.6.1 Applying Taylor's formula

Now let's apply Taylor's formula. First, we'll have another look at the simplest case $n = 0$. Again, it says that

$$f(b) = f(a) + f'(z)(b-a)$$

for some z in the interval between a and b. The way we would apply this result is to get an idea of the range of $f'(z)$ in the interval. For example, one possible conclusion could be that if $f'(z)$ has a small absolute value everywhere in the interval then the blue error would also be small.

Now let's consider a concrete example which illustrates how one applies Taylor's formula. Here is the third partial sum

$$S_3 = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}$$

of the Maclaurin series of e^x. So, Taylor's formula tells us that

$$e^x - \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}\right) = \frac{e^z x^4}{4!},$$

and therefore

$$\left| e^x - \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}\right)\right| = \left|\frac{e^z x^4}{4!}\right| = \frac{e^z x^4}{4!},$$

for some z in the interval between 0 and x. We are interested in estimating how large the expression on the right side can be. If x is positive, then in our interval e^z is maximal at the far right of the interval, that is, at $z = x$. Therefore,

$$\frac{e^z x^4}{4!} \leq \frac{e^x x^4}{4!}.$$

So, for example, for $x = 1$ this estimate tells us that the difference is less than

$$\frac{e}{24} = 0.11326....$$

Let's check this on a calculator:

$$e = 2.71828...$$

and

$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} = 2.666666...$$

and

$$\left| e - \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!}\right)\right| = 0.051615....$$

which is indeed less than the predicted 0.11326....

So far so good. But there is more. If you have a close look you'll notice that the same arguments yield that in general

$$\left| e - \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!}\right)\right| < \frac{e}{(n+1)!}.$$

As n goes to infinity the right side goes to 0. This implies that

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots .$$

or, in other words,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

is true for $x = 1$.

In fact, it's not terribly hard to refine our argument to show that this is true for all x, which then proves the above famous identity.[8]

In fact, most proofs of functions being equal to one of their Taylor series use Taylor's formula.

As a second example, let's ponder the Taylor series for $\ln(x)$ at 1 which we derived on page 41

$$(x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 + \cdots + \frac{(-1)^{n-1}}{n}(x-1)^n + \cdots .$$

Applying Taylor's formula, we get

$$\left| \ln(x) - \left((x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \cdots + \frac{(-1)^{n-1}}{n}(x-1)^n \right) \right| = \left| \frac{f^{(n+1)}(z)}{(n+1)!}(x-1)^{n+1} \right|$$

In this case we have

$$|f^{(n+1)}(z)| = \frac{n!}{z^{n+1}},$$

and so

$$\left| \frac{f^{(n+1)}(z)}{(n+1)!}(x-1)^{n+1} \right| = \frac{n!}{z^{n+1}(n+1)!}|x-1|^{n+1} = \frac{1}{z^{n+1}(n+1)}|x-1|^{n+1}.$$

We are first interested in finding the z in the interval $[1,x], x>1$ at which the expression on the right is maximal. This is clearly the case for $z=1$. This means that the expression on the right is less than

$$\frac{1}{(n+1)}(x-1)^{n+1}.$$

For fixed $1 < x \leq 2$ this expression will go to 0 as n goes to infinity, which then shows that

$$\ln(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \cdots$$

for $1 < x \leq 2$. To show the same for $0 < x < 1$ is also not difficult. Finally, it turns out that the Taylor series diverges for $x > 2$. So, in summary, the above identity holds for $0 < x \leq 2$ and does not for $x > 2$.

[8] For positive x we find that

$$\left| e^x - \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{1}{n!}\right)\right| \leq \frac{e^x x^{n+1}}{(n+1)!}.$$

For negative x we find that

$$\left| e^x - \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{1}{n!}\right)\right| \leq \frac{|x^{n+1}|}{(n+1)!}.$$

Then the fact that $\lim\limits_{x\to\infty} \dfrac{x^n}{n!} = 0$ for all x is all that is needed to see at a glance that the right sides of these inequalities tend to 0 are n goes to infinity.

While we are here, let's write down the error estimate for that particularly interesting value of $x = 2$.

$$\left| \ln(2) - \left(1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{(-1)^{n-1}}{n}\right) \right| < \frac{1}{(n+1)}.$$

So, for example, for $n = 10$ this says that we can be sure the difference between $\ln(2)$ and

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots - \frac{1}{10}$$

is definitely less than $1/10$. Let's check this

$$\ln(2) = 0.69314718...$$

and

$$1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{1}{10} = 0.645635...$$

So the difference is 0.0475123 which is less then 0.1.

Note, that when I am calculating $\ln(2) = 0.69314718...$ I am using a calculator. How does this calculator come up with its answer? Well, its answer will most likely be based on some partial sum of a series and the point at which this power series gets chopped off was determined by the manufacturers of my calculator using Taylor's theorem. So, some of what we've been doing above is really another instance of a snake biting its own tail.

2.7 A bit of bad news

So far, whenever we looked at the Taylor series of a function, we found that the series is actually equal to the function, at least in some interval. However, this is definitely not always the case. In fact, given any function $f(x)$ and its associated Taylor series at a point a, it is fairly easy to see that there are infinitely many functions that have the SAME Taylor series at $x = a$ AND take on different values except at that point. Obviously, at most one of all these infinitely many functions can be equal to the Taylor series in some interval containing $x = a$.

In particular, there are infinitely many functions that have the Maclaurin series

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

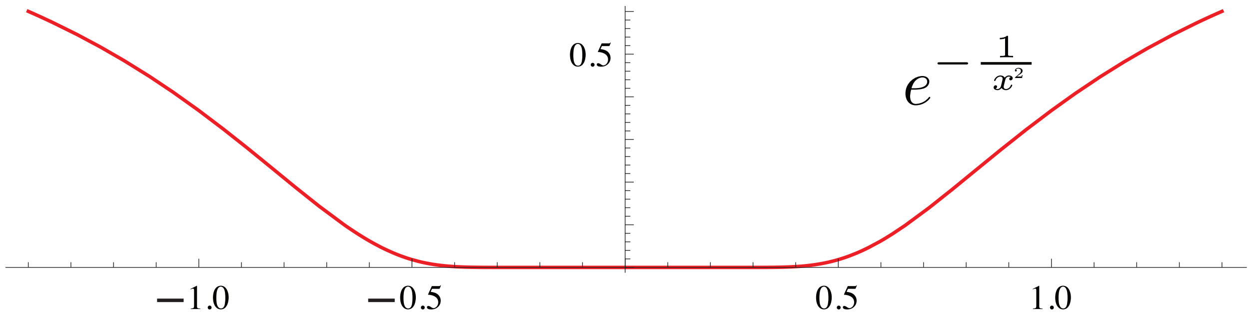
and that take on different values for all $x \neq 0$. So, at first glance we really seem to be very lucky that among all these functions the one we are interested in is e^x.

Want to know more? Okay, let's have a look at the function

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

It turns out that this function is infinitely often differentiable and that its value and all its derivatives are equal to 0 at $x = 0$. This means that just like the zero function its Maclaurin series vanishes. This also means that all its infinitely many multiples also have vanishing Maclaurin series.

Now here is a nice trick: Simply add all these special functions to e^x and you get infinitely many functions that have the same Maclaurin series as e^x and at the same time take on different values for all $x \neq 0$. .



2.8 Further applications of Taylor series

Apart from their use in approximating functions by polynomials Taylor series have all sorts of other important applications. In this section we'll highlight a few of these applications.

2.8.1 The number e

First up is a very pretty proof of the fact that e is an irrational number.

Theorem 2.8.1 (e is irrational) *The base of the natural logarithm e is an irrational number, that is, e cannot be written as a fraction of (positive) integers a/b.*

Proof. Here is a **proof by contradiction**. Let's assume that we can write e as a fraction. This means that there are integers a and b such that

$$\frac{a}{b} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots .$$

Let's multiply this identity by $b!$. This gives

$$a(b-1)! = b! + b! + \frac{b!}{2!} + \frac{b!}{3!} + \frac{b!}{4!} + \cdots + \frac{b!}{b!} + \frac{b!}{(b+1)!} + \frac{b!}{(b+2)!} + \cdots .$$

It is clear that both the green expression and the red sum are integers. This implies that the infinite sum

$$\frac{b!}{(b+1)!} + \frac{b!}{(b+2)!} + \cdots (= \textit{green} - \textit{red})$$

is also an integer. However,

$$\begin{aligned}
\frac{b!}{(b+1)!} + \frac{b!}{(b+2)!} + \cdots &= \frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \frac{1}{(b+1)(b+2)(b+3)} + \cdots \\
&\leq \frac{1}{b+1} + \frac{1}{(b+1)^2} + \frac{1}{(b+1)^3} + \frac{1}{(b+1)^4} + \cdots \\
&= \frac{1}{b+1}\left(1 + \frac{1}{b+1} + \frac{1}{(b+1)^2} + \frac{1}{(b+1)^3} + \cdots\right) \\
&= \frac{1}{b+1}\left(\frac{1}{1-\frac{1}{b+1}}\right) \quad \text{(here we've used the geometric sum)} \\
&= \frac{1}{b+1}\frac{b+1}{b} \\
&= \frac{1}{b}.
\end{aligned}$$

Now, since e is definitely not an integer the denominator b of our fraction has to be greater than 2. This implies that $1/b$ is less than 1 and therefore NOT an integer. What this means is that, starting with our assumption that e is a rational number, we can deduce that a certain number is an integer and at the same time not an integer. Since this is not possible our original assumption must be wrong and we conclude that e is not a rational number. ■

2.8.2 New Taylor series from old ones

Note that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

is an identity that holds for all values of x. Consequently, new series can be derived by substitution. For example, substituting $-x^2$ for x gives

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \cdots + (-1)^n \frac{x^{2n}}{n!} + \cdots$$

This is a new identity which holds for all values of x. In fact, the series on the right turns out to be the Maclaurin series of e^{-x^2} which would have been very cumbersome to derive using the definition.

2.8.3 Extending real valued functions to complex valued functions: Euler's formula

The sum of an infinite series $\sum c_n$ with complex terms $c_n = a_n + ib_n$ is defined by

$$\sum c_n = \sum a_n + i \sum b_n$$

provided that the two infinite series of real numbers on the right-hand side converge. In this case we say that the infinite series of complex numbers on the left-hand side converges. It can be shown that the exponential series

$$1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \cdots$$

converges for all complex numbers $z = x + iy$. Consequently, the exponential function e^z can be **defined** (for complex as well as for real arguments) via this series:

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \cdots$$

Substituting the number z by $i\theta$ (with θ real), we get

$$\begin{aligned} e^{i\theta} &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \cdots \\ &= 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} - \cdots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right). \end{aligned}$$

Here we've used that $i^2 = -1, i^3 = -i, i^4 = 1$, and so on. We spot the Maclaurin series for $\cos(\theta)$ and $\sin(\theta)$ on the right-hand side and conclude that

$$e^{i\theta} = \cos(\theta) + i\sin(\theta)$$

for every real number θ. This is **Euler's formula**, one of the most famous and useful identities in mathematics.[9]

Setting $\theta = \pi$ gives **Euler's identity**, $e^{i\pi} = \cos(\pi) + i\sin(\pi) = -1$, one of the most beautiful identities in mathematics. The most popular way of writing this identity is

$$e^{i\pi} + 1 = 0.$$

[9] If you replace θ by y and multiply the resulting identity by e^x where x just like y stands for a real number you arrive at another way to define the complex exponential function:

$$e^{x+iy} = e^x(\cos(y) + i\sin(y)).$$

In general, if a Taylor series of a real function is equal to the function in some interval, it is possible to use this identity to extend the function to a complex function. The resulting complex functions have many desirable properties and form the core of the beautiful and very important theory of complex analytic functions.

Let's just point out one more really neat application of Euler's formula, a proof of **de Moivre's formula**,

$$(\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx)$$

which is valid for any real number x and integer n. The proof is a one-liner in which we apply Euler's formula twice:

$$(\cos(x) + i\sin(x))^n = (e^{ix})^n = e^{i(nx)} = \cos(nx) + i\sin(nx).$$

And, of course, for particular values of n de Moivre's formula implies all sorts of important trigonometric identities. For example, setting $n = 2$ we get

$$(\cos(x) + i\sin(x))^2 = (\cos^2(x) - \sin^2(x)) + i\ 2\sin(x)\cos(x) = \cos(2x) + i\sin(2x),$$

and therefore

$$\cos^2(x) - \sin^2(x) = \cos(2x)$$

and

$$2\sin(x)\cos(x) = \sin(2x).$$

2.8.4 The Number π: Life beyond the circle

We begin by massaging

$$\begin{aligned}
\frac{1}{1-x} &= 1 + x + x^2 + x^3 + \cdots \\
&= 1 + x + x^2 + x^3 + \cdots + x^n + x^{n+1}\left(1 + x + x^2 + x^3 + \cdots\right) \\
&= 1 + x + x^2 + x^3 + \cdots + x^n + x^{n+1}\frac{1}{1-x} \\
&= 1 + x + x^2 + x^3 + \cdots + x^n + \frac{x^{n+1}}{1-x}
\end{aligned}$$

We substitute x by $-t^2$. This gives

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - t^6 + \cdots + (-1)^n t^{2n} + \frac{(-1)^{n+1} t^{2n+2}}{1+t^2}.$$

Because $(\tan^{-1})'(t) = 1/(1+t^2)$, integrating both sides of this last equation from $t = 0$ to $t = x$ gives

$$\tan^{-1}(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + R_{2n+1},$$

where

$$|R_{2n+1}| = \left| \int_0^x \frac{t^{2n+2}}{1+t^2} dx \right| \leq \left| \int_0^x t^{2n+2} dx \right| = \frac{|x|^{2n+3}}{2n+3}.$$

We conclude that

$$\lim_{n\to\infty} R_n = 0$$

if $|x| \leq 1$. Hence we obtain that the identity

$$\tan^{-1}(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$$

is valid for $-1 \leq x \leq 1$.[10]

If we substitute $x = 1$ into this identity, we obtain **Leibniz's series**

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

which relates π to the odd natural numbers. Not a single circle in sight!

2.9 Absolute convergence

As we've seen, the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

converges to $\ln(2)$, but if we change all the minus signs in this series to plus signs, we get the divergent harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots$$

On the other hand, the geometric series

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \cdots = \frac{2}{3}$$

converges, as does its associated positive-term series

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2.$$

The next theorem asserts that IF a positive-term series converges, then we may change any subset of the terms into their negatives and the resulting sequence will also always converge.

[10]The right side of this identity turns out to be the Maclaurin series of $\tan^{-1}(x)$; see the remark in blue on page 59.

Theorem 2.9.1 (Absolute Convergence implies Convergence) *If the series $\sum |a_n|$ converges, then so does the series $\sum a_n$.*

Proof. Let's assume that the series $\sum |a_n|$ converges. We have

$$0 \leq a_n + |a_n| \leq 2|a_n|.$$

Then since $\sum |a_n|$ converges, $\sum 2|a_n|$ converges and since the positive term series $\sum 2|a_n|$ converges and dominates the positive terms series $\sum (a_n + |a_n|)$, that last series also converges. We conclude that

$$\sum a_n = \sum ((a_n + |a_n|) - |a_n|) = \sum (a_n + |a_n|) - \sum |a_n|$$

converges. ■

This means that if we are faced with an arbitrary sequence $\sum a_n$ we know that it converges if we can show that $\sum |a_n|$ does. This suggests the following definition.

Definition 2.9.1 (Absolute Convergence) *The series $\sum a_n$* **converges absolutely** *if the series*

$$\sum |a_n| = |a_1| + |a_2| + |a_3| + \cdots + |a_n| + \cdots$$

converges.

Returning to our introductory examples this means that

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \cdots = \frac{2}{3}$$

converges absolutely. On the other, hand

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

converges, but does not converge absolutely. A series that converges but does not converge absolutely is called **conditionally convergent**. So a series is absolutely convergent, conditionally convergent, or divergent.

For a positive-term series $\sum a_n$ we have seen how we can use comparisons with other positive-term sequences or integrals to establish convergence or divergence of $\sum a_n$. On the other hand, we have not encountered any generally applicable ways of establishing whether a given series $\sum b_n$ containing infinitely many positive and negative terms converges. So, faced with a series like this it seems like a good idea to first see what we can say about $\sum |b_n|$.

Example 2.9.1 *The series*

$$\sum_{n=1}^{\infty} \frac{\cos(n)}{n^2} = \cos(1) + \frac{\cos(2)}{4} + \frac{\cos(3)}{9} + \cdots$$

has a fairly crazy distribution of positive and negative terms. So let's consider

$$\sum_{n=1}^{\infty} \frac{|\cos n|}{n^2}$$

Since

$$0 \leq \frac{|\cos(n)|}{n^2} \leq \frac{1}{n^2}$$

we can see that the positive-term series $\sum_{n=1}^{\infty} \frac{|\cos(n)|}{n^2}$ is dominated by the convergent p-series

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

We conclude that the series we are interested in is absolutely convergent.

In general, a series S splits up into the series S^+ of its positive terms and the series S^- of its negative terms. These two series are either convergent or divergent to $(\pm)$ infinity and the following theorem holds.

Theorem 2.9.2 (The two parts of a series) *If both S^+ and S^- are convergent, then S is absolutely convergent. If only one of S^+ and S^- is convergent, then S is divergent. If both S^+ and S^- are divergent, then S is either divergent or conditionally convergent.*

2.10 The ratio test

Remember that a geometric series is a series of the form

$$\sum_{n=0}^{\infty} ar^n.$$

Then it is clear that the quotient of consecutive terms is equal to

$$\frac{ar^{n+1}}{ar^n} = r.$$

In other words the terms of a geometric series grow by a factor of r from term to term.

Here is an idea. If we are facing a sequence the growth of whose terms approaches a constant ρ then we should be able to establish its absolute convergence or divergence by comparing it with a suitable geometric series. And this is indeed the case.

Theorem 2.10.1 (The Ratio Test) *Let the terms a_n of a series be non-zero from a certain n onward and suppose that the limit*

$$\rho = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

either exists or is infinite. Then the infinite series $\sum a_n$

1. *converges absolutely if $\rho < 1$; or*
2. *diverges if $\rho > 1$.*

If $\rho = 1$, the ratio test does not tell us anything about the convergence or divergence of our series.

Proof. In this proof we implement our idea from the introduction. If $\rho < 1$, choose a (fixed) number r with $\rho < r < 1$. Then the limit in our theorem implies that there exists a natural number N such that $|a_{n+1}| \leq r|a_n|$ for all $n \geqq N$. It follows that

$$|a_{N+1}| \leq r|a_N|,$$

$$|a_{N+2}| \leq r|a_{N+1}| \leq r^2|a_N|,$$

$$|a_{N+3}| \leq r|a_{N+2}| \leq r^3|a_N|,$$

and in general that

$$|a_{N+k}| \leq r^k|a_N|$$

for $k \geqq 0$.

Hence the series

$$|a_N| + |a_{N+1}| + |a_{N+2}| + \cdots$$

is dominated by the geometric series

$$|a_N|(1 + r + r^2 + r^3 + \cdots),$$

and the geometric series converges because $|r| < 1$. We conclude that the series $\sum |a_n|$ converges, and that therefore the series $\sum a_n$ converges absolutely.

Divergence in the case $\rho > 1$ is established in a similar way. ■

To see that $\sum a_n$ may either converge or diverge if $\rho = 1$, consider the divergent series $\sum(1/n)$ and the convergent series $\sum(1/n^2)$. Convince yourself that for both series, the value of the ratio ρ is 1.

Example 2.10.1 *Here is a series to which the ratio test can be applied successfully.*

$$\sum_{n=1}^{\infty} \frac{(-1)^n 2^n}{n!} = -2 + \frac{4}{2!} - \frac{8}{3!} + \frac{16}{4!} - \cdots$$

Now

$$\rho = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \left| \frac{\frac{(-1)^{n+1}2^{n+1}}{(n+1)!}}{\frac{(-1)^n 2^n}{n!}} \right| = \lim_{n\to\infty} \frac{2}{n+1} = 0.$$

Because $\rho < 1$, the series converges absolutely.

Example 2.10.2 *Let's consider $\sum_{n=1}^{\infty} \frac{n}{2^n}$. Here*

$$\rho = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \frac{n+1}{2^{n+1}} \frac{2^n}{n} = \lim_{n\to\infty} \frac{n+1}{2n} = \frac{1}{2}.$$

We conclude that the series converges absolutely.

Example 2.10.3 *Let's consider*

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

Here

$$\rho = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \frac{3^{n+1}}{(n+1)^2} \frac{n^2}{3^n} = \lim_{n\to\infty} \frac{3n^2}{(n+1)^2} = 3.$$

Since $\rho > 1$, this series diverges.

2.11 Power series

Definition 2.11.1 *An infinite series*

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots$$

with the constant coefficients $a_0, a_1, a_2, \ldots$ is called a **power series***.*

Earlier we found that certain functions can be written as power series

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

In the following we investigate power series in general.

2.11.1 Convergence

To start with let's see what we can say in general about the values of x for which a power series converges.

Obviously, every power series converges for $x = 0$. Because of the way in which the powers of x are part of the nth term of a power series, the ratio test is particularly effective in determining the values of x for which many power series converge.

To be able to apply the ratio test it is necessary that the following limit exists

$$\lim_{n\to\infty} \left| \frac{a_{n+1}x^{n+1}}{a_n x^n} \right| = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| |x|.$$

Then the ratio test tells us that the series converges absolutely if

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| |x| < 1,$$

or

$$|x| < \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right|.$$

Similarly, it tells us that the series diverges if

$$|x| > \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right|.$$

In other words, if

$$R = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right|$$

exists, the series converges absolutely in the interval $(-R, R)$ and diverges in the intervals $(-\infty, -R)$ and (R, ∞). The ratio test does not tell us anything about convergence for $x = -R$ or $x = R$.

It turns out that this is true in general.

Theorem 2.11.1 (Convergence of power series) *If*

$$\sum a_n x^n$$

is a power series, then

1. *The series converges absolutely for all x, or*
2. *The series converges only when $x = 0$, or*
3. *There exists a number $R > 0$ such that $\sum a_n x^n$ converges absolutely if $|x| < R$ and diverges if $|x| > R$.*

The number R in this theorem is called the **radius of convergence** of the power series. We write $R = \infty$ if the series converges for all x and $R = 0$ if it only converges for $x = 0$. The set of all real numbers x for which the series converges is called its **interval of convergence**. This interval is one of the following: $(-R,\ R\ )$, $(-R,\ R\ ]$, $[-R,\ R\ )$, or $[-R,\ R\ ]$.

To figure out which type of interval we are dealing with if R is neither 0 nor infinity, we have to substitute $-R$ and R into the power series and investigate whether the resulting number series converge or diverge.

Example 2.11.1 *Let's find the interval of convergence of the power series.*

$$\sum_{n=1}^{\infty} \frac{x^n}{n \cdot 3^n}.$$

Let's apply the ratio test. Here the general term of the series is $x^n/(n \cdot 3^n)$. Therefore

$$\lim_{n\to\infty} \left| \frac{(n+1)3^{n+1}}{n \cdot 3^n} \right| = \lim_{n\to\infty} \frac{3(n+1)}{n} = 3.$$

This means the radius of convergence of our series is 3. Substituting $x = 3$ into the series results in the divergent harmonic series $\sum(1/n)$, and substituting $x = -3$ we arrive at the convergent series $\sum(-1)^n/n = \ln(2)$. We conclude that the interval of convergence of our power series is $[-3, 3)$.

Example 2.11.2 *Let's consider the power series*

$$\sum_{n=0}^{\infty} \frac{(-2)^n x^n}{(2n)!} = 1 - \frac{2x}{2!} + \frac{4x^2}{4!} - \frac{8x^3}{6!} + \frac{16x^4}{8!} - \cdots$$

We again apply the ratio test.

$$\lim_{n\to\infty} \left| \frac{2^n}{(2n)!} \frac{(2n+2)!}{2^{n+1}} \right| = \lim_{n\to\infty} (2n+1)(2n+2)/2 = \infty$$

for all x, so the series converges for all x.

2.11.2 Differentiating and integrating power series

We can turn a given power series into other power series by differentiating and integrating it. Here is the theorem that tells us what is possible in this respect.

Theorem 2.11.2 (Termwise differentiation and integration) *Suppose that the function $f(x)$ has a power series representation*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$$

with nonzero radius of convergence R. Then f is differentiable on $(-R,\ R)$ and

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \cdots$$

Also,

$$\int_0^x f(t)\ dt = \sum_{n=0}^{\infty} \frac{a_n x^{n+1}}{n+1} = a_0 x + \frac{1}{2} a_1 x^2 + \frac{1}{3} a_2 x^3 + \cdots$$

for each x in $(-R,\ R)$. Moreover, both new power series have the same radius of convergence R.

A very important consequence of this theorem is the fact that a power series representation of a function is unique, that is, if $f(x) = \sum a_n x^n = \sum b_n x^n$ for $|x| < R$, $R > 0$, then $a_n = b_n$ for all n. In particular, the Maclaurin series of a function is its unique power series representation (if there is any). To see this note that setting $x = 0$ in

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 \cdots = b_0 + b_1 x + b_2 x^2 + b_3 x^3 \cdots$$

implies that $a_0 = b_0$. Differentiating both sides gives

$$a_1 + 2a_2 x + 3a_3 x^2 \cdots = b_1 + 2b_2 x + 3b_3 x^2 \cdots .$$

Setting $x = 0$ again yields $a_1 = b_1$. Differentiating again and setting $x = 0$ again yields $a_2 = b_2$, and so on.

Example 2.11.3 *By the above theorem we can differentiate the following identity which is valid for $-1 < x < 1$ on both sides*

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots$$

to arrive at the new identity

$$\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + \cdots$$

which is valid in the same interval.

Example 2.11.4 *Substituting x by $-t$ in the identity*

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots$$

gives

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \cdots + (-1)^n t^n + \cdots$$

Because the derivative of $\ln(1+t)$ *is* $1/(1+t)$*, integration from* $t = 0$ *to* $t = x$ *gives*

$$\ln(1+x) = \int_0^x \frac{1}{1+t} dt = \int_0^x (1 - t + t^2 - \cdots + (-1)^n t^n + \cdots)\ dt.$$

And so

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots + \frac{(-1)^{n+1}}{n}x^n + \cdots$$

for $-1 < x < 1$.

Example 2.11.5 *Since*

$$(\tan^{-1}(t))' = \frac{1}{1+t^2},$$

termwise integration of

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - t^6 - \cdots$$

gives

$$\tan^{-1}(x) = \int_0^x \frac{1}{1+t^2} dt = \int_0^x (1 - t^2 + t^4 - t^6 + t^8 - \cdots)\ dt$$

if x is contained in the interval $(-1, 1)$ *in which the geometric series converges. Therefore*

$$\tan^{-1}(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \frac{1}{9}x^9 - \cdots$$

for $-1 < x < 1$.

2.11.3 Power series in powers of $x - c$

An infinite series of the form

$$\sum_{n=0}^{\infty} a_n (x-c)^n = a_0 + a_1(x-c) + a_2(x-c)^2 + \cdots$$

where c is a constant, is called a power series in (powers of) $x - c$. In the previous sections we can replace x^n by $(x-c)^n$ and everything we said there will stay true. In particular,

1. The power series converges absolutely for all x; or

2. The series converges only when $x - c = 0$, that is, when $x = c$; or

3. There exists a number $R > 0$ such that the series in converges absolutely if $|x - c| < R$ and diverges if $|x - c| > R$.

As in the case of a power series with $c = 0$, the number R is called the radius of convergence of the series.

Chapter 3

Integration

3.1 Some integration techniques

3.1.1 Elementary functions

Have a look at the keys of a scientific calculator. There you'll find everybody's favourite functions: x^n, $\sin(x)$, $\cos(x)$, $\ln(x)$, *constants*, etc. These functions plus all the functions that can be formed by combining them by substitution and the operations of addition, subtraction, multiplication, division and powers are usually referred to as the *elementary functions*. Here are some examples of elementary functions:

$$\cos(x) + \sin(x) \cdot \tan(x), \quad \ln(\ln(\sin(x))), \quad \ln\left(\frac{x^n}{\tan(x^2+1)}\right).$$

Most of the functions that you will encounter in textbooks on mathematics are elementary functions[1]. Although there are lots of functions that are not elementary, the elementary functions are in many respects the most useful functions in science and mathematics and juggling these functions is something that everybody who studies mathematics has to master.

3.1.2 Differentiating and integrating

The derivative of an elementary function is again an elementary function and because of the various simple differentiation rules at our disposal it is a straightforward exercise to find this

[1]...or at least *piecewise elementary* in the sense that the functions are elementary functions when restricted to certain intervals. As a simple example consider the function $|x|$. It is not an elementary function. However, it coincides with the elementary function x when restricted to the interval $[0, \infty)$ and to the elementary function $-x$ when restricted to the interval $(-\infty, 0]$.

new elementary function. Why is this so? To start with, we know that the derivatives of all our basic building blocks ($\sin(x)$, x^n, etc.) are again elementary functions. For example, the derivative of $\sin(x)$ is $\cos(x)$ and the derivative of x^3 is $3x^2$. Furthermore, the rules for finding the derivatives of sums, products, and substitutions mesh in with the way the elementary functions are built—given two elementary functions $f(x)$ and $g(x)$, their sum, product and the combined function $f(g(x))$ are also elementary functions, and we know that the derivatives of these new functions are

$$f'(x) + g'(x), \qquad f'(x)g(x) + g'(x)f(x), \qquad g'(x)f'(g(x)),$$

which are again elementary functions.[2] To summarize:

No problems with derivatives. *The derivatives of elementary functions are elementary functions and finding these elementary functions is easy.*

To *integrate* a function $f(x)$ means finding an *anti-derivative* of $f(x)$, that is a function $F(x)$ whose derivative is $f(x)$. Of course, if $f(x)$ is elementary, what you are really looking for is an $F(x)$ that can be written as an elementary function—to identify $F(x)$ as some function you "know". However, this is not always possible because there are many elementary functions whose integrals are not elementary functions. For example, one very important example of an elementary function whose integral is not elementary is e^{-x^2}. Again, what this means is that the integral of this function, which is a perfectly normal well-behaved function cannot be written as a simple combination of everybody's favourite functions.

There is another problem. Even if the integral of an elementary function is an elementary function, it is usually not easy to see which function this is. The main reason behind both problems is that unlike with differentiation, not all the different ways of combining elementary functions come with integration rules that mesh in with them. For example, given two elementary functions $f(x)$ and $g(x)$ with integrals $F(x)$ and $G(x)$, there are no general rules that tell you how to express the integrals of $f(x) \cdot g(x)$ or $f(g(x))$ in terms of $f(x), g(x), F(x)$ and $G(x)$. To summarize:

Problems with integrals. *The integral of an elementary function is not necessarily elementary, and even if it is elementary, it may not be easy to find.*

All this sounds pretty grim. However, there are some tricks that can help you to: 1) recognize that a certain elementary function has an elementary anti-derivative and 2) find this anti-derivative.

[2] The derivatives of powers of the form $f(x)^{g(x)}$ can be dealt with using the same trick that we applied to take care of the indeterminate forms involving powers 0^0, 1^0 and ∞^0. Simply, rewrite $f(x)^{g(x)}$ in the form $e^{g(x)\ln(f(x))}$ and go on autopilot from there.

Rules of the integration game

Let's play a game. Here are the rules which you have to follow precisely if you want to win:

1. I present you with a function that I know has an elementary anti-derivative, for example, $\sin(x)$.
2. You have to find one such anti-derivative, for example, $-\cos(x)$.
3. Give your answer in the form

$$\int \sin(x)\ dx = -\cos(x) + C.$$

The constant of integration: One really silly way to lose the game

The function $-\cos(x)$ is just *one* anti-derivative of the function $\sin(x)$, and so are $-\cos(x)+5$ and $-\cos(x)-32$. In general, adding a constant to an anti-derivative of a function gives another anti-derivative of the same function and any two anti-derivatives of the same function only differ by some constant. This fact is captured by writing the $+C$ (= "plus a constant") at the end of your answer.

Before we start playing make sure to remind yourself of the following basic integrals—all your answers will depend on them.

$$\int x^n\ dx = \tfrac{1}{n+1}x^{n+1} + C, \qquad n \neq -1$$

$$\int \sin(x)\ dx = -\cos(x) + C$$

$$\int \cos(x)\ dx = \sin(x) + C$$

$$\int e^x\ dx = e^x + C$$

$$\int \tfrac{1}{x}\ dx = \ln|x| + C$$

$$\int \tfrac{dx}{\sqrt{1-x^2}} = \arcsin(x) + C$$

$$\int \tfrac{dx}{1+x^2} = \arctan(x) + C$$

The two main tricks that we will be focusing on in this lecture are derived from the product rule and the chain rule for differentiation.

3.1.3 Integration by parts

Our first trick is based on the product rule for differentiation.

Recall that given two differentiable functions $f(x)$ and $g(x)$ the derivative of their product is

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

Now, integrating both sides gives

$$f(x)g(x) + C = \int f(x)g'(x)\,dx + \int g(x)f'(x)\,dx.$$

Therefore,

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

This gives a way to turn the task of finding the integral on the left into the task of finding the integral on the right. Our hope is that the second integral is easier to find than the first. This will depend on the choices we make for $f(x)$ and $g'(x)$. If the function I give you contains a product of functions, this rule, called **integration by parts**, may be the one you should consider using. However, there are usually a few different possible choices and picking the right one can be tricky. Let's look at a couple of examples.

Example

Find

$$\int xe^x\,dx.$$

using integration by parts.

Answer:

We have to split the *integrand* xe^x into two pieces. You may not realize this, but there are actually infinitely many choices. The obvious one,

$$x \cdot e^x,$$

the not so obvious one

$$1 \cdot (xe^x),$$

the really outlandish tricky one

$$(x \sin x) \cdot \frac{e^x}{\sin(x)}$$

and infinitely many variations of this last choice (replace $\sin(x)$ by any function you like).

However, let's just see what we get if we go for the first obvious choice of splitting things up, by splitting the integrand into x and e^x. Even now that we have made this first choice, there is one more choice to be made—shall we set

$$f(x) = x, \qquad g'(x) = e^x,$$

or the other way around? Have another look at the formula we wish to apply:

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

Well, we will differentiate $f(x)$ and integrate $g'(x)$. When we multiply the two resulting functions to get the integrand on the right, we would like to get a product that is simpler than the one we started with. Obviously, it does not matter whether we differentiate or integrate e^x since the result will be e^x in both cases. However, differentiating x gives the simpler function 1, whereas integrating x yields the more complicated function $\frac{1}{2}x^2$. Therefore,

$$f(x) = x, \qquad g'(x) = e^x$$

appears to be the obvious choice and we get

$$f'(x) = 1, \qquad g(x) = e^x.$$

Substituting all this into our formula

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

gives

$$\int xe^x\,dx = xe^x - \int e^x\,dx.$$

Now we just have to complete the simple integral on the right to get the solution to our problem:

$$\int xe^x\,dx = xe^x - e^x + C.$$

Don't forget the $+C$!!

Example

Find

$$\int x\cos(x)\,dx.$$

Answer:

The same considerations as in the previous example suggest to try

$$f(x) = x, \qquad g'(x) = \cos(x).$$

Then

$$f'(x) = 1, \qquad g(x) = \sin(x).$$

Substituting these functions in our formula

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

gives

$$\int x\cos(x)\,dx = x\sin(x) - \int \sin(x)\,dx = x\sin(x) + \cos(x) + C.$$

Don't forget the $+C$!!

Example

Find

$$\int x\ln(x)\,dx.$$

Answer:

Looks just like the previous two examples. So, we switch to autopilot and make the obvious choice

$$f(x) = x, \qquad g'(x) = \ln(x).$$

However, there is a bit of a problem—who knows what the integral of $\ln(x)$ is?[3] Looks like we are stuck. So, let's try

$$f(x) = \ln(x), \qquad g'(x) = x.$$

Then

$$f'(x) = \frac{1}{x}, \qquad g(x) = \frac{1}{2}x^2.$$

Substituting these functions in our formula

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

gives

$$\int x\ln(x)\,dx = \frac{1}{2}x^2\ln(x) - \frac{1}{2}\int x\,dx = \frac{1}{2}x^2\ln(x) - \frac{1}{4}x^2 + C.$$

[3] Actually, this is not really a dead end since you can use integration by parts again to deal with $\int \ln(x)\,dx$. Can you do it?

Example

Find
$$\int \frac{\ln(x)}{x}\,dx.$$

Answer:

This is a tricky one. As in the previous example, our first attempt is to split up the integrand in the natural way
$$\ln(x)\cdot\frac{1}{x},$$
and since setting $g'(x)=\ln(x)$ gives a dead end (at least for the moment), we try
$$f(x)=\ln(x), \qquad g'(x)=\frac{1}{x}.$$
Then
$$f'(x)=\frac{1}{x}, \qquad g(x)=\ln(x).$$
Substituting in our formula
$$\int f(x)g'(x)\,dx = f(x)g(x)-\int g(x)f'(x)\,dx$$
gives
$$\int \frac{\ln(x)}{x}\,dx = (\ln(x))^2-\int \frac{\ln(x)}{x}\,dx.$$
The integrand on the right is exactly the same as the one on the left. That does not look good since we were really looking for something simpler on the right side. However, we can add $\int \frac{\ln(x)}{x}\,dx$ on both sides and then divide by 2 to arrive at the answer to our problem
$$\int \frac{\ln(x)}{x}\,dx = \frac{1}{2}(\ln(x))^2+C.$$

Neat trick! But don't forget the $+C$!!

3.1.4 Substitution

You already know this second trick from school, but let's just have another look at it anyway. Let's say that the function that I give you contains one simple function substituted in another simple function resulting in a fairly terrible combined function, like, for example,
$$\ldots\sin(3x)\ldots, \quad \ldots(\ln(x))^7\ldots, \text{or } \ldots\sin(\ln(x))\ldots.$$

You don't know the integral of $\sin(3x)$, but you do know what the integral of $\sin(x)$ is. In a situation like this, the first thing you should try is a **substitution** or **change of variable**, something that you've all been drilled to do in school. This means that you make the second function, in this case $3x$, into a new variable u and thereby change the integral to look like this

$$\ldots \sin(u) \ldots,$$

which looks a lot more doable than before. Looks like a good idea, but there are a few more things that you need to do to make this work. Let's "just do it" a couple of times.

Example

Find

$$\int \sin(3x)\, dx.$$

Answer:

We start by making our substitution

$$u = 3x\ .$$

Now the integral looks like this

$$\int \sin(u)\, dx$$

That does not look right, since there should really be a du at the end to make this into a "proper" integral. Well, considering u as a function in x we get

$$u'(x) = \frac{du}{dx} = 3,$$

or

$$dx = \frac{du}{3}.$$

Let's substitute this in

$$\int \sin(u)\, dx$$

and we get

$$\begin{aligned}
\int \sin(3x)\, dx &= \int (\sin(u)) \left(\frac{1}{3} du\right) \\
&= \frac{1}{3} \int \sin(u)\, du \\
&= \frac{1}{3} (-\cos(u)) + C.
\end{aligned}$$

Now, we flip back to the variable x,

$$\int \sin(3x)\,dx = -\frac{1}{3}\cos(3x) + C.$$

Example

Find

$$\int xe^{x^2}\,dx.$$

Answer:

Choose a substitution that targets the ugly bit in the integral. Thus put $u = x^2$. Then $\frac{du}{dx} = 2x$ and $dx = \frac{du}{2x}$. This gives

$$\int \frac{1}{2}e^u\,du = \frac{1}{2}e^u + C = \frac{1}{2}e^{x^2} + C$$

Example

Find

$$\int e^{2x}\cos(e^x)\,dx.$$

Answer:

Set $u = e^x$. Then $\frac{du}{dx} = e^x$ and $dx = \frac{du}{e^x}$. By substituting everything in sight we get

$$\int e^{2x}\cos(e^x)\,dx = \int u\cos(u)\,du.$$

We've already solved the integral on the right using integration by parts. So, we can immediately continue

$$\int e^{2x}\cos(e^x)\,dx = \int u\cos(u)\,du = u\sin(u) + \cos(u) + C = e^x\sin(e^x) + \cos(e^x) + C.$$

Chain rule

This is all very well, but, seriously, do you really know what you are doing when you are performing the last trick? For example, what does dx actually mean? When we perform our trick we manipulate this symbol like an ordinary number or a variable, which it actually is not at all. What we have done here definitely does work (sigh of relief). However, the pseudo-arguments that we used to justify what we are doing do not amount to a proof that this is really the case. Our trick is really just a very streamlined version of a theoretical argument that we will not look at in detail in this unit. However, I would like to give you at least a hint as to what is really happening here.

Our first trick, integration by parts, was based on the product rule for differentiation. Although this is not obvious from what we have done, our second trick is based on the chain rule for differentiation:

Given two differentiable function $f(x)$ and $g(x)$, the **chain rule** tells us that the derivative of $f(g(x))$ is

$$(f(g(x)))' = g'(x)f'(g(x)).$$

For example, if $f(x) = \sin x$ and $g(x) = x^2$, then $f(g(x)) = \sin(x^2)$, and the derivative of $\sin(x^2)$ is

$$(\sin(x^2))' = (x^2)' \cos(x^2) = 2x \cos(x^2).$$

We "integrate" the chain rule to arrive at the following rule:

$$\int g'(x)f'(g(x))\, dx = f(g(x)) + C.$$

In turn this means that if I hand you an elementary function of the form $n'(x)m'(n(x))$, then the right side tells you what its integral is.

If you have a close look at this formula you will see that the integrand on the left side contains the sort of nested function that our second trick deals with and our trick really is just a refined version of this formula. You can use the formula alone to deal with the simplest integrals of this kind. For example, let's have another look at our first example and solve it using this formula.

Example

Find

$$\int \sin(3x)\, dx.$$

(Second) Answer:

Okay, we want to use the formula

$$\int g'(x)f'(g(x))\, dx = f(g(x)) + C.$$

So, our guess is that

$$f'(x) = \sin(x), \qquad g(x) = 3x.$$

This implies that

$$f(x) = -\cos(x), \qquad g'(x) = 3.$$

Our formula tells us that

$$\int g'(x)f'(g(x))\,dx = f(g(x)) + C.$$

Substituting we get

$$\int 3\sin(3x)\,dx = -\cos(3x) + C.$$

We can pull the 3 in front of the integration sign and divide on both sides to get our answer

$$\int \sin(3x)\,dx = -\frac{1}{3}\cos(3x) + C.$$

You can deal with our second substitution example in the same way. However, once you try this on our third example you will find that things get quite complicated.

3.2 Proper and improper integrals

IMPORTANT: Most of what is being covered in the rest of this chapter are ideas that we've already been using without making a big deal out of them, for example, in the chapter on the integral comparison test for series. In fact, in the lectures I won't spend much time on all this and will only mention the most important facts summarized in the rest of this chapter in passing.

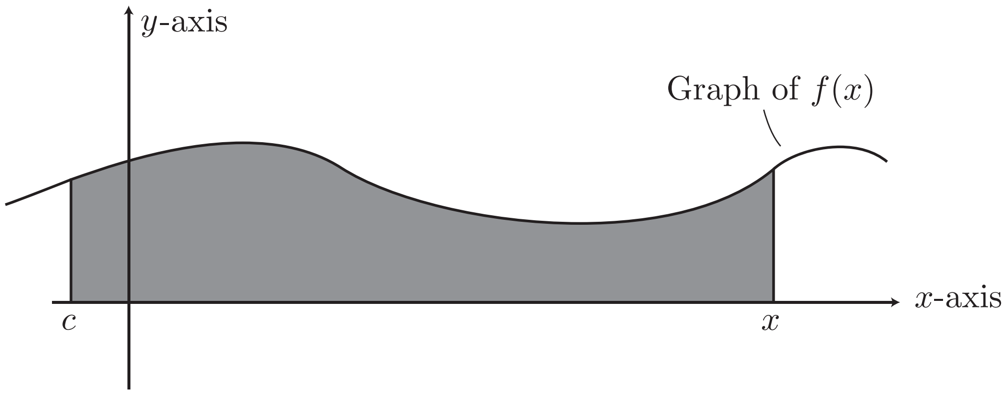
This part of the notes was written a couple of years ago and I've only modified it slightly to make it fit into these new lecture notes. The overall integration with the rest of the lecture notes is still a little bit rough around the edges.

3.2.1 Geometrical interpretation of integrals in a nutshell

The following diagram shows the graph of a positive continuous function and the area under the graph in between two points c and x on the x-axis shaded in gray.

Let's fix a, make x into a variable, and let $F(x)$ denote the area under the graph depending on this variable. Then, and this is the neat part, the function $F(x)$ can be shown to be one of the anti-derivatives of the function $f(x)$ that we started with, that is

$$F'(x) = f(x).$$

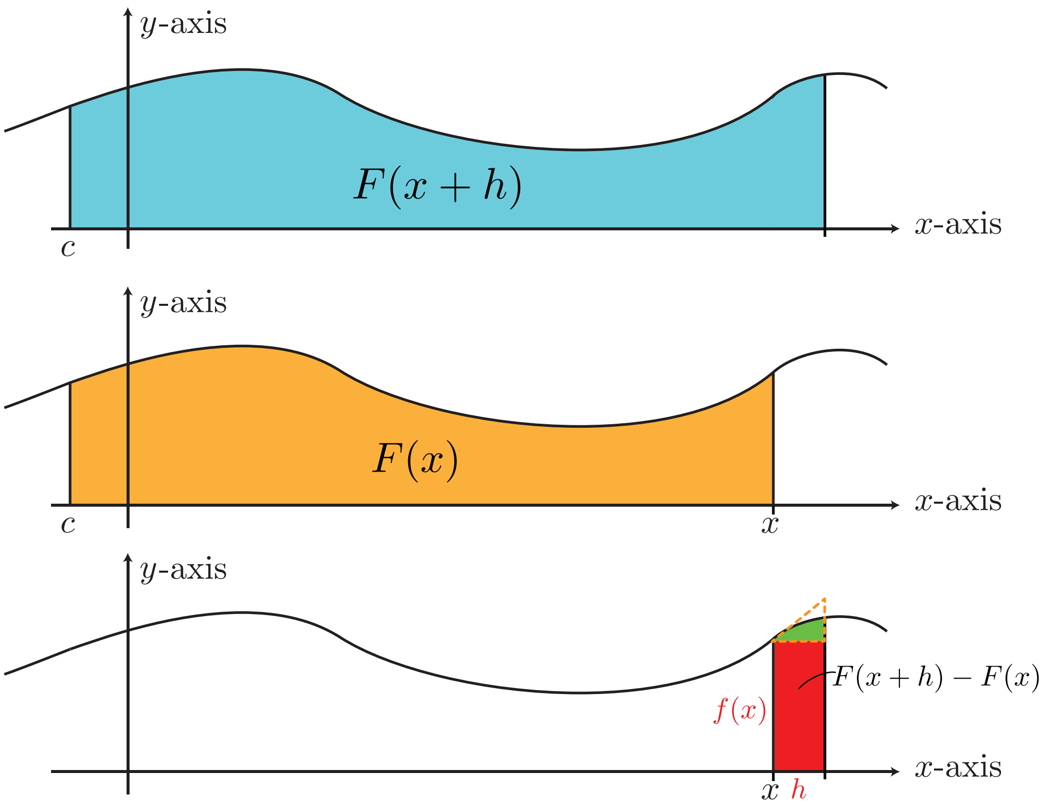


This result is known as the **Fundamental theorem of calculus**. It is, as the name suggests, one of the most important results in calculus. You should all know the simple idea behind the rigorous proof of this result. So, here we go.

Sketch of a proof of the Fundamental Theorem of Calculus. We are interested in the limit

$$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h}.$$

As shown in the following diagram, the difference in the numerator is just the area of the red rectangle plus the green corner. On the other hand, the red area is equal to $f(x) \cdot h$.



This means that

$$F'(x) = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} \frac{f(x)\cdot h + \text{green area}}{h} = f(x) + \lim_{h\to 0} \frac{\text{green area}}{h}.$$

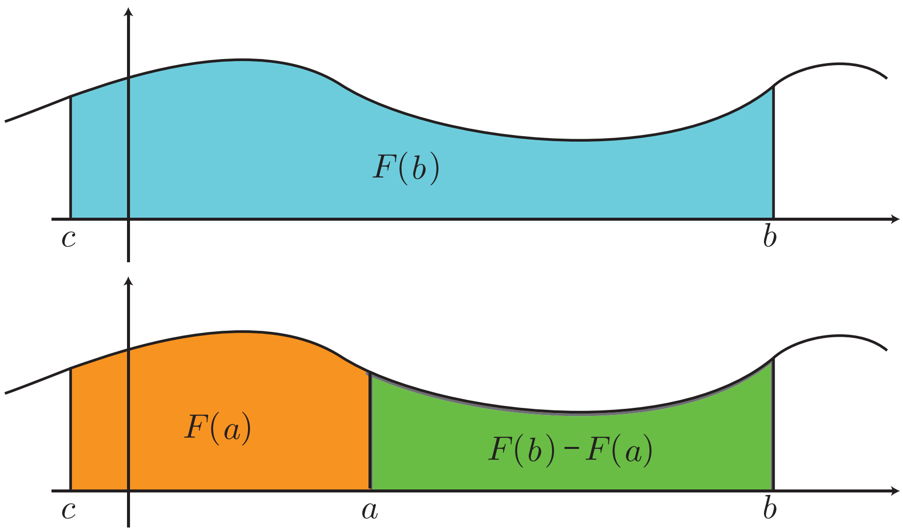
However, it is obvious[4] in the special case that we are looking at, and it can be shown in general that the limit on the right involving the green area is equal to 0. Hence

$$F'(x) = f(x).$$

How neat is that? ■

The fundamental theorem implies that we can use integrals to calculate area. For example, in the following diagram I have highlighted three areas. By definition the orange one is equal to $F(a)$, the larger blue one is equal to $F(b)$, and, therefore, the area in green is simply

$$F(b) - F(a).$$



It is for this reason that this area is usually written as

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

Example

Let $f(x) = x^2$. Then

$$\int x^2\, dx = \frac{1}{3}x^3 + C$$

[4]Just in case this is not obvious to you note that the green area is definitely less than the orange dotted triangle whose area is $f'(x)h^2/2$. Therefore, $\lim_{h\to 0} \frac{\text{green area}}{h} = \lim_{h\to 0} f'(x)h/2 = 0$.

and the area between the two points a and b on the x-axis is simply

$$\int_a^b x^2\,dx = \left[\frac{1}{3}x^3\right]_a^b = \frac{1}{3}b^3 - \frac{1}{3}a^3.$$

The expression

$$\int f(x)\,dx$$

is usually referred to as the **indefinite** integral of $f(x)$, whereas

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

is called a **definite integral**.

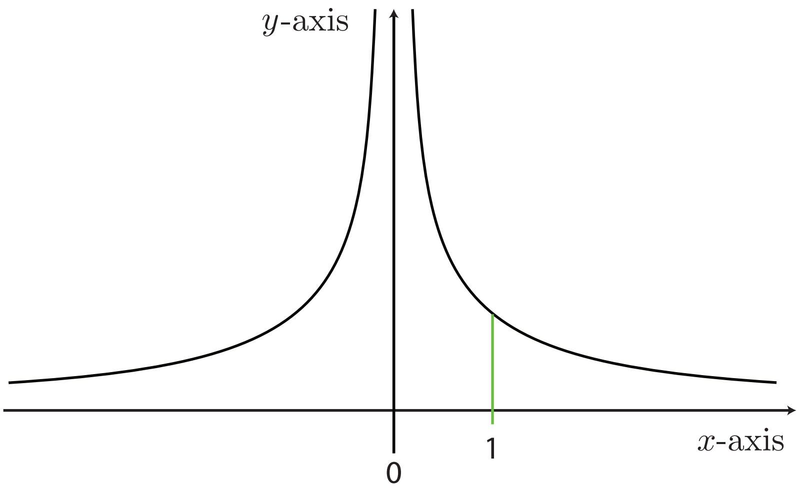
IMPORTANT: Note that we started with a positive continuous function and that our geometrical interpretation referred to this type of function. However, everything we said stays true for arbitrary continuous functions and other "integrable" functions as long as we modify what we mean by area in a reasonable way. Most importantly, if the function dips below the x-axis, then the area between the graph and the x-axis is interpreted as "negative" area.

3.2.2 Improper integrals

To start with, the definite integral $\int_a^b f(x)\,dx$ is only defined if both a and b are real numbers and $f(x)$ is "integrable" in the closed interval $[a, b]$. In particular, "integrable" means that the function $f(x)$ is bounded (not infinite) everywhere in this interval. So what happens if we push things to a limit by setting $a = -\infty$, or $b = \infty$, or if the function $f(x)$ becomes unbounded at one or several points inside the interval $[a, b]$? When this happens the definite integral is called an **improper** integral.

To motivate things, let's have a close look at the graph of the function $\frac{1}{x^2}$ shown in the following diagram. Here are some of its most important features:

1. The function is defined everywhere except for $x = 0$.
2. $$\lim_{x\to 0^+} \frac{1}{x^2} = \lim_{x\to 0^-} \frac{1}{x^2} = \infty$$
   which just means that as we approach the critical point $x = 0$ from the right ($x \to 0^+$) the value of the function goes to ∞ and, similarly, as we approach the critical point $x = 0$ from the left ($x \to 0^-$) the value of the function also goes to ∞.
3. The function is positive wherever it is defined.



Now let's calculate some areas using the integral of this function

$$\int_a^b \frac{1}{x^2}\,dx = \left[-\frac{1}{x}\right]_a^b.$$

As long as a and b are both positive or both negative, the interval $[a,b]$ does not contain the critical point $x=0$ and, in this case, this expression gives the correct area under the graph. However, what about

$$\int_{-1}^1 \frac{1}{x^2}\,dx = \left[-\frac{1}{x}\right]_{-1}^1 = -\frac{1}{1} - \left(-\frac{1}{-1}\right) = -2\ ?$$

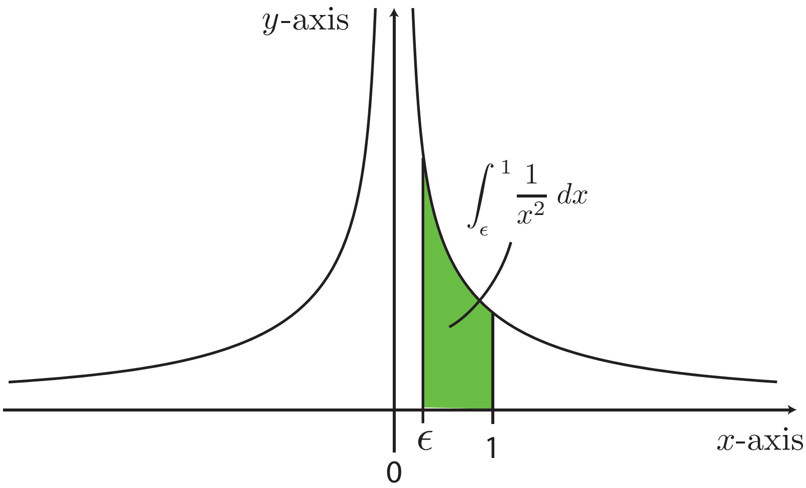
That certainly does not make any sense! And what about

$$\int_0^1 \frac{1}{x^2}\,dx = \left[-\frac{1}{x}\right]_0^1 = -\frac{1}{1} + \frac{1}{0} = \ ?$$

Here we encounter a division by 0 which, of course, is not allowed. However, what this expression seems to say is that the area we are interested in is "$-1+\infty = \infty$", that is, infinitely large, which at least seems possible. Well, the mathematically correct way to check whether this is true is to sneak up on the correct area using a limit of finite areas that we can calculate. Let's consider

$$\int_\epsilon^1 \frac{1}{x^2}\,dx$$

where $0<\epsilon<1$. This definite integral is well-defined and is equal to the area of the green region in the following diagram.



As we let ϵ go to 0, this integral should approach the area we are interested in. This means that what we should be considering is

$$\lim_{\epsilon\to 0^+}\int_\epsilon^1 \frac{1}{x^2}\,dx.$$

No problem,

$$\lim_{\epsilon\to 0^+}\int_\epsilon^1 \frac{1}{x^2}\,dx = -\frac{1}{1} + \lim_{\epsilon\to 0^+}\frac{1}{\epsilon} = \infty.$$

Therefore it makes sense to say

$$\int_0^1 \frac{1}{x^2}\,dx = \lim_{\epsilon\to 0^+}\int_\epsilon^1 \frac{1}{x^2}\,dx = \infty.$$

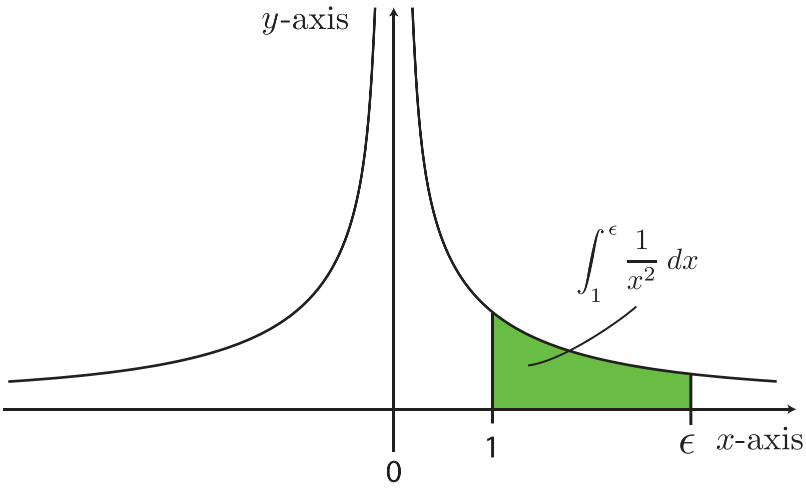
Similarly, it makes sense to define

$$\int_1^\infty \frac{1}{x^2}\,dx$$

as the limit of the area

$$\int_1^\epsilon \frac{1}{x^2}\,dx$$

as we let ϵ go to infinity.



$$\int_1^\infty \frac{1}{x^2}\,dx = \lim_{\epsilon\to\infty} \int_1^\epsilon \frac{1}{x^2}\,dx = \lim_{\epsilon\to\infty} \left(-\frac{1}{\epsilon} - \left(-\frac{1}{1}\right)\right) = 1.$$

The general case

We distinguish a few different types of improper integrals.

Type 1a: An improper integral of the form
$$\int_a^\infty f(x)\,dx,$$
where a is a real number and $\int_a^\epsilon f(x)\,dx$ is a proper integral for all $\epsilon > a$.

We define
$$\int_a^\infty f(x)\,dx = \lim_{\epsilon\to\infty} \int_a^\epsilon f(x)\,dx$$
if the limit on the right exists. If the limit is finite, then the improper integral is called **convergent**. If the limit does not exist, or if it is infinite, then the improper integral is called **divergent**. If the limit is $\pm\infty$ we say that the integral **diverges to infinity** and write
$$\int_a^\infty f(x)\,dx = \pm\infty.$$
We've already seen one example of a convergent improper integral of this type:
$$\int_1^\infty \frac{1}{x^2}\,dx = 1.$$
Here is another example of an improper integral:
$$\int_0^\infty \cos(x)\,dx.$$

Let's consider

$$\int_0^{\epsilon} \cos(x)\, dx = [\sin(x)]_0^{\epsilon} = \sin(\epsilon).$$

Since the limit $\lim_{\epsilon \to \infty} \sin(\epsilon)$ does not exist, our integral is divergent.

And one more simple example:

$$\int_0^{\infty} dx.$$

Here

$$\int_0^{\epsilon} dx = [x]_0^{\epsilon} = \epsilon,$$

and therefore

$$\int_0^{\infty} dx = \lim_{\epsilon \to \infty} \epsilon = \infty.$$

This means that this integral diverges to infinity.

Type 1b: We define $\int_{-\infty}^{b} f(x)\, dx$ analogously:

$$\int_{-\infty}^{b} f(x)\, dx = \lim_{\epsilon \to -\infty} \int_{\epsilon}^{b} f(x)\, dx.$$

Type 2a: An improper integral of the form

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

where both a and b are real numbers, $\lim_{x \to b^-} f(x) = \pm\infty$ and $\int_a^{b-\epsilon} f(x)\, dx$ is a proper integral for all small $\epsilon > 0$.

We define

$$\int_a^b f(x)\, dx = \lim_{\epsilon \to 0^+} \int_a^{b-\epsilon} f(x)\, dx$$

if the limit on the right exists. If the limit is finite, then the improper integral is called *convergent*. If the limit does not exist, or if it is infinite, then the improper integral is called *divergent*.

We've already seen one example of a divergent improper integral of this type:

$$\int_{-1}^{0} \frac{1}{x^2}\, dx = \infty.$$

Type 2b: An improper integral of the form

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

where both a and b are real numbers, $\lim_{x\to a^+} f(x) = \pm\infty$ and $\int_{a+\epsilon}^{b} f(x)\, dx$ is a proper integral for all small $\epsilon > 0$. We define

$$\int_a^b f(x)\, dx = \lim_{\epsilon\to 0^+} \int_{a+\epsilon}^{b} f(x)\, dx,$$

etc.

Mixed Type: Rather than giving a formal definition let's consider an example, the improper integral

$$\int_{-\infty}^{\infty} \frac{1}{x^2}\, dx.$$

All of the types of improper integrals discussed so far have only one tricky spot each, the tricky spot is located at one of the two limits, and things are "tame" otherwise. In this example we have to worry about three different critical spots. The way to deal with such an integral is to chop it up into integrals of the four types considered so far, like this:

$$\int_{-\infty}^{\infty} \frac{1}{x^2}\, dx = \int_{-\infty}^{-1} \frac{1}{x^2}\, dx + \int_{-1}^{0} \frac{1}{x^2}\, dx + \int_{0}^{1} \frac{1}{x^2}\, dx + \int_{1}^{\infty} \frac{1}{x^2}\, dx.$$

Depending on the integrals in the sum on the right, we have the following possibilities:

The integral on the left is convergent if and only if *all* the integrals in the sum on the right are convergent.

The integral on the right is divergent if and only if *any* of the integrals in the sum on the right is divergent.

If all the integrals on the right have values (real numbers, ∞, $-\infty$), the integral on the left

- has no value if there is both a ∞ and a $-\infty$ on the right (since it does not make sense to add ∞ and $-\infty$).
- has the value ∞ if there are only ∞'s and real number values on the right. ("$\infty + \infty = \infty$" and "$\infty +$ any real number $= \infty$").
- has the value $-\infty$ if there are only $-\infty$'s and real number values on the right.

So what about our example? The values of the four integrals on the right are $1, \infty, \infty, 1$. Since there are infinities among these values, the integral is divergent. Since there are only ∞'s and real numbers, this integral is equal to ∞.

Here is another example,

$$\int_{-1}^{1} \frac{1}{x}\, dx.$$

There is only one tricky spot to worry about, at $x = 0$. So we chop this integral up into the two pieces

$$\int_{-1}^{1} \frac{1}{x}\, dx = \int_{-1}^{0} \frac{1}{x}\, dx + \int_{0}^{1} \frac{1}{x}\, dx.$$

The values of the integrals on the right are $-\infty$ and ∞. Therefore the integral on the left is divergent and has no value associated with it.

Chapter 4

Differential equations

A **differential equation** is a mathematical equation that relates some function of one or more variables with its derivatives. In this chapter the focus will be on differential equations that only involve one variable. In its most general form such a differential equation looks like this

$$H(x, y, y', y'', \ldots, y^{(n)}) = 0.$$

where y is a function of the variable x and H is a function that combines x, y and all those derivatives of y. Here is a fairly awful example.

$$(\sqrt{y'' + y} + xyy')^{666} - 7y = 0.$$

The **order** of a differential equation is the highest order of a derivative occurring in the equation. For example, our horror equation is of order 2 since y'' makes an appearance but no derivatives of higher order do.

Integrating means solving a differential equation of order 1 of the form

$$y' = f(x).$$

Things don't get much simpler than this. Still, as we have seen, integration is not an easy task. Why was that again? Because the integral of an elementary function is not necessarily elementary, and even if it is elementary, it may not be easy to find.

Of course, if we have problems like this with the simplest imaginable differentiable equations we expect things to get even more tricky when it comes to solving more complicated differential equations. And this is indeed the case. Apart from the problems already highlighted in the case of integration the main additional problem is that a differential equation may not have any solutions. There are lots of other problems and we'll highlight them and put them into perspective as we will come across them.

What to expect in terms of solutions

When you solve $y' = f(x)$ the solutions are of the form $F(x) + C$ where C can be any number. This means we get infinitely many solutions depending on a free parameter C. If you plot the graphs of these infinitely many functions you find that there is exactly one of these graphs though every single one of the points of the plane. In other words, given a pair (x_0, y_0) there is exactly one solution to $y' = f(x)$ such that $y(x_0) = y_0$. It is the same sort of behaviour that you can expect for many first order differential equations. That is, given a first order differential equation

$$H(x, y, y') = 0$$

and a point (x_0, y_0) we will often find that there is exactly one solution $y(x)$ to the equation that also satisfies $y(x_0) = y_0$. In fact, often a differential equation comes together with an **initial condition** $y(x_0) = y_0$. This sort of problem combo,

$$H(x, y, y') = 0, \quad y(x_0) = y_0$$

is referred to as an **initial value problem**. For example, let's solve the trivial initial value problem

$$y' = x^2, \quad y(0) = 5.$$

Here the general solution of the differential equation is $\frac{1}{3}x^3 + C$. Therefore the solution to the initial value problem is the function $\frac{1}{3}x^3 + 5$.

What about higher order differential equations? Again the most trivial examples suggest what to expect.

$$y'' = f(x).$$

The general solution of this differential equation is of the form $F(x)+Cx+D$ where $C, D \in \mathbf{R}$. So here we have two parameters that can be chosen freely. One of the solutions can be pinned down by specifying that its graph should pass through a certain point (x_0, y_0) in the plane and should have a certain slope at this point. So, an initial value problem for a second order differential equation looks like this

$$H(x, y, y', y'') = 0, \quad y(x_0) = y_0, \quad y'(x_0) = y'_0.$$

Let's again look at a simple example.

$$y'' = x, \quad y(0) = 5, \quad y'(0) = -1.$$

Integrating once we get $y' = \frac{1}{2}x^2 + C$ and integrating a second time gives the general solution of the differential equation,

$$\frac{1}{6}x^3 + Cx + D.$$

Then $y(0) = 5$ implies that $D = 5$ and $y'(0) = -1$ implies that $C = -1$. Then the solution to our initial value problem is

$$\frac{1}{6}x^3 - x + 5.$$

In general, the most nicely behaved differential equations of order 1 and 2 have unique solutions to every possible initial value problem and, based on what we just said about orders 1 and 2, it should now be easy to predict what to expect in the case of differential equations of order n.[1]

4.1 Some easy first order stuff: Separation of variables

As with integration, in the first instance we are interested in finding solutions to differential equations in terms of elementary functions. This will be possible only under certain circumstances. Our first job will be to have a look at some simple types of differential equations that can be solved (often in terms of elementary functions). By doing so and looking at lots of examples, we'll also get a fairly good idea as to what to expect in terms of solutions of general differential equations, when they exist.

Okay, so we already ticked off the simplest type which was

$$y' = \frac{dy}{dx} = f(x).$$

Then we have the differential equations of the form

$$y' = \frac{dy}{dx} = f(y).$$

Formally, an equation like this can be dealt with as follows. Rewrite the equation as

$$\frac{dy}{f(y)} = dx$$

and then integrate

$$\int \frac{dy}{f(y)} = \int dx.$$

If

$$G(y) = \int \frac{dy}{f(y)},$$

then the **implicit solution** of our differential equation is

$$G(y) = x + C.$$

Note that what we've done here is really just shuffling symbols around whose meaning we don't entirely understand. We do get the right result, however, all this definitely does not amount to a proof that this really works.

Before we provide a proof though let's consider an example.

[1]The general solution of "nice" differential equations depend on n free variables and any initial condition of the form $y(x_0) = y_0, y'(x_0) = y_0', y''(x_0) = y_0'', \cdots, y^{(n)}(x_0) = y_0^{(n)}$ pins down one of the solutions.

Example 4.1.1 *Let's solve the initial value problem*

$$\frac{dy}{dx} = y^2, \quad y(0) = 1$$

So

$$\int \frac{1}{y^2} dy = \int dx + C.$$

Hence

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

or

$$y(x) = -\frac{1}{x+C}.$$

We find C by substituting the initial values $x = 0, y = 1$.

$$1 = -\frac{1}{0+C}.$$

And so $C = -1$ and the solution to our initial value problem is

$$y(x) = -\frac{1}{x-1}.$$

Separation of variables is the process of collecting everything to do with x on one side of the equation and everything to do with y on the other side of the equation. The most general form of a differential equation for which this works looks as follows and includes the two simple types of differential equations that we've discussed so far as special cases.

$$\frac{dy}{dx} = g(x)h(y).$$

We formally separate the variables as follows.

$$\frac{1}{h(y)} dy = g(x)\, dx$$

and then integrate

$$\int \frac{1}{h(y)} dy = \int g(x)\, dx$$

If the anti-derivatives $F(y) = \int \frac{1}{h(y)} dy$ and $G(x) = \int g(x) dx$ can be found, then the resulting equation

$$F(y) = G(x) + C$$

is an **implicit solution** of our original equation. This means that any differentiable function $y(x)$ defined implicitly by this equation is an actual (explicit) solution of our initial equation.[2]

Example 4.1.2 *Let's solve the differential equation* $\frac{dy}{dx} = -6xy$ *with the initial condition* $y(0) = 7$.

We first separate the variables

$$\frac{1}{y}dy = -6x\,dx.$$

Then

$$\int \frac{1}{y}dy = \int (-6x)\,dx$$

and therefore

$$\ln|y| = -3x^2 + C.$$

This gives the solutions

$$y(x) = \pm e^{-3x^2+C} = \pm e^C e^{-3x^2}.$$

Since the exponential function is positive the only way we can satisfy our initial condition $y(0) = 7$ *is if the sign in front of the exponential function is a plus. Now we substitute* $x = 0$ *and* $y = 7$ *and thereby find that* $C = \ln(7)$. *Therefore the solution to our initial value problem is*

$$y(x) = e^{\ln 7}e^{-3x^2} = 7e^{-3x^2}.$$

IMPORTANT: This example illustrates a very important point. By separating variables and then going on autopilot we found solutions to our original differential equation. However, we did not find all. There was one that we missed out on. It is the solution $y(x) = 0$. We lost this solution when we separated variables in the first step and divided by y. Of course this is not possible if $y(x) = 0$ is actually a solution to the initial equation. The very important lesson to be learned from this is that when you separate variables and divide by $h(y)$, always check whether any solution of $h(y) = 0$ also happens to be a solution of your differential equation. If this is the case then any such solution will usually not show up in the subsequent analysis.

[2] To verify this claim and to justify our formal manipulations that led to this claim, suppose that $y = y(x)$ satisfies $F(y) = G(x) + C$ for some choice of C. Then differentiating both sides

$$[F(y(x))]' = [G(x) + C]'$$

and applying the chain rule on the left gives

$$F'(y(x)) \cdot y'(x) = G'(x).$$

However, this equation can also be written as

$$\frac{1}{h(y(x))}y'(x) = g(x) \text{ or } y'(x) = g(x)h(y(x)).$$

This means that $y(x)$ really does solve our original equation.

Here is another example where this becomes an issue.

Example 4.1.3 *Let's solve*

$$\frac{dy}{dx} = 6x(y-1)^{2/3}.$$

To separate the variables we'll divide by $(y-1)^{2/3}$. However, note that $(y-1)^{2/3} = 0$ if we set $y(x) = 1$. Also, $y(x) = 1$ is clearly a solution to our given differential equation. Okay, so we've found one solution, now let's find the others.

$$\int \frac{1}{3(y-1)^{2/3}}\, dy = \int 2x\, dx.$$

Therefore,

$$(y-1)^{1/3} = x^2 + C.$$

We are in luck, we can solve this implicit solution for the general solution

$$y(x) = 1 + (x^2 + C)^3.$$

A final example illustrates what happens when we cannot solve the implicit solution for y.

Example 4.1.4 *Let's solve the initial value problem*

$$\frac{dy}{dx} = \frac{4-2x}{3y^2-5},\ y(1) = 3.$$

We switch to autopilot and conclude

$$\int (3y^2 - 5)\ dy = \int (4-2x)\ dx.$$

This equation simplifies as follows.

$$y^3 - 5y = 4x - x^2 + C.$$

If we substitute $x = 1$ and $y = 3$ in the implicit solution, we find that $C = 9$. Hence the desired particular solution $y(x)$ is defined implicitly by the equation

$$y^3 - 5y = 4x - x^2 + 9.$$

Here is the graph of the solution curve plotted in red. The diagram also shows some solution curves for C different from 9. [3]

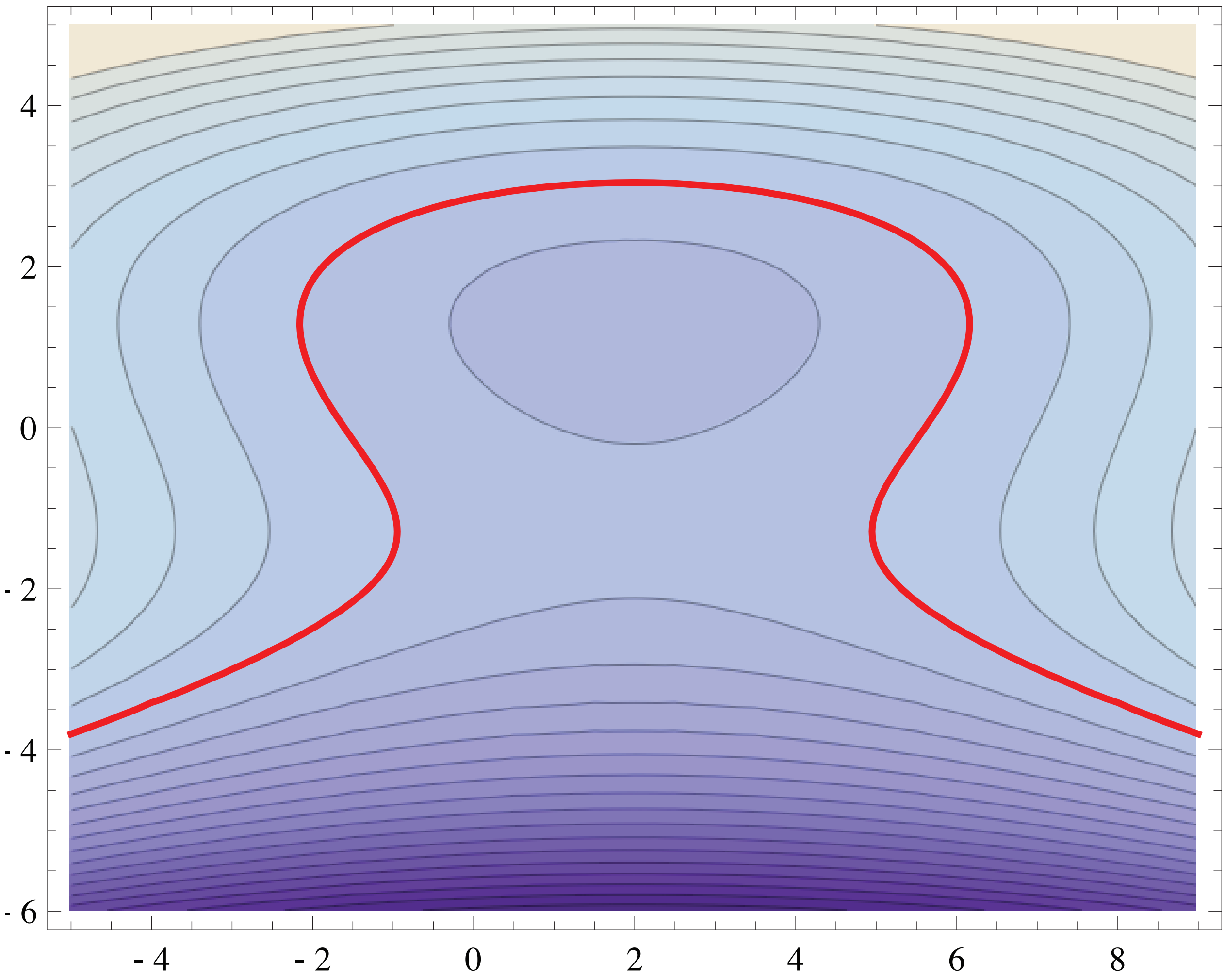
[3] To replicate this plot in *Mathematica* execute the following command.

```
ContourPlot [ y^3 - 5 y - (4 x - x^2 + 9), {x, -5, 9}, {y, -6, 5}, AspectRatio ->
Automatic, Contours -> 25 ]
```

Or to just see the curve defined by the equation and access all the *WolframAlpha* functionality execute

```
WolframAlpha["y^3 - 5 y = 4 x - x^2 + 9"]
```

(You need an internet connection for this last command to work.)



4.2 The integrating factor trick

A first-order differential equation is called **linear** if it can be written in the form

$$y' + P(x)\ y = Q(x)$$

where $P(x)$ and $Q(x)$ are functions of x.

There is a really neat trick that allows us to deal with this type of differential equation. To "discover" this trick let's find the derivative of the product $y(x)$ with a yet to be determined mystery function $I(x)$ using the product rule. Don't ask why, just do it!

$$[y(x)I(x)]' = y'(x)I(x) + y(x)I'(x).$$

Omitting the (x) following the y we can rewrite this equation as

$$[yI(x)]' = I(x)\ y' + I'(x)\ y.$$

The right side of this equation looks very similar to the left side of our differential equation

$$y' + P(x)\ y = Q(x).$$

Let's multiply the differential equation by $I(x)$ to add to this similarity

$$I(x)\ y' + I(x)P(x)\ y = I(x)Q(x).$$

Now, if we are lucky we can choose the function $I(x)$ such that the two expressions in red are equal. And this would be a very good thing to have because then the differential equation could be rewritten as

$$[y(x)I(x)]' = I(x)Q(x).$$

From there finding the solution is easy.

$$y(x) = \frac{1}{I(x)} \int I(x)Q(x)\,dx.$$

So, how do we choose $I(x)$ to be able to perform this miracle? Well, we have to choose $I(x)$ such that

$$I(x)\cdot y' + I(x)P(x)\cdot y = I(x)\cdot y' + I'(x)\cdot y.$$

We've highlighted the difference that needs to be adjusted in blue. We conclude that we have to choose $I(x)$ such that

$$I'(x) = I(x)P(x) \text{ or } \frac{dI}{dx} = IP(x),$$

that is, we have to find an $I(x)$ that solves this differential equation. This can be done by separating variables. So, we just rewrite the last equation as

$$\int \frac{1}{I}\,dI = \int P(x)\,dx$$

and solve by following our nose.

Here is an example.

Example 4.2.1 *Let's consider the differential equation*

$$x^3y' + x^2y = 2x^3 + 1$$

We first have to divide by x^3 on both sides to turn this into the type of equation that we have been discussing, that is,

$$y' + \frac{1}{x}y = 2 + \frac{1}{x^3}.$$

Now we multiply by the mystery function $I(x)$.

$$I(x)y' + I(x)\frac{1}{x}y = I(x)\left(2 + \frac{1}{x^3}\right).$$

We want to choose $I(x)$ such that

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

To do this we separate the variables I and x

$$\int \frac{1}{I}\,dI = \int \frac{1}{x}\,dx.$$

So,

$$\ln|I| = \ln|x| + C.$$

Well, we really only need one of the possible I that satisfy this equation and the obvious choice is $I(x) = x$. Okay, this means that our differential equation can be rewritten as

$$(xy)' = x\left(2 + \frac{1}{x^3}\right) = 2x + \frac{1}{x^2}.$$

Therefore,

$$xy = \int \left(2x + \frac{1}{x^2}\right)dx = x^2 - \frac{1}{x} + C.$$

Now we only have to divide by x to arrive at the general solution of our original differential equation,

$$y(x) = x - \frac{1}{x^2} + \frac{C}{x}.$$

Note that at some point we chose the mystery function among a number of possible candidates. On closer inspection it turns out that it does not matter which of these candidate functions you choose to be $I(x)$, the end result is always the same.

Here are some more examples.

Example 4.2.2 *Let's solve the initial value problem*

$$y' - y = \frac{11}{8}e^{-x/3},\ y(0) = 0.$$

Multiply by $I(x)$ to get

$$I(x)y' - I(x)y = I(x)\frac{11}{8}e^{-x/3}$$

So we need to solve

$$I' = -I$$

and of course one solution is $I(x) = e^{-x}$ This means our differential equation can be rewritten as

$$(e^{-x}y)' = e^{-x}\frac{11}{8}e^{-x/3} = \frac{11}{8}e^{-4x/3}.$$

And so

$$e^{-x}y = \int \frac{11}{8}e^{-4x/3}dx = -\frac{33}{32}e^{-4x/3} + C.$$

Finally, multiplication by e^x gives the general solution

$$y(x) = Ce^x - \frac{33}{32}e^{-x/3}.$$

Substituting $x = y = 0$ now gives $C = \frac{33}{32}$, and therefore the solution to our initial value problem is

$$y(x) = \frac{33}{32}e^x - \frac{33}{32}e^{-x/3} = \frac{33}{32}(e^x - e^{-x/3}).$$

Example 4.2.3 *Let's solve*

$$(x^2+1)y' + 3xy = 6x.$$

We divide both sides of the equation by x^2+1 to turn this equation into the form we know how to deal with.

$$y' + \frac{3x}{x^2+1}y = \frac{6x}{x^2+1}$$

We multiply by $I(x)$

$$I(x)y' + I(x)\frac{3x}{x^2+1}y = I(x)\frac{6x}{x^2+1}$$

Now we need to solve

$$I' = I\frac{3x}{x^2+1}.$$

One solution is

$$I(x) = e^{\int \frac{3x}{x^2+1}dx} = e^{\frac{3}{2}\ln(x^2+1)} = (x^2+1)^{3/2}.$$

Then our differential equation becomes

$$((x^2+1)^{3/2}y)' = 6x(x^2+1)^{1/2}$$

and integration gives

$$(x^2+1)^{3/2}y = \int 6x(x^2+1)^{1/2}dx = 2(x^2+1)^{3/2} + C.$$

Multiplying both sides by $(x^2+1)^{-3/2}$ then gives the general solution

$$y(x) = 2 + C(x^2+1)^{-3/2}.$$

Theorem 4.2.1 (Existence and Uniqueness of solutions) *If the functions $P(x)$ and $Q(x)$ are continuous on an open interval containing the point x_0, then the initial value problem*

$$y' + P(x)y = Q(x),\ y(x_0) = y_0$$

has a unique solution $y(x)$ in this interval.

4.3 Second order linear differential equations

In its most general form a second order linear differential equation looks like this:

$$P(x)y'' + Q(x)y' + R(x)y = S(x).$$

Here $P(x), Q(x), R(x)$ and $S(x)$ are functions in x.

Unlike for first order linear differential equations, there does not exist a general method for solving all second order linear differential equations. So, let's get started by considering some special types of equations like this.

4.3.1 Homogeneous equations

First, if $S(x) = 0$, then we are dealing with a so-called **homogeneous** differential equation. Note that such an equation always has $y(x) = 0$ as one of its solutions. It is also very easy to see that any linear combination $Ay_1(x) + By_2(x)$ of two solutions $y_1(x)$ and $y_2(x)$ is again a solution.

In fact, it turns out that if $y_1(x)$ and $y_2(x)$ are linearly independent, that is, if neither one is a multiple of the other, then over *any* interval in which the coefficient functions $P(x), Q(x), R(x)$ are continuous and $P(x) \neq 0$, *every* solution of the differential equation is of the form $Ay_1(x) + By_2(x)$. We also express this by saying that $Ay_1(x) + By_2(x)$ is the *general solution* of the differential equation (over such an interval).

Now let's simplify things even further and consider the case of *constant coefficients* a, b, and c:

$$ay'' + by' + cy = 0.$$

Let's make an educated guess and check whether

$$y(x) = e^{\lambda x}$$

can be made into a solution of our differential equation if we just choose λ right. First, we need the derivatives of $y(x)$:

$$y'(x) = \lambda e^{\lambda x} \quad , \quad y''(x) = \lambda^2 e^{\lambda x}.$$

We substitute these into our equation and get

$$a\lambda^2 e^{\lambda x} + b\lambda e^{\lambda x} + ce^{\lambda x} = 0.$$

Hence

$$(a\lambda^2 + b\lambda + c)e^{\lambda x} = 0.$$

Remember that $e^{\lambda x} \neq 0$. Hence we need to ensure that

$$a\lambda^2 + b\lambda + c = 0.$$

The two solutions of this quadratic equation are

$$\lambda_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{and} \quad \lambda_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

Let's assume for the moment that $\lambda_1 \neq \lambda_2$ and that both are real numbers.

This means that we have found two distinct solutions of our differential equation:

$$y_1(x) = e^{\lambda_1 x} \quad \text{and} \quad y_2(x) = e^{\lambda_2 x}.$$

In fact, it is easy to see that these two solutions are linearly independent. Consequently,

$$y(x) = Ay_1(x) + By_2(x)$$

is the general solution of the differential equation which is valid for all real numbers x.

Example: Two distinct real roots

Find the general solution of

$$y'' + y' - 6y = 0.$$

First we solve the corresponding quadratic equation

$$\lambda^2 + \lambda - 6 = 0.$$

This gives $\lambda_1 = 2$ and $\lambda_2 = -3$ and thus

$$y(x) = Ae^{2x} + Be^{-3x}$$

is the general solution.

The quadratic equation

$$a\lambda^2 + b\lambda + c = 0$$

arising from our guess $y(x) = e^{\lambda x}$ is known as the **characteristic equation** of the differential equation.

Example: Complex roots

Find the general solution of
$$y'' - 2y' + 5y = 0.$$
First we solve the characteristic equation
$$\lambda^2 - 2\lambda + 5 = 0.$$
This gives $\lambda_1 = 1 - 2i$ and $\lambda_2 = 1 + 2i$. We'll just pretend that things somehow still work just like in the case of two real solutions and see where this leads us. So, our general solution is
$$y(x) = Ae^{\lambda_1 x} + Be^{\lambda_2 x} = Ae^{(1-2i)x} + Be^{(1+2i)x}$$
With *Euler's Formula* $e^{\tau + i\theta} = e^{\tau}(\cos\theta + i\sin\theta)$ we get
$$\begin{aligned} e^{(1-2i)x} &= e^x\left(\cos(2x) - i\sin(2x)\right), \\ e^{(1+2i)x} &= e^x\left(\cos(2x) + i\sin(2x)\right). \end{aligned}$$
Using these identities we can rewrite our general solution as follows:
$$y(x) = e^x\left((A+B)\cos(2x) + i(B-A)\sin(2x)\right).$$
Setting, $A = B = \frac{1}{2}$, we arrive at a first candidate for a *real* solution
$$e^x\cos(2x).$$
And by setting $A = -B = \frac{1}{2}i$ we arrive at a second candidate for a real solution
$$e^x\sin(2x).$$
In fact, it is possible to check that these are really solutions to our differential equation by substituting these candidates into the original equation. Furthermore, since the two solutions are clearly linearly independent, we arrive at the general solution
$$y(x) = e^x\left(A\cos(2x) + B\sin(2x)\right).$$

Example: Equal roots

Find the general solution of
$$y'' + 2y' + y = 0.$$
This time we find just one root for λ,
$$\lambda_1 = \lambda_2 = -1.$$
This gives the nontrivial solution e^{-x} of our differential equation. It turns out that a second solution is given by xe^{-x} (double-check this by substituting this function into the original

equation!). Since both functions are linearly independent both combine into the general solution of our differential equation:

$$(A + Bx)e^{-x}.$$

All these considerations generalize in a straightforward manner into the complete solution for ...

constant coefficient 2nd order homogeneous differential equations

$$ay'' + by' + cy = 0.$$

First solve the characteristic equation

$$a\lambda^2 + b\lambda + c = 0.$$

Let the two roots be λ_1 and λ_2. Then for the general solution $y(x)$ we have:

$$\text{Case 1}: \ \lambda_1 \neq \lambda_2 \text{ both real}$$

$$y(x) = Ae^{\lambda_1 x} + Be^{\lambda_2 x}$$

$$\text{Case 2}: \ \text{Two proper}^4 \text{ complex roots}$$
$$\lambda_{1/2} = \alpha \pm i\beta$$
$$y(x) = e^{\alpha x}\left(A\cos(\beta x) + B\sin(\beta x)\right)$$

$$\text{Case 3}: \ \lambda_1 = \lambda_2 = \lambda$$
$$y(x) = (A + Bx)e^{\lambda x}$$

4.3.2 Non-homogeneous equations

Here is the general constant coefficient *non-homogeneous* linear second order differential equation:

$$ay'' + by' + cy = S(x),$$

where a, b, c are constants and $S(x) \neq 0$ is some given function. This differs from the homogeneous case only in that here we have $S(x) \neq 0$.

[4] Let's not forget that real numbers are also complex numbers!

Let's assume that we know *one* solution $y_p(x)$ of such a differential equation. Then adding *any* solution of the corresponding homogeneous equation

$$ay'' + by' + cy = 0$$

to $y_p(x)$ will result in a new solution of our non-homogeneous equation. In fact, it is also clear that all solutions of the non-homogeneous equation are of this form. This means that the general solution of the non-homogeneous equation is

$$y(x) = y_h(x) + y_p(x)$$

where $y_h(x)$ is the general solution of the homogeneous equation.

Example

Find the general solution of

$$y'' + y' - 6y = 1 + 2x.$$

We proceeds in three steps. First, we solve the homogeneous equation. Second, we find a **particular solution**. Third, we add the two solutions together.

Step 1 : The homogeneous solution

Here we must find the general solution of

$$y'' + y' - 6y = 0.$$

Previously we found the general solution for this homogeneous equation to be

$$y_h(x) = Ae^{2x} + Be^{-3x}.$$

Step 2 : The particular solution

Here we have to find any solution of the original non-homogeneous equation. Since the right-hand side is a polynomial we try a guess of the form

$$y_p(x) = a + bx$$

where a and b are numbers which we have to compute.

We substitute this into the left-hand side of the equation and find

$$(a + bx)'' + (a + bx)' - 6(a + bx) = 1 + 2x.$$

Therefore,

$$b - 6a - 6bx = 1 + 2x.$$

Since this must be true for all x we conclude that

$$b - 6a = 1 \qquad \text{and} \qquad -6b = 2.$$

So, $b = -1/3$ and $a = -2/9$ and thus

$$y_p(x) = -\frac{2}{9} - \frac{1}{3}x.$$

Note that finding a particular solution by this guessing method is often called the **method of undetermined coefficients**.

Step 3 : The general solution

This is the easy bit:

$$y(x) = y_h(x) + y_p(x) = Ae^{2x} + Be^{-3x} - \frac{2}{9} - \frac{1}{3}x.$$

Our job is done!

Undetermined coefficients

How do we choose a workable guess for the particular solution? “Simply” by inspecting the terms in $S(x)$, the right hand side of the differential equation.

Here are some examples,

Guessing a particular solution

$$S(x) = a + bx + cx^2 + \cdots + dx^n$$

$$y_p(x) = e + fx + gx^2 + \cdots + hx^n$$

$$S(x) = (a + bx + cx^2 + \cdots + dx^n)e^{kx}$$
$$y_p(x) = (e + fx + gx^2 + \cdots + hx^n)e^{kx}$$

$$S(x) = (a\sin(bx) + c\cos(bx))e^{kx}$$
$$y_p(x) = (e\cos(bx) + f\sin(bx))e^{kx}$$

Example

What guesses would you make for each of the following?

$$\begin{aligned} S(x) &= 2 + 7x^2 \\ S(x) &= e^{3x}\sin(2x) \\ S(x) &= 2x + 3x^3 + \sin(4x) - 2xe^{-3x} \end{aligned}$$

To check whether your guess for, say, the third right-hand side works, simply unleash *Wolfram Alpha* on the equation

$$y'' + y' - 6y = 2x + 3x^3 + \sin(4x) - 2xe^{-3x}.$$

Exceptions

If $S(x)$ contains terms that are solutions of the corresponding homogeneous equation, then in forming the guess for the particular solution you should multiply that term by x (and by x^2 if the term corresponds to a repeated root of the characteristic equation).

Example

Find the general solution for

$$y'' + y' - 6y = e^{2x}.$$

The homogeneous solution is

$$y_h(x) = Ae^{2x} + Be^{-3x}.$$

Here we have an example of our right-hand side containing a solution of the homogeneous equation. The guess for the particular solution would then be

$$y_p(x) = (a + bx)e^{2x}.$$

Now solve for a and b. Try it and double-check your answer using *Wolfram Alpha*.

We have to stress that unless $S(x)$ belongs to a very restricted class of functions, which includes the examples above, you will not be able to find a general solution of a non-homogeneous equation in terms of elementary functions.

4.3.3 General and more general

It should be clear that the theory we've explored in this chapter provides ideas for how to tackle much more complicated linear differential equations. Here are just some examples that illustrate how some of our ideas generalize:

In general, any solution of an nth order linear differential equation

$$P_n(x)y^{(n)} + P_{n-1}(x)y^{(n-1)} + ... + P_1(x)y' + P_0(x)y = S(x)^5$$

is of the form

$$y(x) = y_h(x) + y_p(x),^6$$

where $y_h(x)$ is the linear combination of n linearly independent solutions $y_1(x), y_2(x), \dots, y_n(x)$ of the corresponding homogeneous equation, and $y_p(x)$ is a particular solution of the full equation (if $S(x) \neq 0$).

If all the coefficient functions on the left are constants,

$$p_n y^{(n)} + p_{n-1}y^{(n-1)} + ... + p_1(x)y' + p_0(x)y = S(x),$$

then our guess $y(x) = e^{\lambda x}$ first reduces our problem to finding the solutions of the characteristic equation

$$p_n\lambda^{(n)} + p_{n-1}\lambda^{(n-1)} + ... + p_1(x)\lambda + p_0 = 0.$$

And a complete solution of the homogeneous equation can be derived based on these solutions. For example, in the simplest case of n distinct roots $\lambda_1, \lambda_2, \dots, \lambda_n$ this general solution is just

$$A_1e^{\lambda_1 x} + A_2e^{\lambda_2 x} + \ldots + A_ne^{\lambda_n x}.$$

Or, whenever we have a root λ of order k, then part of the general solution is

$$(A_1 + A_2x + A_3x^2 + A_4x^3 + \ldots + A_nx^{k-1})e^{\lambda x}.$$

Finally, to find a particular solution of a non-homogeneous system you are usually best off trying to make an educated guess.

Okay, here is just one example. Let's solve

$$y^{(5)} - 8y^{(4)} + 24y^{(3)} - 34y'' + 23y' - 6y = 108x^2.$$

The corresponding homogeneous equation yields the characteristic equation

$$\lambda^5 - 8\lambda^4 + 24\lambda^3 - 34\lambda^2 + 23\lambda - 6 = (\lambda - 1)^3(\lambda - 2)(\lambda - 3) = 0.$$

[5] Here $y^{(n)}$ stands for the nth derivative of the unknown function y. For example, $y^{(2)} = y''$.

[6] Note that this is even true for the general solution of first-order linear equations that we arrived at using the integrating factor trick.

This means that the general solution of the homogeneous equation is

$$y_h(x) = (A + Bx + Cx^2)e^x + De^{2x} + Ee^{3x}.$$

Since the right side of the non-homogeneous equation is the polynomial $108x^2$ it is easy to "guess" the particular solution

$$y_p(x) = -18x^2 - 138x - 325.$$

Adding the two gives the general solution of our non-homogeneous equation

$$y(x) = (A + Bx + Cx^2)e^x + De^{2x} + Ee^{3x} - 18x^2 - 138x - 325.$$

Dealing with complex roots is a bit messier but also not a big deal.