Calculus 1: MA141

2022 Lecturer: Roger Tribe

Contents

Chapter 0. Up periscope

This first chapter is an up-persicope to show you where Calculus 1 and 2 are headed, along with a few ideas from year two maths modules. The very careful proof writing in the these two modules is quite a change from school mathematics, and takes time to get used to, often practising on rather simple, or even obvious, statements. At the risk of stealing thunder from my fellow lecturers in probability, statistics, differential equations, combinatorics e.t.c. I think it is important to keep some of the eventual high points in mind. I think this section should be read quickly (without worrying about details) - it is not needed for revision, but we will meet some of the examples again later and also in exercises. The theory in this module really starts with Chapter 1.

10 examples

Here are ten examples we should meet during Calculus 1 and Calculus 2.

1. **Stirling's formula.** Stirling's approximation for factorials

$$n! \approx n^n e^{-n} n^{1/2} \sqrt{2\pi}$$

is enormously useful in many counting problems (and therefore also in various probability problems).

I have left it here with a $\approx$ symbol, just to mean 'approximately equal to'.

We will have to (i) make a precise statement (ii) prove it.

How on earth is π going to emerge?

2. **The number e.** The limit

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

can be taken as a definition of the number e. If we see a different definition we had better check they really are the same. For example we need to reconcile the above limit with the power series

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

which would indicate that $e = \sum_{k=0}^{\infty} \frac{1}{k!}$.

The related limit

$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda} \quad \text{for any } \lambda \in \mathbf{R}$$

arises early in probability modules, for example in Poisson approximations.

3. **Newton Raphson.** You may well have seen the Newton Raphson method in action at school. Here it is as a scheme for calculating $\sqrt{2}$:

$$a_1 = 3/2, \quad a_{n+1} = \frac{a_n^2 + 2}{2a_n} \quad \text{for } n \geq 1.$$

   My first approximation was 3/2. Working to 12 decimal places I get

$$a_2 = 1.41666666666, \quad a_3 = 1.41421568627, \quad a_4 = 1.41421356237$$

   and we are finished to my level of precision. It is an example of a quadratic algorithm, meaning the the nth error, that is the distance between nth approximation and the true value of $\sqrt{2}$, approximately squares each operation. So the number of decimal places that are correct roughly doubles each time you iterate the algorithm. Finding a quadratic algorithm to approximate π was a more recent discovery. Cubic, quartic, ... algorithms (where the error cubes or quarts...) are now known.

4. **The number π.** There are many formulae that lead to π, for example

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

   How many terms of this infinite series need you add before you correctly find the first 3 decimal places of π? We'll see some quicker ways to evaluate π accurately. We'll also meet a third constant (after π and e) that commonly arises: the Euler-Mascheroni constant γ.

5. **Infinite series.** How do you answer a person in the street who says 'if you keep on adding strictly positive numbers for ever you will reach infinity'? (You might be wise to steer clear of them). We know this isn't true since we know

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 2$$

   as this is just an infinite geometric series.

   For most infinite series, even when they converge to a finite total, there is no simple exact formula for the total using our commonly met functions x^a / exp / log / sin / cos . . ..
   But we will learn techniques to be sure the total is finite, and to estimate how many terms we need to add to approximate the total to any desired accuracy.

   One famous infinite series with an simple exact total is

$$\frac{\pi^2}{6} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \ldots$$

   This is known as Euler's identity. There are quite a few proofs but none very short.

6. **Riemann's re-arrangement theorem.** A well known infinite series is

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \ldots = \log(2).$$

(We will always use logarithms to the base e in the this module unless we state otherwise (why?)). Riemann argued that if you add the numbers up in a different order you may get a different total. For example

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \ldots = \frac{3}{2}\log(2).$$

Note here we are adding two of the positive terms each time and then one of the negative terms, but we still end up using up all the positive terms $1, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \ldots$ and all the negative terms $-\frac{1}{2}, -\frac{1}{4}, -\frac{1}{6}, -\frac{1}{8}, \ldots$ (that is we don't miss any out). Mind-boggling? In fact Riemann's argument shows that we can rearrange the terms to get any final total we want. Adding up infinitely many numbers can be slippery.

On the other hand, however, if we rearrange the terms of the geometric series

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \ldots$$

we will always end up with the same total once we have added them all up. How can we tell if the order matters?

7. **Asymptotics.** Approximate answers are often possible for hard problems if they are 'close' to a solvable problem. Here is a simple example. We can solve the equation

$$x^3 - x = 0$$

as we can factor $x^3 - x = x(x-1)(x+1)$ and we find the three solutions $x = -1, 0, 1$. If we change the problem to ask for solutions to

$$x^3 - x + \epsilon = 0$$

for a small constant ϵ the problem is harder (I know there is a formula to solve cubics but I have forgotten it...). But thinking about the graph of $x^3 - x + \epsilon$ suggests, for small ϵ, that there are still three roots and they lie close to $-1, 0, 1$. Let's make an Ansatz (that is a unjustified assumption) the there is a root $r(\epsilon)$ near 1 which be written as

$$r(\epsilon) = 1 + a\epsilon + b\epsilon^2 + \ldots$$

for some as yet unknown constant coefficients a and b - like the first few terms of a series in the increasing powers of the variable ϵ. Substituting this into $x^3 - x + \epsilon = 0$ I get

$$\begin{aligned} 0 &= (1 + a\epsilon + b\epsilon^2 + \ldots)^3 - (1 + a\epsilon + b\epsilon^2 + \ldots) + \epsilon \\ &= \epsilon(2a+1) + \epsilon^2(3a^2 + 2b) + \epsilon^3(\ldots) + \ldots \end{aligned}$$

where I have collected up terms in increasing powers of ϵ. Then we equate the coefficient of ϵ and then ϵ^2 (and then...) to zero, which suggest $a = -1/2$ and $b = -3a^2/2 = -3/8$. Something must go wrong if ϵ is too big as the root becomes complex, but the formula will be accurate for small ϵ.

8. **Gaussian tails.** Gaussian (also called Normal) random variables permeate probability and statistics. A Gaussian $N(0,1)$ variable has the famous bell shaped density

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \text{for } x \in \mathbf{R}.$$

Tail probabilities are the probabilities

$$\Pr[N(0,1) > z] = \int_z^\infty f(x)dx = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-x^2/2} dx$$

It is a famous result that there are no simple anti-derivatives for the function $f(x)$ (in terms of familiar functions x^n/ exp / log / sin / cos ...), so there is no simple formula for this integral (and this is the typical situation for most integrals). When I was a student I was given a large book with values of the integral that had been calculated numerically and recorded. But we'll see good approximations such as

$$\int_z^\infty e^{-x^2/2} dx \approx e^{-z^2/2} \left(\frac{1}{z} - \frac{1}{z^3} + \frac{3}{z^5} \cdots \right)$$

that give very accurate answers when z is large. Note this looks like a few terms of a series in increasing powers of $\frac{1}{z}$.

9. **A Fourier synthesis formula.** The formula, valid for $x \in (\pi, \pi)$,

$$\frac{4}{\pi} \left(\sin(x) + \frac{1}{3}\sin(3x) + \frac{1}{5}\sin(5x) + \frac{1}{7}\sin(7x) + \ldots \right) = \begin{cases} 1 & \text{when } 0 < x < \pi, \\ 0 & \text{when } x = 0, \\ -1 & \text{when } -\pi < x < 0, \end{cases}$$

is an example of Fourier's synthesis formula for expressing functions as mixtures of trigonometric (sin and cosine) functions. Fourier needed such formulae when trying to solve certain differential equations.

Fourier's formula has a surprising feature: with a mixture of very nice infinitely differentiable functions $\sin(x), \sin(2x), \ldots$ you can create a function that is certainly not differentiable at $x = 0$ - there is a jump. Fourier claimed, but could not fully prove, that he could represent any function as such a mixture, and gave a recipe for doing so. He was essentially correct, and analogues of this type of formula went on to be used extensively throughout many areas of mathematics.

The book *'A radical approach to real analysis'*, by David Bressoud (see the moodle list of recommended books) blames this formula as 'the start of all the trouble' - as a trigger for the re-examination of calculus. Mathematicians took 80 years to re-develop calculus very rigorously, allowing them to resolve difficulties such as the truth of Fourier's formulae. At Warwick we will now drag you through this re-development in one year.

10. **Special functions.** You have already met the most familiar special functions $\exp(x), \sin(x), \cos(x), \log(x)$. You will encounter more in your degree, and while many of them started when investigating solutions of differential equations, they arise in many (or most) areas of mathematics, including combinatorics, statistics and probability. Two that are met early in statistics are the Gamma and Beta functions which are defined by integrals:

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx, \qquad B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx.$$

For general $p, q > 0$ these integrals cannot be evaluated in terms of the simpler special functions, but they arise so often that they deserve their own place as named functions.

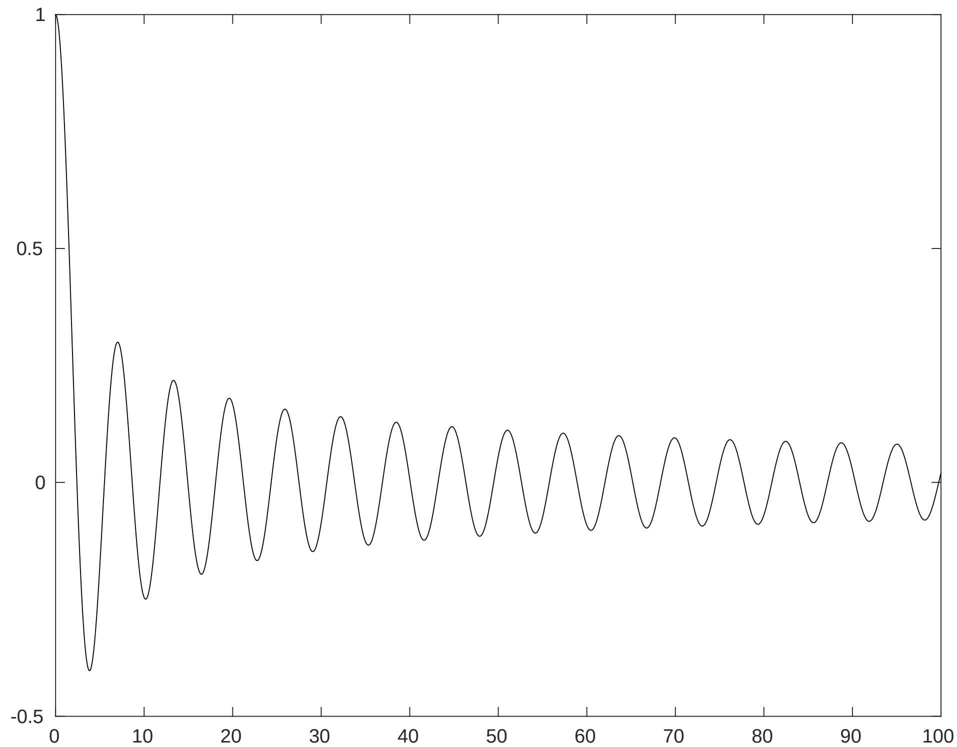
Here is one you will not have met: the Bessel function $J_0(x)$ (technically the 0th order Bessel function of the first kind). It was used to describe the shape of the surface of a vibrating circular drum. As is quite common, it can be defined in different but equivalent, ways: (i) it solves a simple differential equation ($x^2 \frac{d^2 J_0(x)}{dx^2} + x \frac{dJ_0(x)}{dx} + x^2 J_0(x) = 0$); (ii) it can be expressed as an integral $J_0(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sin(t)) dt$; (iii) it has an explicit formula as a power series

$$J_0(x) = \sum_{n=0}^\infty \frac{(-1)^n x^{2n}}{2^{2n}(n!)^2} = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^4(2!)^2} - \ldots$$

Power series is a key topic in Calculus 2, and the power series does give good approximations for small x, but it is not useful for describing the behaviour for large x. A good approximation for large x is

$$J_0(x) \approx \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{x}} \cos\left(x - \frac{\pi}{4}\right).$$

Can you see this in the picture below, plotted below for $x \in [0, 100]$?



Chapter 1. Limits

I have taken the material in this chapter from chapter 3 of last years lecture notes by Ian Melbourne et al., with a few things rearranged or omitted. We need to add a good number of pictures, which I will draw in lectures and which you could add in by hand in the margins, and I have marked places to do this with the symbol (P).

Sequences

A *sequence* is a list of numbers written in a definite order so that we know which number comes first, which number comes second, etc. Here are some simple examples: (P)

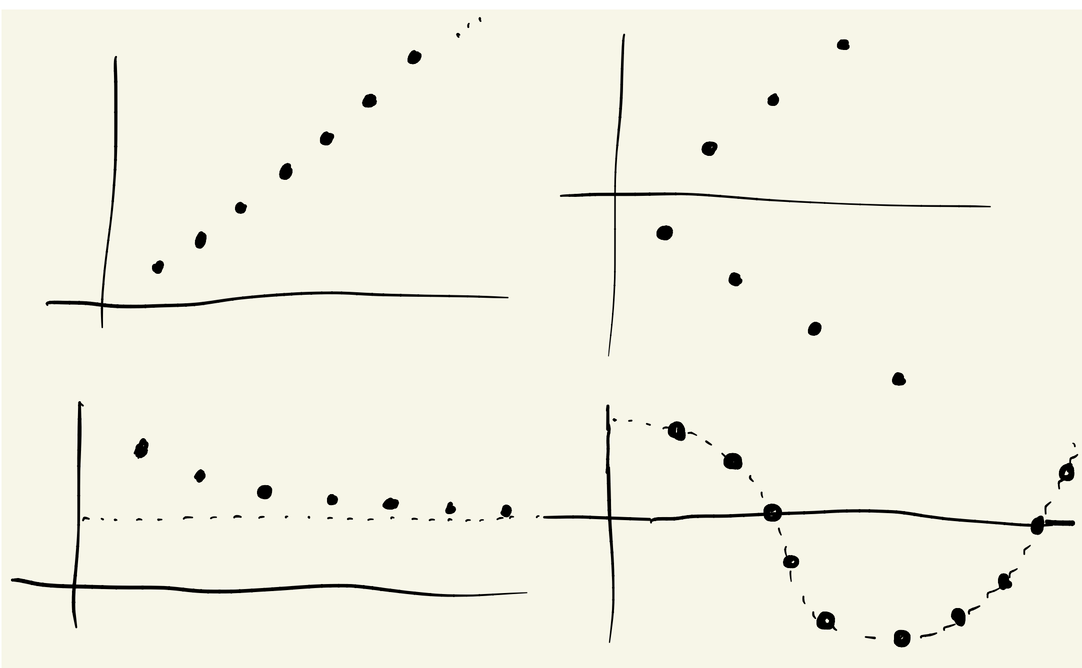
(i) $(1\,,2\,,3\,,4\,,5\,,\ldots)$

(ii) $(2\,,\frac{3}{2}\,,\frac{4}{3}\,,\frac{5}{4}\,,\frac{6}{5},\ldots)$

(iii) $(-1\,,2\,,-3\,,4\,,-5\,,\ldots)$

(iv) $(\cos\frac{\pi}{3}\,,\cos\frac{2\pi}{3}\,,\cos\pi\,,\cos\frac{4\pi}{3}\,,\cos\frac{5\pi}{3}\,,\ldots)$

Here is a rough sketch of the first terms of the four different sequences above :



In general, we denote a sequence by

$$(a_n : n \geq 1) \;=\; (a_1, a_2, a_3, a_4, a_5, \ldots).$$

A particular sequence can be defined by giving a formula for the nth term a_n; the four sequences considered above correspond to

$$(i)\ a_n = n, \quad (ii)\ a_n = \frac{n+1}{n}, \quad (iii)\ a_n = (-1)^n n, \quad (iv)\ a_n = \cos(n\pi/3).$$

We will be interested in describing how the elements of a sequence behave as n increases.

Definition. *A sequence $(a_n : n \geq 1)$ is:*

- ***strictly increasing*** *if $a_{n+1} > a_n$ for all $n \geq 1$;*
- ***increasing*** *if $a_{n+1} \geq a_n$ for all $n \geq 1$;*
- ***strictly decreasing*** *if $a_{n+1} < a_n$ for all $n \geq 1$;*
- ***decreasing*** *if $a_{n+1} \leq a_n$ for all $n \geq 1$;*
- ***monotonic*** *if it is increasing or decreasing or both;*

Example. *We can consider the four sequences listed above. We note that the sequence in (i) is strictly increasing. The sequence in (ii) is strictly decreasing. The sequences in (iii) and (iv) are not monotonic. The only sequences that are both increasing and decreasing are constant sequences, for example $(3, 3, 3, 3, \ldots)$.*

Definition. *A sequence $(a_n : n \geq 1)$ is:*

- ***bounded above*** *if there exists $M \in \mathbb{R}$ such that $a_n \leq M$ for all $n \geq 1$.*
  In this case, we say that M is an ***upper bound*** *for the sequence $(a_n : n \geq 1)$.*
- ***bounded below*** *if there exists $m \in \mathbb{R}$ such that $a_n \geq m$ for all $n \geq 1$.*
  In this case, we say that m is a ***lower bound*** *for the sequence $(a_n : n \geq 1)$.*
- ***bounded*** *if it is both bounded above and bounded below.*

Example. *We again consider the four sequences listed above. We note that the sequence in (i) is bounded below by 1 and is not bounded above. The sequence in (ii) is is bounded below by 1 and bounded above by 2. The sequence in (iii) is neither bounded below nor above. The sequence in (iv) is is bounded below by -1 and bounded above by 1.*

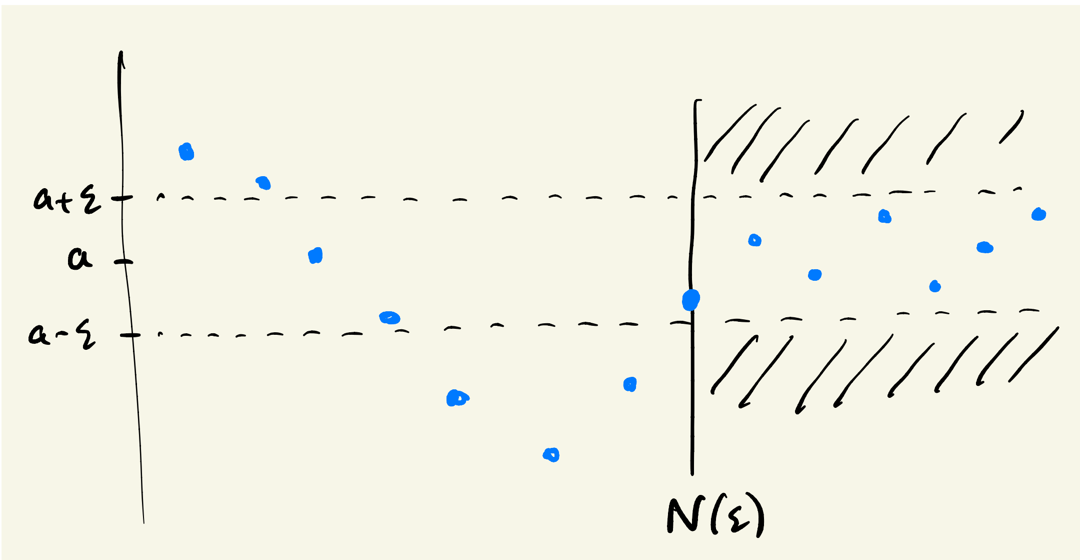
Example. *Consider the sequence defined by $a_n = \sqrt{n+1} - \sqrt{n}$. We claim the sequence $(a_n : n \geq 1)$ is bounded (see lectures).*

Convergence and divergence of sequences

We want to write down a careful definition of the idea that a sequence $(a_n : n \geq 1)$ approaches a limiting value. Here it is - the first key definition of the module.

Definition. *Let $a \in \mathbb{R}$. A sequence $(a_n : n \geq 1)$* ***has the limit a*** *if, for each $\epsilon > 0$, there exists N such that $|a_n - a| < \epsilon$ for all $n \geq N$.*

A sequence $(a_n : n \geq 1)$ is called ***convergent*** *if it has a limit $a \in \mathbf{R}$, and if it does not have a limit it is called* ***divergent****.*



There are a lot of alternative phrases, and several bits of notation, that all are used to mean that a sequence satisfies has the limit a: we can equivalently say that a_n **converges to** a or that a_n **tends to** a as n tends to infinity; as shorthand we often use either of the following notations

$$a_n \to a \text{ as } n \to \infty, \qquad \text{or} \qquad \lim_{n \to \infty} a_n = a.$$

Two simple ways for a sequence to be divergent are defined as follows.

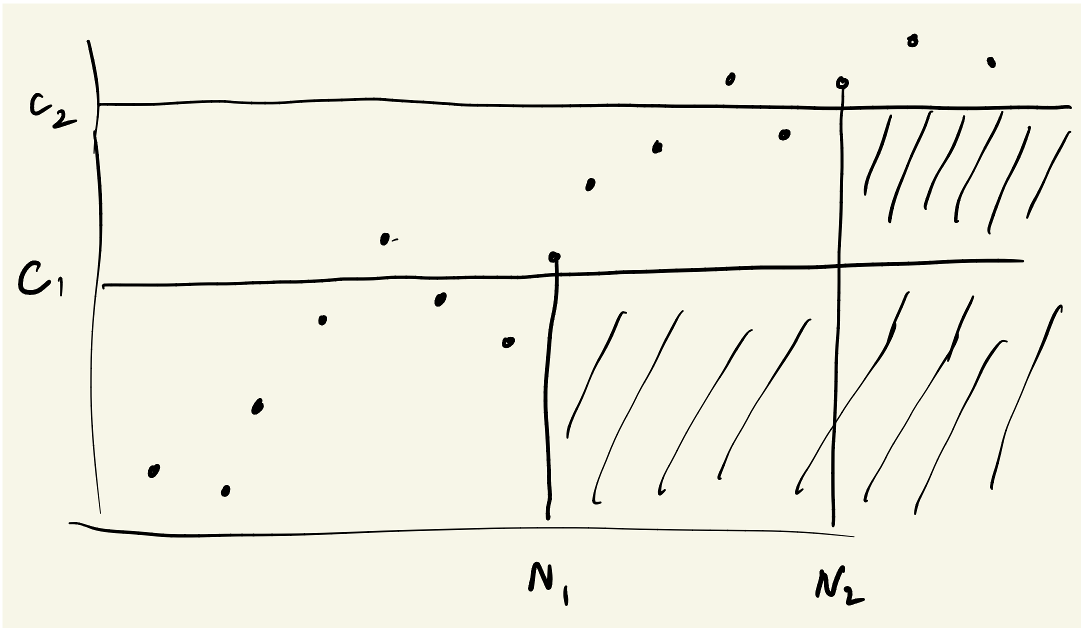
Definition. *A sequence $(a_n : n \geq 1)$* ***diverges to infinity*** *if, for every $C > 0$, there exists N such that $a_n > C$ whenever $n \geq N$.*

A sequence $(a_n : n \geq 1)$ ***diverges to minus infinity*** *if, for every $C < 0$, there exists N such that $a_n < C$ whenever $n \geq N$.*

We will use the shorthand notation

$$a_n \to \infty \text{ as } n \to \infty, \qquad \text{or} \qquad \lim_{n \to \infty} a_n = \infty.$$

and we replace ∞ with $-\infty$ when the sequence diverges to minus infinity.



Both definitions are a bit daunting, asking us to check infinitely many inequalities. We first check on some very simple examples, that the definitions are doing what we want.

Example. *(a)* $a_n = \frac{1}{\sqrt{n}} \to 0$ *as* $n \to \infty$.

(b) $a_n = \frac{n+2}{n} \to 1$ *as* $n \to \infty$.

(c) $a_n = n + (-1)^n \to \infty$ *as* $n \to \infty$.

(d) $a_n = n \sin \frac{n\pi}{2}$ *diverges.*

There is some rough work in italics below, followed by a full answer for (a),(b),(c).

Proof. (a) (*We are asked to consider when* $|a_n - 0| = \left|\frac{1}{\sqrt{n}}\right| < \epsilon$. *This inequality is equivalent to* $\sqrt{n} > \frac{1}{\epsilon}$ *or to* $n > \frac{1}{\epsilon^2}$)

Given $\epsilon > 0$, I choose an integer $N > \frac{1}{\epsilon^2}$. Then for $n \geq N$ we have $\sqrt{n} \geq \sqrt{N} > \frac{1}{\epsilon}$ and so $\frac{1}{\sqrt{n}} < \epsilon$. Since also $\frac{1}{\sqrt{n}} \geq 0$ we have $\left|\frac{1}{\sqrt{n}}\right| < \epsilon$. The definition is satisfied.

(b) (*We are asked to consider when* $|a_n - 1| = \left|\frac{n+2}{n} - 1\right| = \left|\frac{2}{n}\right| < \epsilon$ *for a value* $\epsilon > 0$. *This final inequality is equivalent to* $n > \frac{2}{\epsilon}$.)

Given $\epsilon > 0$, I choose an integer $N > \frac{2}{\epsilon}$. Then for $n \geq N$ we have $|a_n - 1| = \left|\frac{n+2}{n} - 1\right| = \left|\frac{2}{n}\right| < \epsilon$. The definition is satisfied.

(c) (*We are asked to consider when* $a_n = n + (-1)^n > C$ *for a value* $C > 0$. *Since* $(-1)^n$ *is either* $+1$ *or* -1 *this will be true once* $n > C + 1$.)

Given $C > 0$, I choose an integer $N > C + 1$. Then for $n \geq N$ we have $a_n = n + (-1)^n \geq n - 1 \geq N - 1 > C$. The definition is satisfied.

(d) (*The sequence is easy when written out* $(a_n : n \geq 1) = (1, 0, -3, 0, 5, 0, -7, \ldots)$. *Now it is obviously not converging - and nor is it diverging to* ∞ *or to* $-\infty$. *We are asked to rule out the possibility that* $a_n \to a$ *for any* $a \in \mathbf{R}$. How to start?).

We will discuss 'proofs by contradiction' in lectures. Proof omitted for now.

$\square$

Remark. *It is often easier to show the terms of sequence get small, than to show they get close to a non-zero number* a. *But checking that the sequence* $(a_1, a_2, a_3, \ldots)$ *converges to* a *is the same as checking that the sequence* $(a_1 - a, a_2 - a, a_3 - a, \ldots)$ *converges to* 0: *indeed in both cases we need to check for any* $\epsilon > 0$ *there exists* N *so that* $|a_n - a| < \epsilon$ *for* $n \geq N$. *In other words*

$$a_n \to a \text{ as } n \to \infty \quad \text{ is the same as } \quad a_n - a \to 0 \text{ as } n \to \infty.$$

A tiny remark, but we will use it over and over.

The two key workhorses for making limits easier to establish are the next two theorems: the **Algebra of Limits** and the **The Sandwich Theorem**. I hope they seem completely reasonable - that is they must be true or our definitions are probably not right. The proofs will be good examples of how to work with the definitions, and how to argue completely convincingly.

Theorem (Algebra of Limits)**.** *Let* $a, b \in \mathbb{R}$. *Suppose that* $a_n \to a$ *and* $b_n \to b$ *as* $n \to \infty$. *Then*

Sum rule: $a_n + b_n \to a + b$ *as* $n \to \infty$; *and* $ca_n \to ca$, *for any* $c \in \mathbf{R}$.

Product rule: $a_n b_n \to ab$ *as* $n \to \infty$.

Quotient rule: *provided* $b \neq 0$, $a_n/b_n \to a/b$ *as* $n \to \infty$.

Let's see it in action. We will be able to avoid ever choosing $\epsilon > 0$ or finding a suitable N. Rather we will build upon limits we already know, such as $\frac{1}{n} \to 0$ as $n \to \infty$.

Example. *Compute the limit* $\lim_{n\to\infty} a_n$ *where* $a_n = \dfrac{(n^2+1)(4n-3)}{2n^3+4}$.

Proof. We give the proof in full detail. On homework assignments and exams, one is not expected to give such a large amount of detail in all of the steps unless it is specifically asked for.

Dividing the numerator and denominator by n^3,

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

(We can already guess the answer is $(1\times 4)/2 = 2$ since everything else converges to 0.)

By the quotient rule,

$$\lim_{n\to\infty} a_n = \frac{\lim_{n\to\infty}\left[(1+\frac{1}{n^2})(4-\frac{3}{n})\right]}{\lim_{n\to\infty}\left[2+\frac{4}{n^3}\right]}\,.$$

By the product rule $\lim_{n\to\infty}\left[(1+\frac{1}{n^2})(4-\frac{3}{n})\right] = \lim_{n\to\infty}(1+\frac{1}{n^2})\lim_{n\to\infty}(4-\frac{3}{n})$.

We know $\frac{1}{n}\to 0$, so by the product rule $\frac{1}{n^2}\to 0$. Then using the sum rule several times

$$\lim_{n\to\infty} a_n = \frac{\left[1+\lim_{n\to\infty}\frac{1}{n^2}\right]\left[4-3\lim_{n\to\infty}\frac{1}{n}\right]}{\left[2+4\lim_{n\to\infty}\frac{1}{n^3}\right]} \;=\; \frac{(1+0)(4-3\cdot 0)}{2+4\cdot 0} \;=\; 2\,.$$

On a homework assignment or an exam, one could write a shorter version, as follows: 'By the sum, product and quotient rules,

$$\lim_{n\to\infty}\frac{(n^2+1)(4n-3)}{2n^3+4} \;=\; \lim_{n\to\infty}\frac{(1+\frac{1}{n^2})(4-\frac{3}{n})}{2+\frac{4}{n^3}} \;=\; \frac{(1+0)(4-3\cdot 0)}{2+4\cdot 0} \;=\; 2.$$

□

Proof. (Proof of Algebra of Limits). Since $a_n\to a$ we know there exists N_1 so that $|a_n-a|<\frac{1}{2}\epsilon$ for all $n\geq N_1$. Since $b_n\to b$ we know there exists N_2 so that $|b_n-b|<\frac{1}{2}\epsilon$ for all $n\geq N_2$. Now we set $N=\max\{N_1,N_2\}$. Then once $n\geq N$ we have

$$|(a_n+b_n)-(a+b)| = |(a_n-a)+(b_n-b)| \leq |a_n-a|+|b_n-b| < \frac{1}{2}\epsilon+\frac{1}{2}\epsilon = \epsilon$$

using the triangle inequality (that is $|x+y|\leq|x|+|y|$ for all $x,y\in\mathbf{R}$). We have proved the first part of the sum rule.

The second part, that $ca_n \to ca$ is trivial when $c = 0$. So we assume that $c \neq 0$. Then, using $a_n \to a$ we may find N so that $|a_n - a| < \epsilon/|c|$ for all $n \geq N$. Then

$$|ca_n - ca| = |c(a_n - a)| = |c|\,|a_n - a| < \epsilon \quad \text{for all } n \geq N.$$

We next prove a **special case** of the product rule, namely we suppose $a_n \to 0$ and $b_n \to 0$ as $n \to \infty$ and we will show that $a_n b_n \to 0$ as $n \to \infty$. We first note that, given $\epsilon > 0$, there exist $N_1, N_2 \geq 1$ such that $|a_n| < \sqrt{\epsilon}$ for $n \geq N_1$ and $|b_n| < \sqrt{\epsilon}$ for $n \geq N_2$. Again let $N = \max\{N_1, N_2\}$. For $n \geq N$ we have

$$|a_n b_n| \;=\; |a_n||b_n| \;<\; \sqrt{\epsilon}\sqrt{\epsilon} \;=\; \epsilon$$

showing that $a_n b_n \to 0$. We have finished the **special case**.

We will now show that the **general case** stated in the theorem follows from the **special case**. Indeed suppose $a_n \to a$ and $b_n \to b$ as $n \to \infty$. We use the identity (please check)

$$a_n b_n - ab = (a_n - a)(b_n - b) + a(b_n - b) + b(a_n - a).$$

We claim each of the three terms on the right hand side converges to 0. Indeed by the **special case** since $a_n - a \to 0$ and $b_n - b \to 0$ we know $(a_n - a)(b_n - b) \to 0$. By the sum rule we know that since $a_n - a \to 0$ then $b(a_n - a) \to 0$, and since $b_n - b \to 0$ then $a(b_n - b) \to 0$. Adding the three parts of the identity is another use of the sum rule (why can we add three limits?) and we have shown $a_n b_n - ab \to 0$ which is the product rule.

To prove the quotient rule, we we start by supposing $b_n \to b \neq 0$ and we will show that $\frac{1}{b_n} \to \frac{1}{b}$. Note that

$$\frac{1}{b_n} - \frac{1}{b} = \frac{b - b_n}{b b_n}.$$

Now $bb_n \to b^2$ by the sum rule. Taking $\epsilon_1 = b^2/2$, there exists $N_1 \geq 1$ such that $|bb_n - b^2| < b^2/2$ for all $n \geq N_1$. In particular, $-b^2/2 < bb_n - b^2$ so we deduce that $bb_n > b^2/2$ for all $n \geq N_1$. For $n \geq N_1$, it follows that

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| \;=\; \frac{|b - b_n|}{|bb_n|} \;\leq\; \frac{2}{b^2}|b - b_n|\,.$$

By the sum rule $\frac{2}{b^2}(b_n - b) \to 0$. Hence for any $\epsilon > 0$, there exists $N \geq N_1$ such that $\left|\frac{1}{b_n} - \frac{1}{b}\right| < \epsilon$ for all $n \geq N$. That is, $\frac{1}{b_n} \to \frac{1}{b}$.

To finish the proof of the quotient rule, we suppose $a_n \to a$ and $b_n \to b \neq 0$. Then $\frac{1}{b_n} \to \frac{1}{b}$ and so

$$\frac{a_n}{b_n} = a_n \frac{1}{b_n} \to \frac{a}{b}$$

by the product rule. Done. ∎

Theorem (Sandwich Theorem). *Suppose that $a_n \to L$ and $b_n \to L$ as $n \to \infty$, that is they both converge to the same limit $L \in \mathbb{R}$.*

If $a_n \leq c_n \leq b_n$ for all n, then $c_n \to L$ as $n \to \infty$.

Example. *Let $c_n = \frac{n+\sin(n^2+7)}{n}$. Since $-1 \leq \sin(x) \leq 1$ we immediately have*

$$a_n = 1 - \frac{1}{n} = \frac{n-1}{n} \leq c_n \leq \frac{n+1}{n} = 1 + \frac{1}{n} = b_n.$$

But we know $a_n \to 1$ and $b_n \to 1$ and so the Sandwich Theorem guarantees $c_n \to 1$ as $n \to \infty$ (and we never needed to get our hands messy with the sin(n^2+7) term).

Example. *Let $c_n = \sqrt{n+1} - \sqrt{n}$. Recall that in lectures we used the trick*

$$0 \ \leq \ c_n \ = \ \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} \ = \ \frac{1}{\sqrt{n+1} + \sqrt{n}} \ \leq \ \frac{1}{\sqrt{n}}.$$

Hence, we can take $a_n = 0$ and $c_n = \frac{1}{\sqrt{n}}$ in the Sandwich Theorem to see that $c_n \to 0$ as $n \to \infty$.

Proof. (Proof of Sandwich Theorem). Given $\epsilon > 0$, since $a_n \to L$ there exists N_1 so that $|a_n - L| < \epsilon$ when $n \geq N_1$. We use half of this, namely $L - \epsilon \leq a_n$. Since $c_n \to L$ there exists N_2 so that $|c_n - L| < \epsilon$ when $n \geq N_2$. We use half of this, namely $c_n \leq L + \epsilon$. Then for $n \geq \max\{N_1, N_2\}$ we have

$$L - \epsilon \leq a_n \leq b_n \leq c_n \leq L + \epsilon.$$

This shows $b_n \to L$ as $n \to \infty$. □

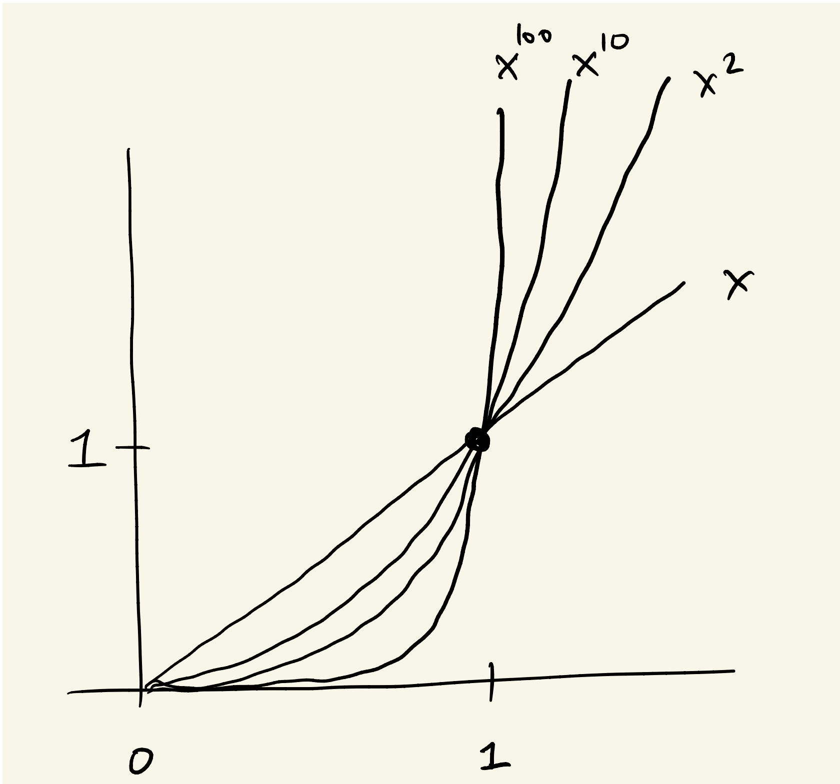
There are similar Algebra of Limits, and Sandwich Theorems for sequences diverging to infinity. Again, one feels they must be true. I have included just some proofs, again as an example of how to work with the definition of diverging to infinity. I'll put some others on the example sheet.

Lemma. *(A comparison lemma or a one sided sandwich). Let $(a_n : n \geq 1)$ and $(b_n : n \geq 1)$ be two sequences such that $b_n \geq a_n$ for all n.*

Suppose that $a_n \to \infty$ as $n \to \infty$. Then $b_n \to \infty$ as $n \to \infty$.

Proof. Let $C > 0$ be given. Since $a_n \to \infty$, it follows that there exists N such that $a_n > C$ whenever $n \geq N$. In particular, $b_n \geq a_n \geq C$ for $n \geq N$. Easy! □

We need a few more basic examples of simple limits with which to make our Sandwiches. We already know how to draw the polynomial functions $f(x) = x^N$ for $N = 1, 2, 3, \ldots$.



So we expect the following basic examples of geometric sequences.

Lemma.

$$\lim_{n\to\infty} x^n = \begin{cases} \infty & \textit{when } x > 1, \\ 1 & \textit{when } x = 1, \\ 0 & \textit{when } 0 \le x < 1. \end{cases}$$

Proof. Take $x > 1$. What about the following proof? We want to show that $x^n > C$ and taking logarithms this becomes $n\log(x) > \log(C)$. So if we take $N = \log(C)/\log(x)$ then $x^n > C$ for all $n \ge N$.

Very convincing. But most books will avoid this as it uses the logarithm function $\log(x)$ and the fact that it is strictly increasing. In most books there is an attempt to prove things in a logical order, starting from basic axioms and building slowly up. I will talk a bit about this in lectures (and see the handout on Proof Writing on the moodle page). We will meet the exponential function and its inverse the logarithm function near the end of the module when we can define $\exp(x)$ by a power series. Both $\exp(x)$ and $\log(x)$ will, of course, be strictly increasing where they are defined.

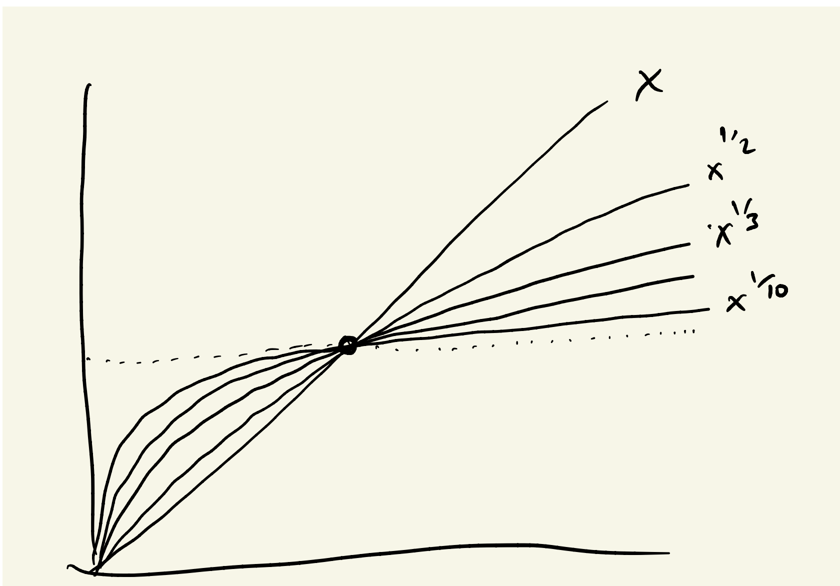
Here is an alternative proof that is commonly presented. For $x > 1$ we can write

$x = 1 + y$ where $y > 0$. Then

$$x^n = (1+y)^n = 1 + ny + \binom{n}{2} y^2 + \ldots + y^n > 1 + ny$$

using the binomial identity. Now we have a comparison with a sequence $1 + ny$ that we know diverges to infinity. I'll do the case case $x \in [0, 1)$ in lectures. $\square$.

Finally our last basic examples of simple limits. Here is my sketch of the nth root functions $f(x) = x^{1/n}$, for $n = 1, 2, 3, \ldots$.



So we expect the following basic limit for nth roots.

Lemma.

$$\lim_{n\to\infty} x^{1/n} = \begin{cases} 1 & \textit{when } x > 1, \\ 0 & \textit{when } x = 0. \end{cases}$$

Proof. There is another proof that just uses Binomial expansions as in the the last lemma - I will put this on the example sheet.

This lemma will be immediate to see by the end of the term, when we have logarithms and the fact that the exponential functions is continuous. Indeed the sequence $a_n = \frac{1}{n}\log(x)$ clearly converges to 0 ($x > 0$ is fixed so that $log(x)$ is just a fixed constant). Then continuity of the exponential will tell us that $x^{1/n} = \exp(a_n) \to \exp(0) = 1$ as $n \to \infty$. $\square$.

Example. *An old chestnut that appears every year: find* $\lim_{n\to\infty}(2^n+3^n)^{1/n}$.

$$3 \leq (2^n+3^n)^{1/n} \leq (3^n+3^n)^{1/n} = 2^{1/n}3.$$

The basic limit $2^{1/n} \to 1$ *and the Sandwich lemma implies that* $\lim_{n\to\infty}(2^n+3^n)^{1/n} = 3$.

One final simple property of limits, that is used over and over again. We give the proof as it is a nice example of proof by contradiction.

Lemma. *Suppose that* $a_n \to a$ *as* $n \to \infty$. *If* $a_n \geq 0$ *for all* $n \geq 1$, *then* $a \geq 0$.

Proof. We argue by contradiction: namely we suppose that $a < 0$. Since $a_n \to a$, given any $\epsilon > 0$ it follows that there exists $N \geq 1$ such that $|a_n - a| < \epsilon$ for $n \geq N$. I choose $\epsilon = -a > 0$. Then $|a_n - a| < \epsilon$ is equivalent to

$$a = -\epsilon < a_n - a < \epsilon = -a$$

implying that $a_n < 0$ - this is a contradiction. $\square$

Corollary. *Suppose that* $a_n \to a$ *and* $b_n \to b$ *as* $n \to \infty$. *If* $a_n \leq b_n$ *for all* n *then* $a \leq b$.

Proof. Define $c_n = b_n - a_n$. Then $c_n \geq 0$ and $c_n \to b - a$. So, by the lemma, $b - a \geq 0$. $\square$

There are lots of other small, completely reasonable (I am trying not to use the word obvious), lemmas that are good exercises in using the definitions. They are building blocks in a pyramid of results that end up helping us with the targets in Chapter 0. I'll ask you to construct careful proofs for some of them on the example sheet. Think of them now as good practice working with the definitions, or as verifying that the definitions are working.

Lemma. *The sequence* $(a_n : n \geq 1)$ *diverges to infinity if and only if the sequence* $(-a_n : n \geq 1)$ *diverges to minus infinity.*

Lemma. *Suppose the sequence* $(a_1, a_2, a_3, \ldots)$ *converges to* $a \in \mathbf{R}$. *Then the shifted sequence* $(a_{100}, a_{101}, a_{102}, \ldots)$, *which starts at the* 100*th term, also converges to* $a \in \mathbf{R}$.

Lemma. *If* $a_n \to \infty$ *then* $1/a_n \to 0$. *If* $a_n > 0$ *for all* n *and* $a_n \to 0$ *then* $1/a_n \to \infty$.

Lemma. *Every convergent sequence is bounded.*

Remark. *The definition of convergence or divergence of a sequence* $(a_n : n \geq 1)$ *only tells us something about the behaviour of a the terms* a_n *for large values of* n*. Changing the first* 100*, or* 1000*, values of the sequence won't change whether it converges or not. A property for a sequence* $(a_n : n \geq 1)$ *is said to hold* ***eventually*** *if it holds for all sufficiently large* n*. For example,* $(a_n : n \geq 1)$ *is eventually increasing if there exists* N *so that* $a_{n+1} \geq a_n$ *for all* $n \geq N$*. We can improve many of our results by asking that the hypotheses only hold eventually. For example:*

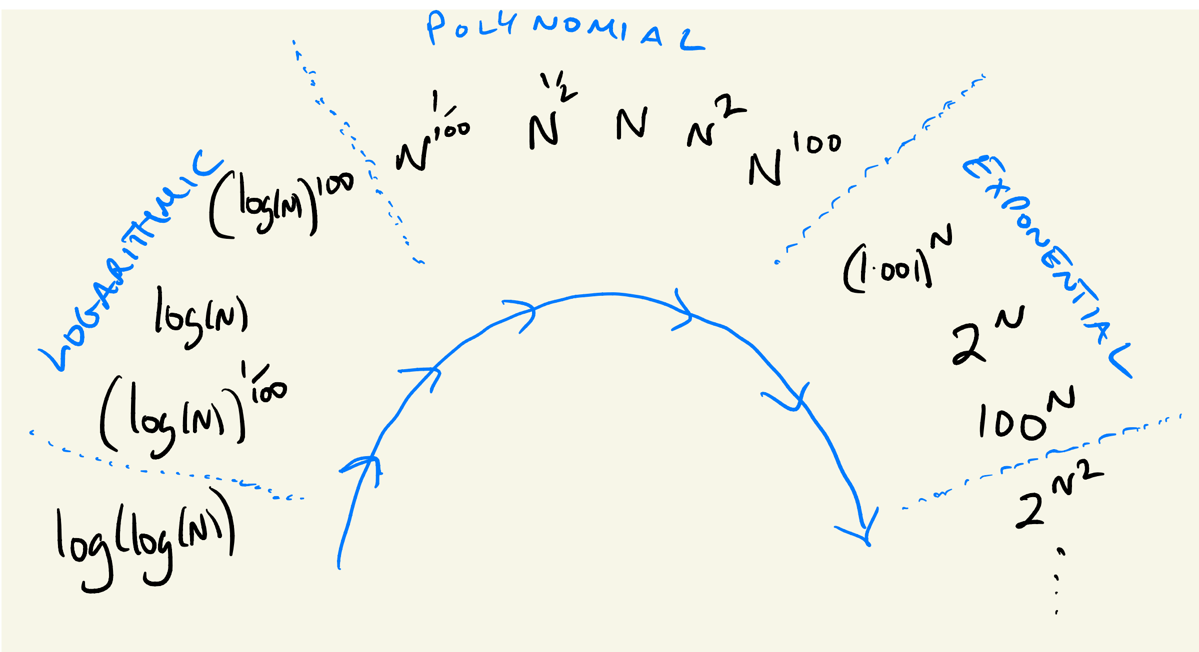
Lemma. *(A comparison lemma or a one sided sandwich). Let* $(a_n : n \geq 1)$ *and* $(b_n : n \geq 1)$ *be two sequences such that* $b_n \geq a_n$ *eventually.*

Suppose that $a_n \to \infty$ *as* $n \to \infty$*. Then* $b_n \to \infty$ *as* $n \to \infty$*.*

The proof hardly changes (but it needs somewhere to involve the N *when the comparison* $b_n \geq a_n$ *starts to hold). We won't halt to go back and improve all the results - but you might meet* ***eventually*** *versions in books.*

Scales of growth

Here is a list of sequences that diverge to ∞.



I have listed them so that if a sequence $(b_n : n \geq 1)$ is to the right of a sequence $(a_n : n \geq 1)$ then it 'grows faster'. So super-exponential growth is faster than exponential growth, which is faster than polynomial growth, which is faster than logarithmic growth. We will check this carefully below.

The list is not at all complete. Where does $2^{\sqrt{n}}$ fit in? Where does $n\log(n)$ fit in? I'll put some questions like this on the example sheet.

We do know where factorials $n!$ fit in: we met in lectures upper and lower bounds in the spirit of Stirling's approximation,

$$n^n e^{-n+1} \leq n! \leq n^{n+1} e^{-n+1}.$$

Check that these show that $n!$ grows faster than exponential, but much slower than 2^{n^2}.

By taking reciprocals we get scales of decay to zero:

$$\dots \quad \frac{1}{N!} \quad \frac{1}{100^N} \quad \frac{1}{2^N} \quad \frac{1}{N^{100}} \quad \frac{1}{N} \quad \frac{1}{N^{\frac{1}{100}}} \quad \frac{1}{\log(N)} \quad \frac{1}{\log(\log(N))} \quad \dots$$

So factorial decay if faster than exponential decay, which is faster than polynomial decay, which itself is faster than logarithmic decay.

We can make this a bit more precise with a bit of notation, as follows.

Definition. *For two sequences* $(a_n : n \geq 1)$ *and* $(b_n : n \geq 1)$ *of strictly positive numbers we write*

$$a_n = o(b_n) \quad \textit{when } \lim_{n\to\infty} \tfrac{a_n}{b_n} = 0.$$

We write

$$a_n = O(b_n) \quad \textit{when there exists } C \textit{ so that } a_n \leq C b_n \textit{ for all } n.$$

Informally $a_n = o(b_n)$ means that a_n grows slower than b_n, or decays faster than b_n; $a_n = O(b_n)$ means that a_n grows, or decays, at most as fast b_n. In words we say a_n is **little o** of b_n, or we say a_n is **big O** of b_n.

These two notations are very common is all areas of science. The computer science students will meet them when they study the speed of computer algorithms. For example the method *bubblesort* will sort a list of n numbers into increasing order using $O(n^2)$ operations; however the method *quicksort* will sort a list of n numbers into increasing order using $O(n \log(n))$ operations. We will use the notation to describe the speed of approach of a limit $a_n \to a$ by writing for example $a_n = a + O(2^{-n})$ when the error $|a_n - a|$ is $O(2^{-n})$ - a fast approximation, or $a_n = a + O(1/n)$ - a slow approximation.

Let's check that $n^p = o(c^n)$ for any $p > 0$ and $c > 1$ - that is polynomials grow slower than exponentials.

Lemma. *For any $p > 0$ and any $c > 1$*

$$\lim_{n\to\infty} \frac{n^p}{c^n} = 0.$$

Proof. Since $c > 1$ we can write $c = 1 + d$ where $d > 0$. Then

$$\frac{n}{c^n} = \frac{n}{(1+d)^n} = \frac{n^2}{1 + nd + \binom{n}{2}d^2 + \ldots + d^n} \leq \frac{n}{\binom{n}{2}d^2} = \frac{2n}{d^2 n(n-1)}.$$

We used the binomial expansion of $(1+d)^n$ here and the inequality follows since all the terms are positive. But we know $\frac{2n}{d^2 n(n-1)} \to 0$ so by the Sandwich Theorem we have proved $\lim_{n\to\infty} \frac{n}{c^n} = 0$ for any value of $c > 1$.

You can see how to adjust the proof to cope with $\lim_{n\to\infty} \frac{n^2}{c^n}$ or $\lim_{n\to\infty} \frac{n^3}{c^n}$ - just keep a different term in the binomial expansion. Alternatively, you can argue using the product rule that

$$\lim_{n\to\infty} \frac{n^2}{c^n} = \lim_{n\to\infty} \frac{n}{(\sqrt{c})^n} \lim_{n\to\infty} \frac{n}{(\sqrt{c})^n} = 0$$

since $\sqrt{c} > 1$. Repeating this (that is using induction) one finds $\lim_{n\to\infty} \frac{n^k}{c^n} = 0$ for any $k = 1, 2, \ldots$ and any $c > 1$.

Finally, if $p > 0$ is not an integer just sandwich $0 \leq \frac{n^p}{c^n} \leq \frac{n^k}{c^n}$ for an integer k larger than p. $\square$

Now we want to see that $\log(n) = o(n^p)$ for any $p > 0$ - that is logarithms grow slower than polynomials. Recall that we use natural logarithms throughout this module. We again will use properties of the logarithm function here although we won't carefully define it until chapter 4.

Lemma. *For any $p > 0$*

$$\lim_{n\to\infty} \frac{\log(n)}{n^p} = 0.$$

Proof. We know $\log(e^m) = m$. If $e^m \leq n \leq e^{m+1}$ then we have $n^p \geq e^{pm}$ and $\log(n) \leq \log(e^{m+1}) = m + 1$. Hence, for such n we have

$$0 \leq \frac{\log(n)}{n^p} \leq \frac{m+1}{e^{pm}}.$$

But we know that $\lim_{m\to\infty} \frac{m+1}{e^{pm}} = 0$, since we just checked that exponentials grow faster than polynomials. So, given $\epsilon > 0$ we can find M so that

$$\tfrac{m+1}{e^{pm}} < \epsilon \text{ for all } m \geq M.$$

This implies that

$$\tfrac{\log(n)}{n^p} < \epsilon \text{ for all } n \geq e^M.$$

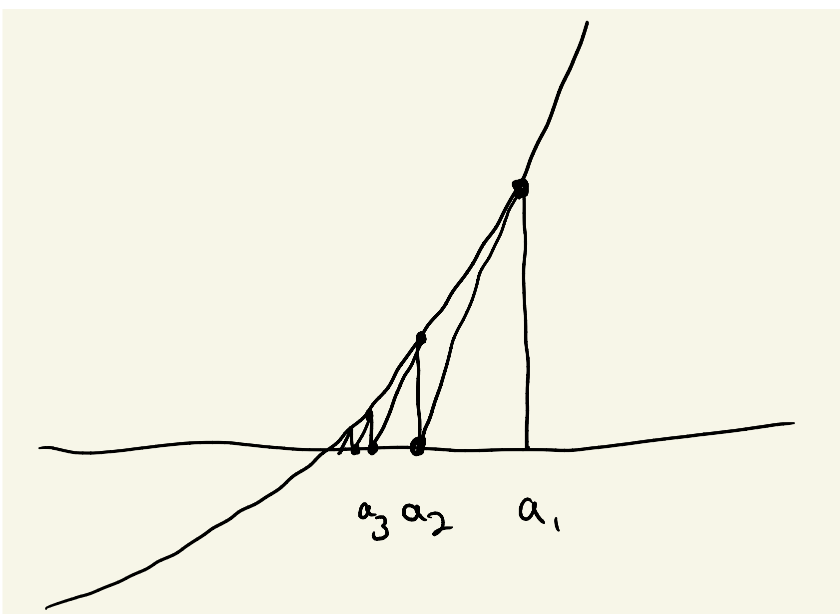
□

Chapter 2. Completeness

We recall the sequences that arise from two well used methods for root finding.

The Newton Raphson method for finding good approximations to a solution r to $f(r) = 0$ is defined by the iteration (in lectures we re-derived this formula)

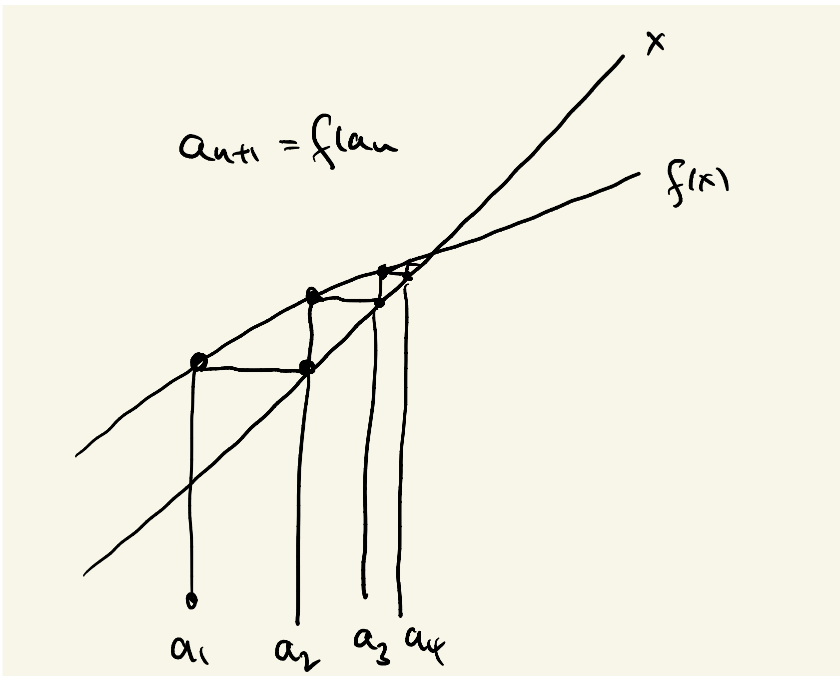
$$a_1 = \text{a good first approximation}, \qquad a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}.$$



In this picture it looks like we started with a_1 larger than the root r satisfying $f(r) = 0$, and it looks like we should get a decreasing sequence $a_1 > a_2 > a_3 > \ldots$ that converges to the root. I sketched a function that is convex, that is $f''(x) > 0$ near the root. If the function is close to being linear at the root (that is f'' is small) then the algorithm is fast (why?). On the example sheets I ask you to sketch the situations (i) the initial guess $a_1 < r$; (ii) where $f''(x) < 0$ near the root; (iii) where the method fails to converge.

A simpler method, that is perhaps useful when it is not possible to calculate f' easily, is just called iteration. The aim to approximate a solution to $f(r) = r$. One simply uses

$$a_1 = \text{a good first approximation}, \quad a_{n+1} = f(a_n).$$



So $a_2 = f(a_1)$, $a_3 = f(f(a_1))$, $a_3 = f(f(f(a_1)))$ ... In this picture we started with a_1 smaller than the root r satisfying $f(r) = 0$, and it looks like we should get a increasing sequence $a_1 < a_2 < a_3 < \dots$ that converges to the root. I sketched a function that satisfied, that is $0 < f'(x) < 1$ near the root. If f is nearly constant at the root, that is f' is small, then the algorithm is fast (why?). On the example sheet I ask you to sketch the situations (i) the initial guess $a_1 > r$; (ii) where $-1 < f'(x) < 0$ near the root (you will see why these diagrams are called cobweb diagrams); (iii) situations where there is no convergence to the root.

Example. *Show the sequence defined by $a_1 = 1$ and $a_{n+1} = \sqrt{2 + a_n}$ converges to 2.*

Sketching the function $f(x) = \sqrt{2+x}$ and $f(x) = x$ it all looks like the picture above. We can guess that $(a_n : n \geq 1)$ should be increasing and that $1 \leq a_n \leq 2$ for all n. It is easy to check both statements by induction. Suppose $1 \leq a_k \leq 2$ for $k = 1, 2, \dots, n$; then

$$a_{n+1} = \sqrt{2 + a_n} \leq \sqrt{2+2} = 2, \quad \textit{and} \quad a_{n+1} = \sqrt{2 + a_n} \geq \sqrt{2+1} \geq 1.$$

Suppose $a_1 \leq a_2 \leq \ldots \leq a_n$; then then

$$a_{n+1} = \sqrt{2 + a_n} \geq \sqrt{2 + a_{n_1}} = a_n.$$

We can directly try anc calculate how close a_n gets to the limit 2.

$$\begin{aligned} 2 - a_{n+1} &= 2 - \sqrt{2 + a_n} \\ &= \frac{2 - a_n}{2 + \sqrt{2 + a_n}} \\ &\leq \frac{2 - a_n}{2 + \sqrt{2 + 1}}. \end{aligned}$$

Thus the error $\mathcal{E}_n$ defined by $\mathcal{E}_n = 2 - a_n$ satisfies $\mathcal{E}_{n+1} \leq c\mathcal{E}_n$ where $c = 1/(2+\sqrt{3}) \approx 0.267$. Since $\mathcal{E}_1 = 2 - 1 = 1$ we find $\mathcal{E}_2 \leq c$, $\mathcal{E}_3 \leq c^2$, ... and in general $\mathcal{E}_n \leq c^{n-1}$. The errors are converging exponentially fast to zero.

Example. *Aim: estimate the speed of convergence for the Newton-Raphson algorithm*

$$a_1 = 3/2 \text{ and } a_{n+1} = \tfrac{a_n}{2} + \tfrac{1}{a_n}$$

to its limit $\sqrt{2}$.

This is the Newton-Raphson algorithm for finding the root $\sqrt{2}$ of $f(x) = x^2 - 2 = 0$. We directly try and calculate how close a_n is to the limit 2.

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

This guarantees us that $a_n \geq \sqrt{2}$ for all n (yes - by induction). Also it shows us the error $\mathcal{E}_n$ defined by $\mathcal{E}_n = 2 - a_n$ satisfies

$$\mathcal{E}_{n+1} \leq \frac{\mathcal{E}_n^2}{2a_n} \leq \frac{\mathcal{E}_n^2}{2\sqrt{2}} \leq \mathcal{E}_n^2.$$

where the final two inequalities I have used the fact the that the denominator $2a_n$ is greater than $2\sqrt{2}$ and then recklessly thrown away $\frac{1}{2\sqrt{2}}$.

Now $\mathcal{E}_1 = \frac{3}{2} - \sqrt{2} \leq \frac{1}{10}$. So $\mathcal{E}_2 \leq \frac{1}{(10)^2}$ and then $\mathcal{E}_3 \leq \frac{1}{(10)^4}$. Iterating I get

$$\mathcal{E}_n \leq \frac{1}{(10)^{2^{n-1}}}.$$

This is faster than exponential decay. Running the algorithm one more step, from a_n to a_{n+1}, the error has changed from $\frac{1}{(10)^{2^{n-1}}}$ to $\frac{1}{(10)^{2^n}}$ - that is the number of accurate decimal places has roughly doubled. (If I hadn't thrown away the $\frac{1}{2\sqrt{2}}$ factor I could have shown the error is even smaller - see the example sheet.)

Both these root finding algorithms above are useful when we do not know a simple explicit formula for the root. This chapter is all about how to establish that a sequence converges when you do not know exactly what the limit is. Here is a very believable first result in this direction.

Theorem. *The Weierstrass Criterion*

A sequence $(a_n : n \geq 1)$ that is increasing and bounded above must converge.

Recall that a monotone sequence is either decreasing or increasing. I have stated the criterion just for increasing ones, but there is an immediate corollary:

A sequence $(a_n : n \geq 1)$ that is decreasing and bounded below must converge.

Indeed if $(a_n : n \geq 1)$ that is decreasing and bounded below then $(-a_n : n \geq 1)$ is increasing and bounded above. The theorem guarantees that $-a_n \to a$ for some a, and then we know $a_n \to -a$ as $n \to \infty$.

This result is so useful it deserves to be called a Theorem. It is worth repeating, spelling out the hypotheses and the conclusion.

Hypotheses:

$(a_n : n \geq 1)$ is increasing, that is $a_1 \leq a_2 \leq a_3 \leq \ldots$

$(a_n : n \geq 1)$ is bounded above, that is there exists $U \in \mathbf{R}$ so that $a_n \leq U$ for all n.

Conclusion:

The limit $\lim_{n\to\infty} a_n = a \in \mathbf{R}$ must exist.

The result is entirely believable - indeed I can tell you what the limit must be. It must be a number a that lies above all the values $a_1, a_2, \ldots$ but it must be the smallest such number. In other words the limit a will be an upper bound for the sequence $(a_n : n \geq 1)$ but it should be the smallest such upper bound.

Amazingly, one cannot prove this Theorem with a simple arithmetic proof - one that only uses the rules for addition/multiplication and the simple properties of inequalities. Indeed in lectures I will waffle about Planet Rational, where they only believe in rational numbers, and where despite agreeing with our definition of convergence this theorem is false in their world. Their problem is that the sequence $a_1, a_2, \ldots$ might be a sequence of increasing rationals that we can see converges to $\sqrt{2}$, but for them $\sqrt{2}$ does not exist so there is no limit. This leads to an interesting discussion about axioms, proofs, models... but I will not go further in the lecture notes, but discuss it a bit more in lectures and in a handout.

To get a proof of the Theorem, which we do in the next section, we will write down a basic natural property, an axiom, for the real numbers that we can appeal to at any time, one that tries to encapsulate that there are no holes - hence it is called **completeness**. At various points we will see we that require this property and I will probably waive my hands up and down in lectures.

Let's look at some examples of the Theorem in use first.

Key example The sequence $a_n = \left(1 + \frac{1}{n}\right)^n$ converges. This has been a motivating example for us and I tried in lectures to convince you (bank interest sequence...) that this should be increasing. Here is one standard argument to check that $(a_n : n \geq 1)$ is both increasing and bounded. The idea is to examine carefully the Binomial expansions for a_n and a_{n+1}.

$$\begin{aligned}
\left(1 + \frac{1}{n}\right)^n &= 1 + n\,\frac{1}{n} + \binom{n}{2}\frac{1}{n^2} + \ldots + \binom{n}{k}\frac{1}{n^k} + \ldots + \frac{1}{n^n}, \\
\left(1 + \frac{1}{n+1}\right)^{n+1} &= 1 + (n+1)\,\frac{1}{n+1} + \ldots + \binom{n+1}{k}\frac{1}{(n+1)^k} + \ldots + \frac{1}{(n+1)^{n+1}}.
\end{aligned}$$

The first expansion has $n+1$ terms and the second expansion has $n+2$ terms. The k-th term in the first expansion for a_n is

$$\binom{n}{k}\frac{1}{n^k} = \frac{n(n-1)\ldots(n-k+1)}{n^k k!} = \frac{1}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{k-1}{n}\right).$$

The k-th term in the second expansion for a_{n+1} is

$$\begin{aligned}
\binom{n+1}{k}\frac{1}{(n+1)^k} &= \frac{(n+1)n(n-1)\ldots(n-k)}{(n+1)^k k!} \\
&= \frac{1}{k!}\left(1 - \frac{1}{n+1}\right)\left(1 - \frac{2}{n+1}\right)\ldots\left(1 - \frac{k-1}{n+1}\right).
\end{aligned}$$

So each of the first n terms in the expansion of a_{n+1} is larger than the corresponding term for a_n. Also there is an extra term $\frac{1}{(n+1)^{n+1}} > 0$ since the expansion for a_{n+1} is one term longer. We conclude $a_{n+1} > a_n$. Phew! There is a simpler way when we have explored logarithms and derivatives as we will be ale to show more simply that the functions $x \log\left(1+\frac{1}{x}\right)$ and its exponential $\left(1+\frac{1}{x}\right)^x$ are increasing for $x > 0$ (by checking the derivative is positive).

Now we check that the sequence $a_n = \left(1+\frac{1}{n}\right)^n$ is bounded. In the expansion for a_n above we can bound the kth term by

$$\binom{n}{k}\frac{1}{n^k} = \frac{1}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right) \leq \frac{1}{k!} \leq \frac{1}{2^{k-1}}$$

since $k! = k(k-1)\dots 2 \geq 2.2.\dots.2$. Then

$$\begin{aligned} a_n &\leq 1+1+\frac{1}{2!}+\frac{1}{3!}+\dots+\frac{1}{n!} \\ &\leq 1+1+\frac{1}{2}+\frac{1}{2^2}+\dots+\frac{1}{2^{n-1}} \leq 3 \end{aligned}$$

where we know the sum of a geometric series $1+\frac{1}{2}+\dots+\frac{1}{2^{n-1}} = (1-\frac{1}{2^n})/(1-\frac{1}{2}) \leq 2$.

Now we can apply the Weierstrass Criterion to conclude $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$ exists.

Definition. *We define the number e to be the limit* $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$.

This is traditionally the way to define the number e in this module. Later we might prefer other ways, for example via the series expansion

$$e = \sum_{k=0}^{\infty}\frac{1}{k!}$$

or, perhaps best of all,

$e = x(1)$ where $(x(t) : t \geq 0)$ is the unique solution to the differential equation $\frac{dx}{dt} = x$ for $t \in \mathbf{R}$ with initial condition $x(0) = 1$.

Of course, we need to check all of these define the same number.

The completeness axiom

We go slowly in our search for least upper bounds. We start by extending the definitions of upper bounds and lower bounds for sequences to arbitrary subsets of $\mathbf{R}$

Definition. *Let $A \subset \mathbb{R}$ be a non-empty set. The set A is called*

(i) ***bounded above*** *if there exists $M \in \mathbb{R}$ such that $a \leq M$ for all $a \in A$, in which case M is called an* ***upper bound*** *for A;*

(ii) ***bounded below*** *if there exists* $m \in \mathbb{R}$ *such that* $a \geq m$ *for all* $a \in A$*, in which case* m *is called a* ***lower bound*** *for* A*;*

(iii) ***bounded*** *if it is bounded above and below.*

Upper and lower bounds are not uniquely determined. If M is an upper bound for A, then any $M' \geq M$ is also an upper bound for A. Similarly if m is a lower bound for A, then any $m' \leq m$ is also a lower bound for A. The 'best possible' bounds are defined as follows.

Definition. *Let* $A \subset \mathbf{R}$ *be a non-empty set bounded from above. A number* $U \in \mathbf{R}$ *is called the* ***supremum*** *or the* ***least upper bound*** *of* A*, and we write* $U = \sup A$*, if the following properties hold:*

(i) U *is an upper bound of* A *and (ii) if* M *is any upper bound of* A*, then* $M \geq U$*.*

Let $A \subset \mathbf{R}$ *be a non-empty set bounded from below. A number* $L \in \mathbf{R}$ *is called the* ***infimum*** *or the* ***greatest lower bound*** *of* A*, and we write* $L = \inf A$*, if the following properties hold:*

(i) L *is a lower bound of* A *and (ii) if* m *is any lower bound of* A*, then* $m \leq L$*.*

Convention

If A *is non-empty but is not bounded from above, we set* $\sup A = \infty$*.*
If A *is non-empty but is not bounded from below, we set* $\inf A = -\infty$*.*

Remark. $\sup A$ *and* $\inf A$ *do not necessarily belong to* A*. Consider* $A = \{x : 0 < x < 1\}$*. Then* $\inf A = 0$ *and* $\sup A = 1$*, neither of which belong to* A*. If* $\sup A \in A$*, we say that* A *has a maximal element and we can write* $\max A$*. Likewise, if* $\inf A \in A$*, we say that* A *has a minimal element and we write* $\inf A$ *also as* $\min A$*. For example, consider* $A = \{x : 0 \leq x \leq 1\}$*. Here* $\min A = \inf A = 0$ *and* $\max A = \sup A = 1$*.*

Example. *Let* $A = \{\frac{1}{n} : n \geq 1\}$*. Then* $\sup A = \max A = 1$ *and* $\inf A = 0$*. There is no minimum element.*

Example. *Let* $A = \{x : x^2 < 2\}$*. Then* $\sup A = \sqrt{2}$*, there is no maximum element, and* $\inf A = \min A = 0$*.*

Example. *Let* $A = \{x : x^3 \leq 2\}$*. Then* $\sup A = \max A = 2^{1/3}$ *and* $\inf A = -\infty$*.*

Example. *Let* $A = \left\{\frac{x}{1+x^2} : x > 0\right\}$*. Then* $\sup A = \max A = \frac{1}{2}$ *(plot the graph - there is a single critical point at* $x = 1$*);* $\inf A = 0$ *and no minimum element.*

You will meet sup(A) and inf(A) being used in many of your maths modules. I will give some glimpses in lectures.

Completeness Axiom for R

A non-empty subset $A \subseteq \mathbb{R}$ that is bounded from above has a least upper bound $\sup A$.

A non-empty subset $A \subseteq \mathbb{R}$ that is bounded from below has a greatest lower bound $\inf A$.

When working with a supremum we often use the following simple lemma.

Lemma. ***A supremum is always nearly attained.***
Let A be a non-empty subset which is bounded from above. For every $\epsilon > 0$, there exists $a \in A$ such that

$$\sup A - \epsilon < a \leq \sup A\,.$$

Proof. The fact that $a \leq \sup A$ for all $a \in A$ is immediate, since $\sup A$ is an upper bound for A. It remains to show that for all $\epsilon > 0$ there exists $a \in A$ satisfying $a > \sup A - \epsilon$. Argue by contradiction: suppose there exists $\epsilon > 0$ such that there is no $a \in A$ such that $a > \sup A - \epsilon$. Then $a \leq \sup A - \epsilon$ for all $a \in A$. Hence $\sup A - \epsilon$ is an upper bound of A. But $\sup A - \epsilon < \sup A$, which contradicts the fact that $\sup A$ is the least upper bound of A. □

Now we can return and finish the proof of the Weierstrass Criterion. We restate it here.

Theorem. ***The Weierstrass Criterion.***
A sequence $(a_n : n \geq 1)$ that is increasing and bounded above must converge.

Proof. Let $A = \{a_1, a_2, a_3, \ldots\} = \{a_n : n \geq 1\}$. Since $(a_n : n \geq 1)$ is bounded the set A is bounded. Clearly A is non-empty (since $a_1 \in A$). Hence by the Completeness Axiom $a = \sup A$ is a well-defined real number. We will now check that $a_n \to a$ as $n \to \infty$.

Let $\epsilon > 0$. By the lemma above there exists $N \geq 1$ such that $a - \epsilon \leq a_N$. Since a_n is increasing we have $a - \epsilon \leq a_n$ for all $n \geq N$. Since a is an upper bound for $A = \{a_1, a_2, \ldots\}$ we have $a_n \leq a < a + \epsilon$ for all n. We have shown that $|a_n - a| < \epsilon$ for $n \geq N$ and hence proved that $a_n \to a$. □

Example. *Consider the sequence $(a_n : n \geq 1)$ given recursively by*

$$a_1 = 1, \quad a_{n+1} = \tfrac{1}{5}(a_n^2 + 6) \text{ for } n \geq 1\,.$$

Show that the sequence a_n has the following properties:

$$\textit{(i)}\ \ 0 \le a_n < 2 \textit{ for all } n; \quad \textit{(ii) } a_n \textit{ is increasing.}$$

It follows from the Weierstrass Criterion that $a = \lim_{n\to\infty} a_n$ exists.

Now (iii) compute a.

Solution. (i) We show by induction on n that $0 \le a_n < 2$ for all n. The base case $n = 1$ holds since $a_1 = 1$. Suppose now that $a_n \in [0, 2)$. We want to show that $a_{n+1} \in [0, 2)$. Since $a_n \ge 0$, it is immediate that $a_{n+1} \ge \frac{1}{5} \cdot 6 \ge 0$. Furthermore, since $0 \le a_n < 2$, it follows that $a_n^2 < 4$ and hence

$$a_{n+1} = \tfrac{1}{5}(a_n^2 + 6) < \tfrac{1}{5}(4 + 6) = 2\,.$$

Hence, $a_{n+1} \in [0, 2)$, as was claimed.

(ii) We need to show that

$$\tfrac{1}{5}(a_n^2 + 6) \ge a_n$$

for all n. Equivalently,

$$(a_n - 2)(a_n - 3) = a_n^2 - 5a_n + 6 \ge 0.$$

By part (i), $a_n - 2 < 0$ and $a_n - 3 < 0$ so the inequality holds for all n.

Let $a = \lim a_n$. Then the shifted sequence $a_{n+1} \to a$ also and by the product and sum rule for limits

$$\tfrac{1}{5}(a_n^2 + 6) \to \tfrac{1}{5}(a^2 + 6).$$

Hence, taking limits on both sides of the equation $a_{n+1} = \frac{1}{5}(a_n^2 + 6)$ we find

$$a = \tfrac{1}{5}(a^2 + 6)\,.$$

The solutions of this equation are $a = 2$ and $a = 3$. But we already know that $a_n < 2$ and so the limit $a \le 2$. Hence $a = 2$ must be the limit. □

The Cauchy property

For an increasing sequence $(a_n : n \ge 1)$ we have a simple way of checking convergence - we just need to check it is bounded above. We do not need to know what the limit is before we start. Similarly for decreasing sequences we need to check if they are bounded below. For a general sequence, that is not increasing or decreasing, is there a method for checking convergence? Yes - it is called the Cauchy property.

Definition. *A sequence* $(a_n : n \geq 1)$ *has the* ***Cauchy Property*** *if, given* $\epsilon > 0$*, there exists* $N \geq 1$ *such that*

$$|a_n - a_m| < \epsilon \quad \textit{for all } n, m \geq 1 \textit{ with } n \geq N \textit{ and } m \geq N.$$

This is a bit harder than the definition of a limit, as we have to check all possible pairs of terms, a_n and a_m, and show they get close. Informally, this means that the terms 'get closer to each other' the further we go in the sequence. I find it difficult to draw a useful picture - suggestions welcome.

This property is harder to check than just boundedness above or below. However the theorem below shows that it is an *exact* test for convergence, and it still has the big advantage that you do not need to know what the limit is before you start - it guarantees a limit a must exist.

Theorem. ***Cauchy's Criterion.*** *A sequence* $(a_n : n \geq 1)$ *has the Cauchy property if and only if it is convergent.*

You will meet the Cauchy property many times in higher level modules, and its extension to functions $(f_n : n \geq 1)$ satisfying a Cauchy property and random variables $(X_n : n \geq 1)$ satisfying a Cauchy property. All these extensions have their root in this theorem - a highlight of the module. This chapter will end with two proofs of this theorem, but neither is very short. Let's see it in action before we work through the proofs.

Example. ***Contracting sequences.***

We call a sequence $(a_n : n \geq 1)$ *contracting* if the exists a contraction factor $0 < \gamma < 1$ so that

$$|a_{n+1} - a_n| \leq \gamma |a_n - a_{n-1}| \quad \text{for all } n \geq 2.$$

In words, the successive differences shrink by at least γ. This property implies that $(a_n : n \geq 1)$ has the Cauchy property - here is why. We have $|a_3 - a_2| \leq \gamma |a_2 - a_1|$ and $|a_4 - a_4| \leq \gamma |a_3 - a_2| \leq \gamma^2 |a_2 - a_1|$. By induction we get

$$|a_{n+1} - a_n| \leq \gamma^{n-1} |a_2 - a_1| \quad \text{for all } n.$$

Suppose $m < n$. We use the triangle inequality repeatedly to see

$$\begin{aligned} |a_n - a_m| &= |a_n - a_{n-1} + a_{n-1} - a_m| \\ &\leq |a_n - a_{n-1}| + |a_{n-1} - a_m| \\ &\leq \ldots \\ &\leq |a_n - a_{n-1}| + |a_{n-1} - a_{n-2}| + \ldots + |a_{m+1} - a_m|. \end{aligned}$$

Substitute in the estimate just above to get

$$|a_n - a_m| \quad \leq \quad |a_2 - a_1| \left(\gamma^{n-2} + \gamma^{n-3} + \ldots + \gamma^{m-1}\right).$$

But the right hand side is a geometric series which we can sum:

$$|a_n - a_m| \quad \leq \quad |a_2 - a_1| \frac{\gamma^{m-1}}{1-\gamma}.$$

I have taken the sum if the infinite geometric series here as it will be an upper bound for any n. Now we have the Cauchy property, since $\gamma^m \to 0$ as $m \to \infty$. Indeed for any $\epsilon > 0$ I can choose M so that

$$|a_2 - a_1| \frac{\gamma^{M-1}}{1-\gamma} \leq \epsilon$$

and then

$$|a_n - a_m| \leq \epsilon \quad \text{for all } m, n \geq M.$$

I think checking the contraction property is often easier than directly checking the Cauchy property. I will show in lectures (by glimpsing forward and using some Calculus) that sequences coming from iterated function algorithm often have the contraction property.

Now the rest of this chapter is proofs - and probably the hardest proofs in the module. Buckle up.

Proof. **First proof of Cauchy's Criterion.** The easy part is to show that a convergent sequence has the Cauchy property. The harder, but useful, direction is to show that if a sequence has the Cauchy property then it must converge.

Easy part. Suppose that $(a_n : n \geq 1)$ is convergent. Denote its limit by a. Given $\epsilon > 0$, there exists $N \geq 1$ such that $|a_n - a| < \epsilon/2$ whenever $n \geq N$. In particular, by the triangle inequality, for all $m, n \geq N$,

$$|a_n - a_m| \leq |a_n - a| + |a - a_m| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Therefore $(a_n : n \geq 1)$ has the Cauchy property. Done.

Hard part. Now we suppose $(a_n : n \geq 1)$ has the Cauchy property, and we must prove it is convergent. We will warm up by checking first that $(a_n : n \geq 1)$ must be bounded. We choose $\epsilon = 1$ in the definition of the Cauchy property. There exists N so that $|a_n - a_m| < 1$ for all $n, m \geq N$. In particular, choosing $m = N$,

$$a_N - 1 < a_n < a_N + 1 \quad \text{for all } n \geq N.$$

So the sequence $(a_N, a_{N+1}, a_{N+2}, \ldots)$ is bounded above and below. But as we saw in chapter 1 the first $N-1$ terms will not prevent the whole sequence being bounded. We write U for an upper bound and L for a lower bound.

Now to show $(a_n :\geq 1)$ is convergent we will (i) write down a formula for the limit a and (ii) prove that the sequence converges to a.

The Cauchy property tells us that eventually the sequence doesn't 'vary' that much. Consider the tail of the sequence, the subsequence $(a_n, a_{n+1}, a_{n+2}, \ldots)$ from term number n onwards. This is bounded by U and so has a least upper bound

$$U_n = \sup\{a_n, a_{n+1}, a_{n+2}, \ldots\}.$$

The intuition is that U_n should get closer and closer to the limit as n grows. As we look further along the sequence these upper bounds will decrease (why?), that is $U_{n+1} \leq U_n$ for all n. We know $U_n \geq a_n \geq L$ so that the sequence $(U_1, U_2, \ldots)$ is bounded below. So the Weierstrass Criterion tells us that the limit

$$a = \lim_{n\to\infty} U_n = \lim_{n\to\infty} \sup\{a_n, a_{n+1}, a_{n+2}, \ldots\}$$

must exist. This is step (i) the promised formula or the likely limit. The final step (ii) is to check we are right, that is to show that $a_n \to a$.

Fix $\epsilon > 0$. We use the Cauchy property to find N so that

$$|a_n - a_m| \leq \epsilon \quad \text{for all } n, m \geq N. \qquad (\star)$$

Since $U_n \to a$ there exists $N_1 \geq N$ so that $|U_{N_1} - a| < \epsilon$. The supremum $U_{N_1} = \sup\{a_{N_1}, a_{N_1+1}, a_{N_1+2}, \ldots\}$ must be 'nearly attained' (see the lemma after the definition of supremum) so there exists $N_2 \geq N_1$ so that

$$U_{N_1} - \epsilon < a_{N_2} < U_{N_1}$$

and thus $|a_{N_2} - U_{N_1}| < \epsilon$. The triangle inequality gives

$$|a_{N_2} - a| \leq |a_{N_2} - U_{N_1}| + |U_{N_1} - a| < 2\epsilon.$$

Combining this with the Cauchy property $(\star)$ above we find

$$|a_n - a| \leq |a_n - a_{N_2}| + |a_{N_2} - a| < \epsilon + 2\epsilon = 3\epsilon$$

for $n \geq N_2$, finishing the proof that $a_n \to a$ as $n \to \infty$. □

Non-examinable Remark. The argument above needs a series of steps done in exactly the right order. If you check each step you may have an increased admiration for Cauchy.

One reason I put the proof in the lecture notes is that it gives you a first view at two notions related to limits: *limsup* and *liminf* defined as

$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} \sup\{a_n, a_{n+1}, a_{n+2}, \ldots\},$$
$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty} \inf\{a_n, a_{n+1}, a_{n+2}, \ldots\}.$$

We used $\limsup_{n\to\infty} a_n$ as our guess for what the limit must be in the above proof. These notions used to be in every first year calculus module, and you may well meet them if you take further analysis or probability modules. They allow one to describe sequences that do not converge, for example that oscillate. For example the non-convergent sequence $a_n = (-1)^n \frac{n+1}{n}$ has (please check) that

$$\limsup_{n\to\infty} a_n = +1, \qquad \liminf_{n\to\infty} a_n = -1.$$

They are used in proofs when arguing that limits exist, since (as can be checked), for a bounded sequence $(a_n : n \geq)$, both $\limsup_{n\to\infty} a_n$ and $\liminf_{n\to\infty} a_n$ always exist and the limit $\lim_{n\to\infty} a_n$ exists precisely when $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$. An expert's summary of the proof above is 'The Cauchy property ensures that $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$ and so the limit must exist'.

Why will I give a second proof of Cauchy's Criterion? The second proof uses the ideas of subsequences, which turn out to be rather useful and you will meet in later maths and stats modules (for example MA260, ST318, ST342).

Definition. *A **subsequence** of $(a_n : n \geq 1)$ is a sequence of the form $(a_{n_1}, a_{n_2}, a_{n_3}, \ldots)$ where $1 \leq n_1 < n_2 < n_3 < \ldots$ is a strictly increasing sequence of integers.*

In the above definition, the k-th term of the sequence $(a_{n_k} : k \geq 1)$ is a_{n_k}.

Example. *Let $(a_n : n \geq 1)$ be a sequence. Examples of subsequences are the following:*

- $(a_2, a_4, a_6, a_8, \ldots)$ *which you could write as* $(a_{2n} : n \geq 1)$.
- $(a_1, a_3, a_5, a_7, \ldots)$ *which you could write as* $(a_{2n-1} : n \geq 1)$.
- $(a_{101}, a_{102}, a_{103}, \ldots)$ *which you could write as* $(a_{100+n} : n \geq 1)$.

The key step in the second proof of Cauchy's Criterion is a result about subsequences that is so often used in later modules that it merits itself being called a Theorem.

Theorem. ***The Bolzano Weierstrass Theorem***

Every bounded sequence $(a_n : n \geq 1)$ has a convergent subsequence.

Example. *The sequence* $(1, -1, 1, -1, \ldots)$ *defined by* $a_n = (-1)^{n+1}$ *has the convergent subsequence* $(a_2, a_4, a_6, a_8, \ldots) = (-1, -1, -1, \ldots)$.

A more interesting example is $a_n = \sin(n)$. This is bounded above by 1 and below by -1 so the Bolzano Weierstrass Theorem tells us it must have a convergence subsequence. But it is not so easy to write down a specific such subsequence.

We will prove The Bolzano Weierstrass Theorem as the final proof of this chapter. We first show that it leads to a short second proof Cauchy's Theorem.

Proof. **Second proof Cauchy's Theorem** We give a second proof for the harder implication. That is we suppose that $(a_n : n \geq 1)$ has the Cauchy property and we show it converges. We also steal the first step, where we checked the Cauchy property implies that $(a_n : n \geq 1)$ must be bounded.

Now we may use the Bolzano Weierstrass theorem to conclude that $(a_n : n \geq 1)$ has a convergent subsequence $(a_{n_1}, a_{n_2}, a_{n_3}, \ldots)$, whose limit we denote by a. We will now show that the whole sequence converges to a.

Choose $\epsilon > 0$. Using the Cauchy property, there exists $N \geq 1$ such that $|a_n - a_m| < \epsilon$ whenever $m, n \geq N$. Since $a_{n_k} \to a$ as $k \to \infty$ there exists K so that $|a_{n_k} - a| < \epsilon$ whenever $k \geq K$. We may choose K so large that $n_K \geq N$. Then for $n \geq N$ the triangle inequality gives

$$|a_n - a| \leq |a_n - a_{n_K}| + |a_{n_K} - a| < \epsilon + \epsilon = 2\epsilon$$

and therefore $a_n \to a$. $\square$

The final debt we need to pay is ...

Proof. **Proof of the Bolzano Weierstrass Theorem**. I copy here the beautiful proof from last year's lecture notes. We need a_k and b_k as notation in this proof so we start by supposing that $(x_n : n \geq 1)$ is a bounded sequence. We shall assume that $0 \leq x_n \in 1$ for all n thereby simplifying the notation. You can make the small changes in the proof needed when you only know $L \leq x_n \leq U$ for all n.

Let us now define new sequences $(a_n : n \geq 1)$ and $(b_n : n \geq 1)$ as follows. Let $a_1 = 0, b_1 = 1$. We observe that at least one of the following statements is true.

(i) $x_n \in [0, \frac{1}{2}]$ for infinitely many $n \geq 1$.

(ii) $x_n \in [\frac{1}{2}, 1]$ for infinitely many $n \geq 1$.

If (i) is true, we set $a_2 = 0, b_2 = \frac{1}{2}$. Otherwise, we set $a_2 = \frac{1}{2}, b_2 = 1$.

We now continue this algorithm with $[a_1, b_1]$ replaced by $[a_2, b_2]$, etc. More precisely, suppose that for $k \geq 1$ we are given $0 \leq a_k \leq b_k \leq 1$ such that $x_n \in [a_k, b_k]$ for infinitely many $n \geq 1$. Then at least one of the following statements is true.

(i) $x_n \in [a_k, \frac{a_k+b_k}{2}]$ for infinitely many $n \geq 1$.

(ii) $x_n \in [\frac{a_k+b_k}{2}, b_k]$ for infinitely many $n \geq 1$.

If (i) is true, we set $a_{k+1} = a_k$, $b_{k+1} = \frac{a_k+b_k}{2}$. Otherwise, we set $a_{k+1} = \frac{a_k+b_k}{2}$, $b_{k+1} = b_k$. In either case,

$$x_n \in [a_{k+1}, b_{k+1}] \text{ for infinitely many } n \geq 1$$

and

$$b_{k+1} - a_{k+1} = \tfrac{b_k - a_k}{2} = \frac{1}{2^{k-1}}.$$

To deduce the last equality we used the fact that

$$\tfrac{a_k+b_k}{2} - a_k = b_k - \tfrac{a_k+b_k}{2} = \tfrac{b_k-a_k}{2}\,.$$

Furthermore,

$$a_k \leq a_{k+1}, \quad b_{k+1} \leq b_k\,.$$

Summarising, we have constructed sequences a_k, b_k taking values in $[0,1]$ which satisfy the following properties.

(i) $a_1 = 0$, $b_1 = 1$.

(ii) $b_k - a_k = \frac{1}{2^{k-1}}$ for all k.

(iii) a_k is increasing.

(iv) b_k is decreasing.

(v) For all k we have that $x_n \in [a_k, b_k]$ for infinitely many $n \geq 1$.

Note that $(a_k : k \geq 1)$ is a bounded increasing sequence. Therefore by Weierstrass's Criterion it has a limit $a \in \mathbb{R}$. Similarly, since $(b_k : k \geq 1)$ is a bounded decreasing sequence, it has a limit $b \in \mathbb{R}$. Letting $n \to \infty$ in property (ii), $a = b$.

By induction on k, we can construct a strictly increasing sequence of positive integers n_k such that

$$x_{n_k} \in [a_k, b_k]\,.$$

Namely, given such $n_1 < n_2 < \ldots < n_{k-1}$, by property (v) there exists $n_k > n_{k-1}$ such that $x_{n_k} \in [a_k, b_k]$ (this interval contains infinitely many terms of the sequence by construction). In particular, $(x_{n_1}, x_{n_2}, x_{n_3}, \ldots)$ is a subsequence of $(x_1, x_2, x_3, \ldots)$.

By the Sandwich Lemma, $x_{n_k} \to a$ as $k \to \infty$. We have found a convergent subsequence. $\square$

Chapter 3. Infinite series

Convergence and divergence of infinite series

Let's start with an old chestnut:

$$\begin{aligned}
1 &= 1+0+0+0+\dots \\
&= 1+(-1+1)+(-1+1)+(-1+1)+\dots \\
&= (1-1)+(1-1)+(1-1)+(1-1)+\dots \\
&= 0+0+0+0+\dots \\
&= 0.
\end{aligned}$$

Collapse of all mathematics? It can only be the fact that we have infinite sums that is causing a problem somewhere. Our next aim is to make sense of infinite sums

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

Our knowledge of limits allows us to make everything precise.

Definition. *Define the associated sequence of **partial sums***

$$S_n = \sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n\,. \tag{2}$$

*We say that the series $\sum_{n=1}^{\infty} a_n$ **converges** if there exists $A \in \mathbb{R}$ such that $S_n \to A$ as $n \to \infty$. In this case we say $\sum_{n=1}^{\infty} a_n$ **converges to** A and we write*

$$\sum_{n=1}^{\infty} a_n = A$$

*and we call A the **sum** of the series.*

Definition. *We say that a series is **divergent** if it is not convergent. If $S_n \to \infty$, we say that the series $\sum_{n=1}^{\infty} a_n$ **diverges to** ∞ and we write*

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

Remark. *Since we have cast our definition of convergence of a series* $\sum a_n$ *in terms of limits we can immediately apply the results we have built up for limits. For example the sum rule for limits implies that:*

$$\text{if } \sum_{n=1}^{\infty} a_n = A \text{ and } \sum_{n=1}^{\infty} b_n = B \text{ then } \sum_{n=1}^{\infty}(a_n + b_n) = A + B.$$

Of course! To check this look at the partial sums $S_n = \sum_{k=1}^{n} a_k$ *and* $T_n = \sum_{k=1}^{n} b_k$. *Both series converge so that* $S_n \to A$ *and* $T_n \to B$ *as* $n \to \infty$. *The sum rule tells us* $S_n + T_n \to A + B$. *But* $S_n + T_n$ *equals the partial sum* $\sum_{k=1}^{n}(a_k + b_k)$.
In the same way for any constant $C \in R$ *we have* $\sum_{n=1}^{\infty} C\, a_n = C \sum_{n=1}^{\infty} a_n$.

Example. *Geometric series*

We need to check that our familiar example, infinite geometric series, fits this definition. Fix $x \in \mathbb{R}$ *and we consider the series*

$$\sum_{n=1}^{\infty} x^{n-1} = 1 + x + x^2 + \dots$$

Lemma. *The geometric series is convergent if* $|x| < 1$, *in which case its sum is* $\frac{1}{1-x}$, *and it is divergent if* $|x| \geq 1$.

Proof. Recall the standard argument, as follows. If $x = 1$ then $S_n = n \to \infty$ so the series is divergent. Otherwise, write

$$S_n = 1 + x + x^2 + \cdots + x^{n-1}$$

and so

$$xS_n = x + x^2 + x^3 + \cdots + x^n.$$

Subtracting these equations many terms cancel leaving $(1-x)S_n = 1 - x^n$. Since $x \neq 1$,

$$S_n = \frac{1 - x^n}{1 - x}\,.$$

If $|x| < 1$, then our basic limit $x^n \to 0$ implies $S_n \to \frac{1}{1-x}$. If $|x| > 1$, then $x^{2n} \to \infty$ so $(1-x)S_{2n} = 1 - x^{2n} \to -\infty$. Therefore the series diverges. Finally, if $x = -1$, then $S_{2n} = 0$ and $S_{2n+1} = 1$ for all n and so the series diverges. $\square$

Example. *Telescoping series* *Consider the series*

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \frac{1}{1.2} + \frac{1}{2.3} + \frac{1}{3.4} + \ldots = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \ldots$$

There is a trick! Using the idea of partial fractions we have $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$. *This shows*

$$\begin{aligned} S_N &= \sum_{n=1}^{N} \frac{1}{n(n+1)} \\ &= \frac{1}{1.2} + \frac{1}{2.3} + \frac{1}{3.4} + \ldots + \frac{1}{N(N+1)} \\ &= \frac{1}{1} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{N} - \frac{1}{N+1} \\ &= 1 - \frac{1}{N+1}. \end{aligned}$$

A lucky break - the terms almost all cancel and long sum collapses like a collapsing telescope into just two terms. Now we see that $S_N = 1 - \frac{1}{N+1} \to 1$ *as* $N \to \infty$ *so that*

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

Example. *We consider the* ***harmonic series*** *which is*

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

Lemma. *The harmonic series is divergent.*

Proof. Idea: we can regroup the terms of the harmonic series in blocks as follows:

$$1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq 1/2} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{\geq 1/2} + \underbrace{\frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \frac{1}{12} + \frac{1}{13} + \frac{1}{14} + \frac{1}{15} + \frac{1}{16}}_{\geq 1/2} + \cdots$$

We now make this more precise. We use this idea to estimate the partial sums:

$$S_2 = 1 + \frac{1}{2}, \quad S_4 > 1 + \frac{1}{2} + \frac{1}{2}, \quad S_8 > 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}, \quad S_{16} > 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}.$$

This leads to the guess

$$S_{2^n} \geq 1 + \frac{n}{2} \quad \text{for all } n \geq 1.$$

We prove the guess by induction. The base case $n = 1$ holds since $S_{2^1} = 1 + \frac{1}{2}$. For the step,

$$S_{2^{n+1}} = S_{2^n} + \underbrace{\frac{1}{2^n + 1} + \frac{1}{2^n + 2} + \cdots + \frac{1}{2^{n+1}}}_{2^n \text{ terms, all of which are at least } \frac{1}{2^{n+1}}}.$$

Therefore,

$$S_{2^{n+1}} \geq S_{2^n} + \frac{2^n}{2^{n+1}} = S_{2^n} + \frac{1}{2}\,.$$

Using this and the induction hypothesis,

$$S_{2^{n+1}} \geq 1 + \frac{n}{2} + \frac{1}{2} = 1 + \frac{n+1}{2}\,.$$

This verifies our guess above. Hence $S_{2^n} \geq 1 + \frac{n}{2} \to \infty$. But, since the terms in the series are positive, the partial sums are increasing and it follows that S_n diverges. $\square$

Example. *Find a value of $N \geq 1$ such that*

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N} \geq 10\,.$$

Proof. We showed that $S_{2^n} \geq 1 + n/2$. Solving $1 + n/2 \geq 10$ leads us to the choice $n = 18$. Then $S_{2^{18}} \geq 1 + \frac{18}{2} = 10$. This guarantees that you could take $N = 2^{18} = 262144$. $\square$

The harmonic series diverges to infinity but very slowly. Indeed $\sum_{n=1}^{10^{43}} \frac{1}{n} < 100$.

Lemma. *Suppose that the series $\sum_{n=1}^{\infty} a_n$ converges. Then $a_n \to 0$ as $n \to \infty$.*

Proof. If the series $\sum_{n=1}^{\infty} a_n$ converges, then the sequence of partial sums $S_n = \sum_{k=1}^{n} a_k$ converges to s the sum of the series. But

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

We know the shifted sequence $S_{n-1} \to s$ as $n \to \infty$ also. So the sum rule tells us $a_n \to s - s = 0$. $\square$

This simple lemma above is commonly used in its contrapositive ($P \Rightarrow Q$ is equivalent to Not $Q \Rightarrow$ Not P.) Thus we get a **Necessary Condition for Convergence**:

if a_n does not converge to 0 then the series $\sum_{n=1}^{\infty}$ must diverge.

The examples above suggest that the terms a_n in a series should be converging to zero sufficiently fast for the infinite series $\sum_{n=1}^{\infty} a_n$ to converge. The telescoping series shows that $a_n = \frac{1}{n(n+1)}$ is fast enough, but the harmonic series shows that $a_n = \frac{1}{n}$ is too slow.

Some convergence tests

Sad Fact. For most convergent series there will not be a simple formula for the sum. (Similar sad facts hold for most integrals - though you have a few more tools available). So we reduce our aims: we will become good at deciding whether a series converges or not. Also when the exact sum it becomes more important to estimate the the speed that the partial sums S_n converge, so that we know how to approximate the exact sum efficiently.

When all the terms a_n of the series $\sum_{n=1}^{\infty} a_n$ are non-negative, we have a simple criterion for convergence, as follows.

Theorem. ***Boundedness criterion.*** *Suppose that $a_n \geq 0$ for all $n \geq 1$. Then the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the associated sequence of partial sums $S_n = \sum_{k=1}^{n} a_k$ is bounded.*

Proof. Suppose first that the sequence $(S_n : n \geq 1)$ is bounded. Since all the terms a_n of the series are non-negative, it follows that S_n is increasing. Therefore, by Weierstrass's Criterion the sequence $(S_n : n \geq 1)$ must converge - this is the definition that the series $\sum_{n=1}^{\infty} a_n$ converges.
Conversely, suppose that the series $\sum_{n=1}^{\infty} a_n$ converges. Then the sequence of partial sums $(S_n : n \geq 1)$ is convergent, and therefore bounded. $\square$

This implies a hugely useful criterion for convergence in the case when the series have non-negative terms: compare the series with one you already know converges. This won't help us find the sum, but it will help us show the series is convergent.

Theorem. ***Comparison Test for Series.*** *Suppose that $0 \leq a_n \leq b_n$ for all $n \geq 1$.*

(i) If $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges and $\sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} b_n$.

(ii) If $\sum_{n=1}^{\infty} a_n$ diverges to infinity, then $\sum_{n=1}^{\infty} b_n$ diverges to infinity.

Proof. (i) Suppose $B = \sum_{n=1}^{\infty} b_n$. Then

$$\sum_{k=1}^{n} a_k \leq \sum_{k=1}^{n} b_k \leq B.$$

Hence $\sum_{n=1}^{\infty} a_n$ is convergent by the Boundedness Criterion above. Moreover

$$\sum_{n=1}^{\infty} a_n = \lim_{n\to\infty} \sum_{k=1}^{n} a_k \leq B.$$

For part (ii) note that since $a_n \geq 0$ the partial sums S_n are increasing and the only way $\sum_{n=1}^{\infty} a_n$ to diverge is if it diverges to infinity. Then part (ii) is just the contrapositive of part (i) (check!). $\square$

Example.

$$\textit{(i)} \sum_{n=1}^{\infty} \frac{3^n + 7^n}{2^n + 10^n} \ \textit{converges.} \quad \textit{(ii)} \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \ \textit{diverges.} \quad \textit{(iii)} \sum_{n=1}^{\infty} \frac{1}{n^2} \ \textit{converges.}$$

Proof. (i) We note that, for all $n \geq 1$ we have

$$0 \leq \frac{3^n + 7^n}{2^n + 10^n} \leq \frac{3^n + 7^n}{10^n} = \left(\frac{3}{10}\right)^n + \left(\frac{7}{10}\right)^n.$$

Since $\frac{3}{10} \in (0,1)$ and $\frac{7}{10} \in (0,1)$, it follows that $\sum_{n=1}^{\infty} (\frac{3}{10})^n$ and $\sum_{n=1}^{\infty} (\frac{7}{10})^n$ converge. By the sum rule $\sum_{n=1}^{\infty} [(\frac{3}{10})^n + (\frac{7}{10})^n]$ converges. By the Comparison Test we deduce that $\sum_{n=1}^{\infty} \frac{3^n+7^n}{2^n+10^n}$ converges.

(ii) We know that $\frac{1}{\sqrt{n}} \geq \frac{1}{n}$ for all $n \geq 1$ and that the series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges (the harmonic series example). Hence, by the Comparison Test the series $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ diverges.

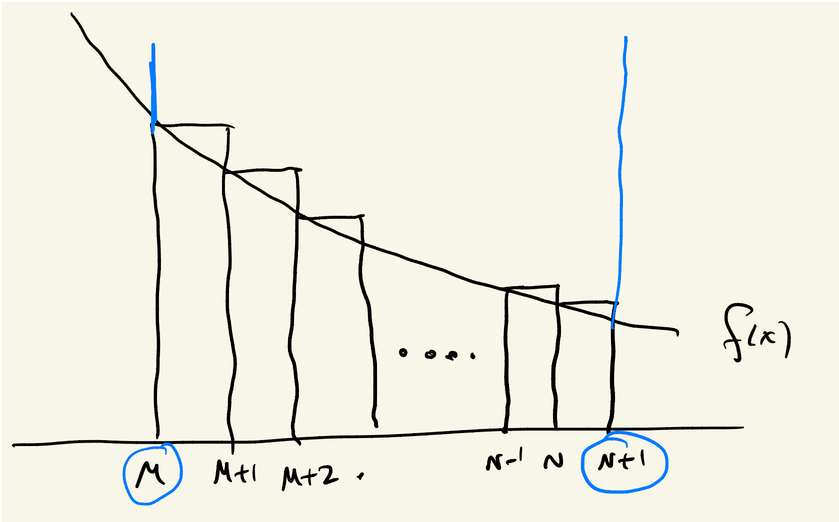
(iii) I claim $\frac{1}{n^2} \leq \frac{2}{n(n+1)}$ - please check by cross multiplying. But we saw that $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ converges (it was a telescoping series). So $\sum_{n=1}^{\infty} \frac{2}{n(n+1)}$ also converges and by the Comparison Test we now know that $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges. $\square$

Rather delicate comparisons can be made using our skills with integration. We used these ideas in week one when discussing Stirling's formula.

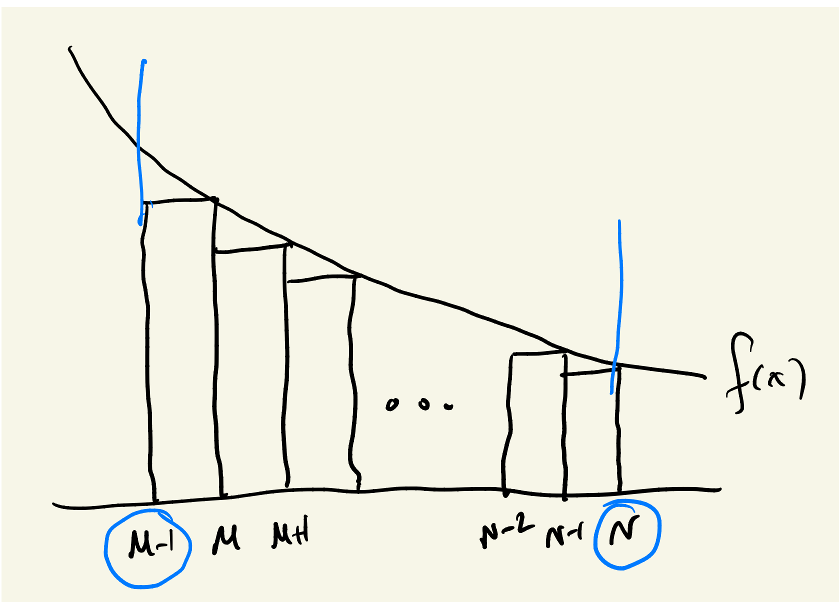
Lemma. *Integral Bounds*
*Suppose that the function $f : (0,\infty) \to \mathbb{R}$ is **decreasing** (i.e. $f(x) \leq f(y)$ for all $x \geq y$). Then for all $1 \leq m < n$,*

$$\int_m^{n+1} f(x)\,\mathrm{d}x \leq \sum_{k=m}^{n} f(k) \leq \int_{m-1}^{n} f(x)\,\mathrm{d}x\,. \tag{3}$$

Proof. Two pictures say it all:



The area of all the blocks is $\sum_{k=m}^{n} f(k)$. This is greater than the area under the curve between m and $n+1$.



The area of all the blocks is again $\sum_{k=m}^{n} f(k)$. This is less than the area under the curve between $m-1$ and n. $\square$

Example. *Let $p > 0$. The series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if $p > 1$ and diverges if $0 < p \leq 1$.*

Proof. We consider the function $f : (0, \infty) \to \mathbb{R}$ given by $f(x) = \frac{1}{x^p}$. Note that f is non-negative and decreasing.

For $p \in (0,1)$ we use the lemma to see that

$$S_N = \sum_{n=1}^{N} \frac{1}{n^p} \geq \int_1^{N+1} \frac{1}{x^p} dx = \left. \frac{x^{1-p}}{1-p} \right|_1^{N+1} = \frac{(N+1)^{1-p} - 1}{1-p}.$$

Since $1-p > 0$ the term $(N+1)^{1-p}$ diverges to infinity as $N \to \infty$. Thus the partial sums S_N also diverge to ∞ (by Comparison Lemma for sequences) showing that $\sum_{n=1}^{\infty} \frac{1}{n^p} = \infty$.

For $p > 1$ we use the lemma to see that

$$\sum_{n=2}^{N} \frac{1}{n^p} \leq \int_1^{N} \frac{1}{x^p} dx = \left. \frac{x^{1-p}}{1-p} \right|_1^{N} = \frac{N^{1-p} - 1}{1-p} = \frac{1 - N^{1-p}}{p-1} \leq \frac{1}{p-1}.$$

(Why did we start at $n = 2$ here?) So the partial sums $S_N = \sum_{n=1}^{N} \frac{1}{n^p}$ are bounded above by $1 + \frac{1}{p-1}$, and this implies that $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges.

We already know the answer for $p = 1$, by the beautiful blocking argument earlier, but it would follow using the integral bounds as well - please check. □

The easiest comparison is to compare to a geometric series. The Ratio Test below is commonly used recipe that makes this comparison quick to do.

Theorem. *Ratio test.*
Suppose that $(a_n : n \geq 1)$ is a sequence with $a_n > 0$ for all $n \geq 1$ and so that $\frac{a_{n+1}}{a_n} \to \ell$. Then $\sum_{n=1}^{\infty} a_n$ converges if $0 \leq \ell < 1$ and diverges if $\ell > 1$.

Proof. I'll write the proof somewhat backwards to make it clear that we are just comparing with a geometric series.

If we knew

for some $\gamma < 1$ the terms $a_n \leq C\gamma^n$ for all n

we could directly apply the Comparison Test for Series and conclude that $\sum_{n=1}^{\infty} a_n$ converges by comparison with the geometric series $C \sum_{n=1}^{\infty} \gamma^n = C\gamma/(1-\gamma)$.

If we knew

for some $\gamma < 1$ the terms $a_n \leq C\gamma^n$ for all $n \geq N$ $\quad (\star)$

we could directly apply the Comparison Test for Series and conclude that $\sum_{n=N}^{\infty} a_n$ converges by comparison with the geometric series $C \sum_{n=N}^{\infty} \gamma^n = C\gamma^N/(1-\gamma)$. But adding on the first $N-1$ terms $\sum_{n=1}^{N-1} a_n$ cannot change the convergence and we can still conclude that $\sum_{n=1}^{\infty} a_n$ converges.

If we knew

$$\text{for some } \gamma < 1 \text{ the ratios } \tfrac{a_{n+1}}{a_n} \leq \gamma \text{ for all } n \geq N \quad (\star\star)$$

then we get $a_{N+1} \leq a_N\gamma$, $a_{N+2} \leq a_N\gamma^2$, ..., and in general $a_{N+n} \leq a_N\gamma^n = a_N\gamma^{-N}\gamma^{n+N}$ for all $n \geq 1$. But this implies the condition $(\star)$ above by taking $C = a_N\gamma^{-N}$. So we again conclude that $\sum_{n=1}^{\infty} a_n$ converges.

Finally, if we knew

$$\text{the ratios } \tfrac{a_{n+1}}{a_n} \to l < 1 \text{ as } n \to \infty$$

then we can pick $\epsilon = (1-l)/2$ and conclude that there exists N so that

$$\text{for all } n \geq N \text{ we have } \tfrac{a_{n+1}}{a_n} \leq l + \epsilon = \tfrac{1+l}{2}$$

which is the condition $(\star\star)$ with $\gamma = \frac{1+l}{2} < 1$. We conclude that that $\sum_{n=1}^{\infty} a_n$ converges.

Next suppose that $\ell > 1$. Note that then $1 < \frac{1+\ell}{2} < \ell$. We want to use the limit $\frac{a_{n+1}}{a_n} \to \ell$ to see there exists N so that for $n \geq N$

$$\frac{a_{n+1}}{a_n} > \frac{1+\ell}{2}.$$

What ϵ have I picked? It must be $\epsilon = \ell - \frac{1+\ell}{2} = \frac{\ell-1}{2} > 0$. Now we find

$$a_{N+1} > \frac{1+\ell}{2}a_N, \quad a_{N+2} > \left(\frac{1+\ell}{2}\right)^2 a_N, \quad a_{N+3} > \left(\frac{1+\ell}{2}\right)^3 a_N, \ldots.$$

In general $a_{N+n} > \left(\frac{1+\ell}{2}\right)^n a_N$, so the terms in the series diverge to infinity. We saw that there is then no hope that the series $\sum a_n$ converges. $\square$

Example. *The series $\sum_{n=1}^{\infty} \frac{1}{n!}$ converges.*

Proof. Let $a_n = \frac{1}{n!}$. Then $\frac{a_{n+1}}{a_n} = \frac{1}{n+1} \to 0$. Hence, the series $\sum_{n=1}^{\infty} \frac{1}{n!}$ converges by the ratio test. $\square$

Example. *The series $\sum_{n=1}^{\infty} \frac{n^2}{2^n}$ converges.*

Proof. Let $a_n = \frac{n^2}{2^n}$. We compute

$$\frac{a_{n+1}}{a_n} = \frac{\frac{(n+1)^2}{2^{n+1}}}{\frac{n^2}{2^n}} = \frac{1}{2} \cdot \frac{(n+1)^2}{n^2} \to \frac{1}{2}.$$

By the ratio test, it follows that the series $\sum_{n=1}^{\infty} \frac{n^2}{2^n}$ converges. $\square$

Example. *What if $\frac{a_{n+1}}{a_n} \to 1$ as $n \to \infty$? We get no conclusion - it is possible that $\sum_{n=1}^{\infty} a_n$ converges and possible that $\sum_{n=1}^{\infty} a_n$ diverges. For example if $a_n = \frac{1}{n}$, the series diverges, while if $a_n = \frac{1}{n^2}$, the series converges. In both cases $\frac{a_{n+1}}{a_n} \to 1$.*

Absolute convergence

Series with both positive and negative terms are harder - there will be some cancelation that may help the series to converge. If we replace the terms a_n by their absolute values $|a_n|$ this will eliminate all cancelation, and should make it more likely that the series diverges. This is indeed true, as the next result shows.

Theorem. ***Absolute Convergence Theorem***

If the series $\sum_{n=1}^{\infty} |a_n|$ converges, then the series $\sum_{n=1}^{\infty} a_n$ converges.

Remark. *To prove this, we need to show that the partial sums $S_n = \sum_{k=1}^{n} a_k$ converge. But these will not be increasing so we cannot use the Weierstrass Criterion. You will not be surprised that we will use the Cauchy Criterion.*

Proof. Define the sequences of partial sums

$$S_n = \sum_{k=1}^{n} a_k, \qquad T_n = \sum_{k=1}^{n} |a_k| \, .$$

By assumption, $\sum_{n=1}^{\infty} |a_n|$ is convergent, so the sequence T_n is convergent. In particular, by the Cauchy Criterion, $(T_n : n \geq 1)$ is a Cauchy sequence. Therefore, given $\epsilon > 0$, there exists $N \geq 1$ such that $|T_n - T_m| < \epsilon$ whenever $m, n \geq N$.

We may assume, without loss of generality, that $m > n$. By the triangle inequality, for $m > n \geq N$,

$$|S_m - S_n| = \left| \sum_{k=n+1}^{m} a_k \right| \leq \sum_{k=n+1}^{m} |a_k| = T_m - T_n = |T_m - T_n| < \varepsilon.$$

But this shows that $(S_n : n \geq 1)$ is Cauchy and therefore convergent by the Cauchy Criterion again. Hence, $\sum_{n=1}^{\infty} a_n$ converges. □

This gives us a way of establishing convergence for series with both positive and negative terms: apply the ideas from the previous section to show $\sum_{n=1}^{\infty} |a_n|$ converges, and the Absolute Convergence Theorem implies that $\sum_{n=1}^{\infty} a_n$ converges.

Example. *The series $\sum_{n=1}^{\infty} \frac{\cos(n)}{n^2}$ converges.*

Proof. We note that for all $n \geq 1$ we have

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

Since we know the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges, we obtain by the Comparisons Test for Series that $\sum_{n=1}^{\infty} \left| \frac{\cos(n)}{n^2} \right|$ converges. Then the Absolute Convergence Theorem guarantees that $\sum_{n=1}^{\infty} \frac{\cos(n)}{n^2}$ converges. $\square$

Corollary. ***Improved Ratio Test*** *Suppose that* $(a_n : n \geq 1)$ *is a sequence such that* $a_n \neq 0$ *for all* $n \geq 1$ *and such that* $\frac{|a_{n+1}|}{|a_n|} \to \ell$*. Then* $\sum_{n=1}^{\infty} a_n$ *converges if* $0 \leq \ell < 1$ *and diverges if* $\ell > 1$*.*

Proof. Suppose first that $0 \leq \ell < 1$. Then by the Ratio Test (for positive series) it follows that $\sum_{n=1}^{\infty} |a_n|$ converges. By the Absolute Convergence Theorem we deduce that the series $\sum_{n=1}^{\infty} a_n$ converges.
Suppose now that $\ell > 1$. In this case, in the proof of Ratio Test for the positive series $\sum_{n=1}^{\infty} |a_n|$ we showed that $|a_n| \to \infty$ as $n \to \infty$. So $a_n \not\to 0$ and the series $\sum_{n=1}^{\infty} a_n$ cannot converge. $\square$

You will make use of the Ratio Test frequently in term 2 when you study power series.

Example. *Determine for which values of* $x \in \mathbb{R}$ *the series* $\sum_{n=1}^{\infty} \frac{x^n}{n}$ *is convergent.*

Proof. We observe that the series converges for $x = 0$. In what follows, we consider the case when $x \neq 0$. For $n \geq 1$, let $a_n = \frac{x^n}{n}$. In particular $a_n \neq 0$ for all $n \geq 1$.
We compute the ratio

$$\frac{|a_{n+1}|}{|a_n|} = \frac{\frac{|x|^{n+1}}{n+1}}{\frac{|x|^n}{n}} = \frac{n}{n+1} \cdot |x| \to |x| \, .$$

By the Ratio Test the series $\sum_{n=1}^{\infty} \frac{x^n}{n}$ is convergent when $|x| < 1$ and is divergent when $|x| > 1$. We need to see what happens when $|x| = 1$, i.e. when $x = 1$ and $x = -1$. For $x = 1$, we obtain the series $\sum_{n=1}^{\infty} \frac{1}{n}$ which is divergent. When $x = -1$, we obtain the series $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ and we are stuck - none of our results so far apply. (In the next few pages we will see that it is indeed convergent.) $\square$

Example. *Let's look at the example from the introductory chapter*

$$J_0(x) := \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{2^{2n}(n!)^2}.$$

For each $x \in \mathbf{R}$ *we are trying to define the function* J_0 *at the point* x *by an infinite series. We can fix* x *and use the Ratio Test to see if the series converges: we see*

$$\frac{|a_{n+1}|}{|a_n|} = \frac{\frac{|x|^{2(n+1)}}{2^{2(n+1)}((n+1)!)^2}}{\frac{|x|^{2n}}{2^{2n}(n!)^2}} = \frac{|x|^2}{2^2(n+1)^2}.$$

Note that taking the absolute values has removed the factors $(-1)^n$ and turned x^{2n} into $|x|^{2n}$. Now we see that for any value of x we will get $\frac{|a_{n+1}|}{|a_n|} \to 0$ and the Ratio Test tells us the series converges. Thus $J_0(x)$ is properly defined for any value of x.

Are there series where $\sum_{n=1}^{\infty} a_n$ converges but $\sum_{n=1}^{\infty} |a_n|$ diverges? Yes - and these are trickier. But there is one small class that we can treat quite easily.

Theorem. ***Alternating Series Theorem*** *Suppose that a sequence $(a_n : n \geq 1)$ satisfies both*

$$(i)\ a_n \to 0 \text{ as } n \to \infty, \quad \text{and} \quad (ii)\ (a_n : n \geq 1) \text{ is decreasing.}$$

Then the series

$$\sum_{n=1}^{\infty} (-1)^{n+1} a_n = a_1 - a_2 + a_3 - a_4 + \dots \tag{4}$$

converges to a limit A.

Moreover, we have the error bound describing the speed of convergence

$$|S_n - A| \leq a_{n+1} \quad \text{for all } n.$$

Remark. *We call a series of the form described here an* ***alternating series*** *since the terms $(-1)^{n+1} a_n$ alternate in sign. Note we are asking that the size of the terms decreases. The simplest example is*

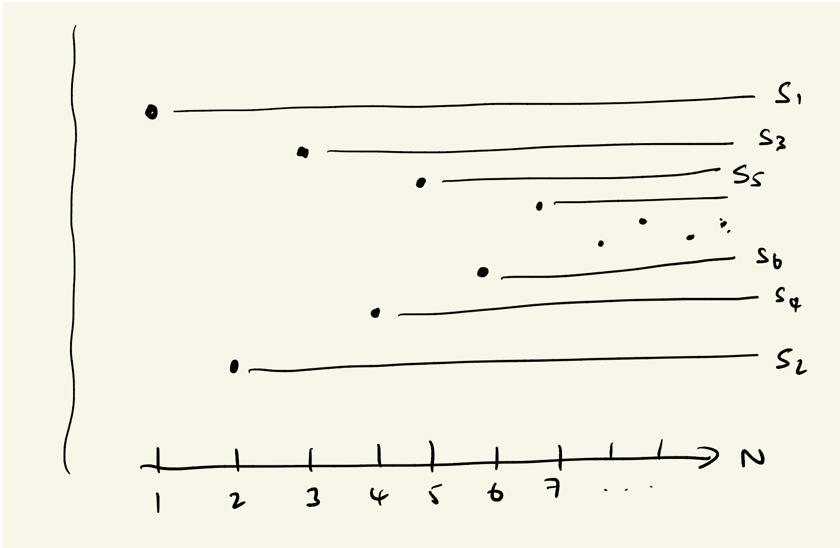
$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

The Theorem guarantees this converges, and we know that without the minus signs it would diverge. (We will show later that the sum is exactly calculable as $\log(2)$*.)*

Proof. The picture shows the key idea: that $S_2 \leq S_4 \leq S_6 \leq \dots \quad \dots \leq S_5 \leq S_3 \leq S_1$. We can check this by induction: the induction hypothesis is

$$S_2 \leq S_4 \leq \ldots \leq S_{2n} \leq S_{2n-1} \leq \ldots \leq S_3 \leq S_1.$$

This holds for $n = 1$ since $S_2 = S_1 - a_2 \leq S_1$. To go from n to $n+1$ we need to check three inequalities: $S_{2n} \leq S_{2n+2} \leq S_{2n+1} \leq S_{2n-1}$. Each of them is easy: $S_{2n+2} - S_{2n} = a_{2n+1} - a_{2n+2} \geq 0$; $S_{2n+2} - S_{2n+1} = -a_{2n+2} \leq 0$; $S_{2n+1} - S_{2n-1} = -a_{2n} + a_{2n+1} \leq 0$.



Thus the subsequence $(S_2, S_4, S_6, \ldots)$ is increasing and bounded above by S_1 and so must converge (by Weierstrass Criterion) to a limit L. Similarly the subsequence $(S_1, S_3, S_5, \ldots)$ is increasing and bounded below by S_2 and must converge to a limit U. But

$$0 \leq S_{2n+1} - S_{2n} = a_{2n+1} \to 0$$

so by the Sandwich Theorem $U - L = 0$, that is U and L are equal, and we call them A. Now it won't surprise you that the whole sequence $S_n \to A$ as $n \to \infty$ and so the series $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ converges.

The error bound stated in the Theorem is hidden in our calculations:

$$0 \leq S_{2n+1} - A \leq S_{2n+1} - S_{2n} = a_{2n+1} \quad \text{and} \quad 0 \leq A - S_{2n} \leq S_{2n-1} - S_{2n} = a_{2n}.$$

$\square$

Example. *In the introduction we mentioned the series*

$$\frac{\pi^2}{12} = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \ldots = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n^2}.$$

Can we say anything about this series? The Alternating Series Theorem applies and we know it must converge. We cannot (yet) prove that the sum is $\frac{\pi^2}{8}$ - for that you might take some modules with complex variable integration or with Fourier series in year two. But the error bound applies and if we sum the first 99 terms to get $S_{99} = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n^2}$ we know that we will be close to the infinite sum $\frac{\pi^2}{8}$ - indeed we know

$$\left| S_{99} - \frac{\pi^2}{8} \right| \leq \frac{1}{100^2}. \qquad \textit{Indeed I make } S_{99} \approx 0.8225175 \textit{ and } \frac{\pi^2}{8} \approx 0.8224670.$$

A series for which $\sum_{n=1}^{\infty} a_n$ converges but $\sum_{n=1}^{\infty} |a_n|$ diverges is called *conditionally convergent*, since they converge but the sum of their absolute values does not. The cancelation between the positive and the negative terms manages to make the series converge. These are the series that Riemann showed can behave strangely.

Theorem. *Riemann's Rearrangement Theorem.*

Suppose the series $\sum_{n=1}^{\infty} a_n$ converges but the series $\sum_{n=1}^{\infty} |a_n|$ diverges. Then for any total B there is a rearrangement of the series $\sum_{n=1}^{\infty} b_n$ tat converge to B.

However if a series satisfies $\sum_{n=1}^{\infty} |a_n|$ converges, than any rearrangement will converge to the same total as $\sum_{n=1}^{\infty} a_n$.

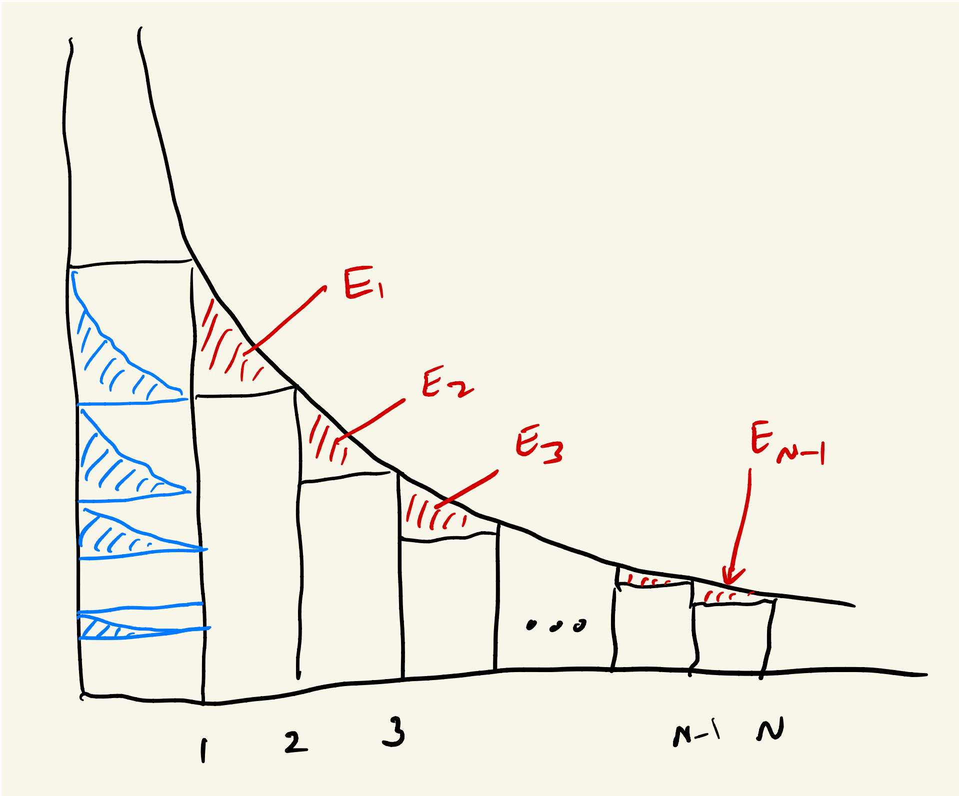
A rearrangement means that the sequence $(b_n : n \geq 1)$ contains all the terms of the sequence $(a_n : n \geq 1)$ but in a different order.

A renowned Warwick maths professor once suggested we omit mentioning conditionally convergent series in first year analysis, as they aren't that important. But the idea of cancelling lots of positive numbers with lots of negative numbers to obtain an interesting limit is important in several areas of analysis - not least in the construction of the stochastic Ito integral used in financial modelling. The proof of the theorem was non-examinable last year and I am happy to leave it so, but I may not be able to resist sketching the proof in lectures - we have all the tools necessary. In the rest of this chapter we will give a concrete example of a rearrangement with a different total to at least convince ourselves that it really does happen.

The **Harmonic Series Revisited**. We can give some nice examples of exactly summable series by re-examining the Harmonic series using the integral bounds technique. We set

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}.$$

I will mention a couple of problems (coupon collecting, quicksort) where the sums H_n occur in their solution. We can get an accurate estimate of H_n by using the integral bounds trick.



Matching areas in the picture we see that

$$\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}=\int_1^n \frac{1}{x}dx - E_1 - E_2 - \ldots - E_{n-1}$$

where the left hand side is the sum of the blocks and the the area under the curve between $x=1$ and $x=n$ with the small red shaded $E_k \geq 0$ areas subtracted.

The blue shaded areas are copies of the red shaded areas translated over to lie in a single column. They do not overlap and together they fill just part of the box $0 \leq x \leq 1,\ 0 \leq y \leq 1$. and we conclude that $E_1 + E_2 + \ldots + E_{n-1} \leq 1$.

Rearranging we find

$$H_n - \log(n) = \mathcal{E}_n.$$

where $\mathcal{E}_n = 1 - E_1 - E_2 - \ldots - E_{n-1} \geq 0$ is decreasing and so must converge by Weierstrass's criterion. We have proved the following lemma.

Lemma. ***Harmonic Series Lemma***

$$H_n = \log(n) + \mathcal{E}_n \qquad \textit{where} \quad \lim_{n\to\infty} \mathcal{E}_n = \gamma.$$

The constant γ is called Euler's constant (or sometimes the Euler-Mascheroni constant) and occasionally arises (for example describing the distribution of the primes in Number Theory). It cropped up for me in a probability problem concerning polynomials with random coefficients last year. You might come across it if you meet the integral $\int_0^\infty \log(x)e^{-x}dx$ whose value is exactly the value $-\gamma$. It's value is about 0.57.... It is unknown whether it is rational or not.

We can use it to resolve an example in the introductory chapter. Consider the series

$$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

The partial sum S_{2n} of the first $2n$ terms will have n positive and n negative terms:

$$\begin{aligned}
S_{2n} &= \left(1 + \frac{1}{3} + \frac{1}{5} + \ldots + \frac{1}{2n-1}\right) - \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \ldots + \frac{1}{2n}\right) \\
&= \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{2n}\right) - 2\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \ldots + \frac{1}{2n}\right) \\
&= \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{2n}\right) - \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right) \\
&= H_{2n} - H_n \\
&= (\log(2n) + \mathcal{E}_{2n}) - (\log(n) + \mathcal{E}_n) \\
&= \log(2) + \mathcal{E}_{2n} - \mathcal{E}_n \\
&\to \log(2) + \gamma - \gamma = \log(2).
\end{aligned}$$

We have used the Harmonic Series Lemma in the last two steps. Since also $S_{2n+1} = S_{2n} + \frac{1}{2n+1} \to \log(2)$ you can check that the whole sequence $S_n \to \log(2)$ and we have identified the exact value of $\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}$ as $\log(2)$.

The claim in the introduction was that we could add the terms in a different order and get a different sum. We tried alternately adding two of the positive terms and then one of the negative terms, that is

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \ldots$$

Consider the partial sum S_{3n}, which will have the first $2n$ positive terms and the first n

negative terms:

$$
\begin{aligned}
S_{3n} &= \left(1+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{4n-1}\right)-\left(\frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\ldots+\frac{1}{2n}\right) \\
&= \left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{4n}\right) \\
&\quad -\left(\frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\ldots+\frac{1}{4n}\right)-\left(\frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\ldots+\frac{1}{2n}\right) \\
&= H_{4n}-\frac{1}{2}H_{2n}-\frac{1}{2}H_n \\
&= (\log(4n)+\mathcal{E}_{4n})-\frac{1}{2}(\log(2n)-\mathcal{E}_{2n})-\frac{1}{2}(\log(n)-\mathcal{E}_n) \\
&= \frac{3}{2}\log(2)+\mathcal{E}_{4n}-\frac{1}{2}\mathcal{E}_{2n}-\frac{1}{2}\mathcal{E}_n \\
&\to \frac{3}{2}\log(2)+\gamma-\frac{1}{2}\gamma-\frac{1}{2}\gamma=\frac{3}{2}\log(2).
\end{aligned}
$$

You will be able now to check that the whole sequence $(S_n : n \geq 1)$ converges to $\frac{3}{2}\log(2)$ verifying the claim. A concrete demonstration of Riemann's Rearrangement theorem.

Chapter 4. Continuous functions

This chapter is taken from the first chunk of last years term 2 module by Keith Ball, as this chunk of material has been moved into our first term Calculus 1 module.

Continuity

We recall from Sets and Numbers some notation for functions. Given two sets A and B a *function* f from A to B assigns an element of B to each element of A. Thus every function comes "equipped" with two sets: A, the set of points where the function is defined, and B, a set to which the values of the function belong. We can draw attention to these sets by writing

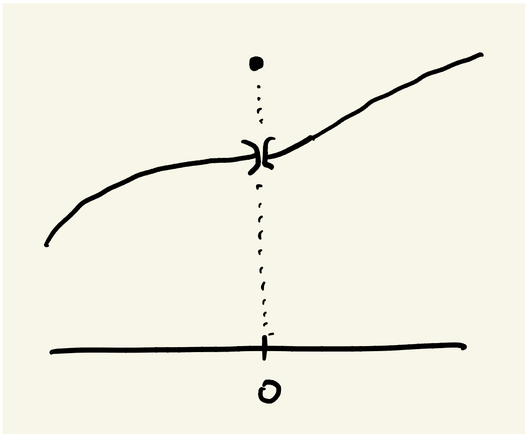
$$f : A \to B.$$

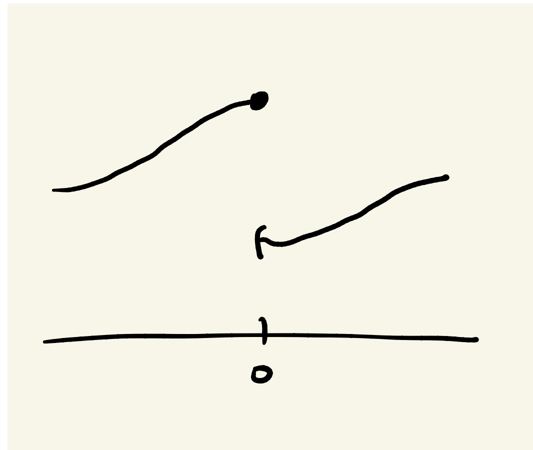
The set A is called the *domain* of f; B is called its *codomain*. For each element x of the domain we write $f(x)$ for the place to which x is sent: the *image* of x under f. The set of values $\{f(x) : x \in A\}$ is called the *range* of f.

An interval of the real line is a subset of $\mathbf{R}$ with the property that if $x < y < z$ and x and z both belong to the subset then so does y. So an interval contains all the points between its ends, but the ends themselves may or may not be included. Examples are

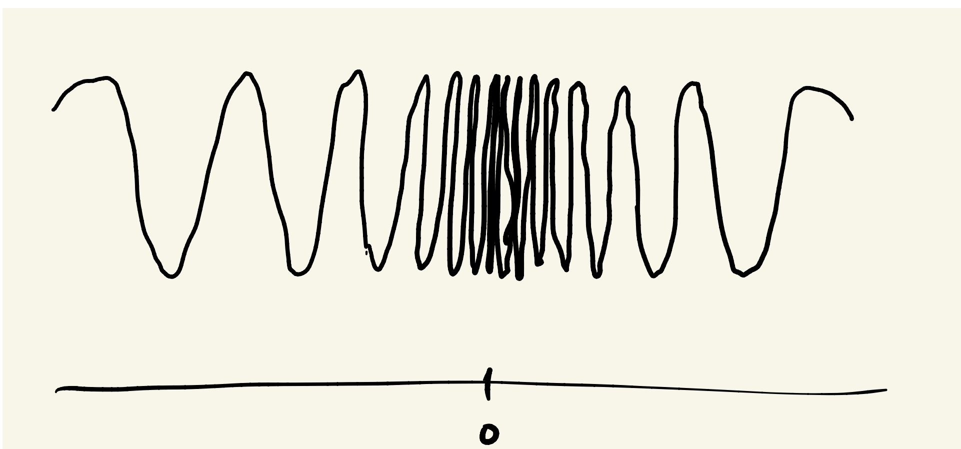
$$\begin{aligned}
\{x : a \leq x \leq b\} &= [a, b] \quad \text{a closed interval} \\
\{x : a < x < b\} &= (a, b) \quad \text{an open interval} \\
\{x : a \leq x < b\} &= [a, b) \quad \text{a half-open interval} \\
\{x : a \leq x\} &= [a, \infty) \quad \text{a half-infinite interval}
\end{aligned}$$

For this course we shall be considering real valued functions whose domains are intervals of the real line. We want to say what it means for such a function to be continuous. At each point of the domain there are essentially three types of problem that can occur. In each of the cases below, the function is discontinuous at 0 (but continuous elsewhere).





The first two are fairly straightforward. In the first case the function has a value at zero that doesn't "match" the rest of the function. In the second case, the function jumps in value as we pass through 0.



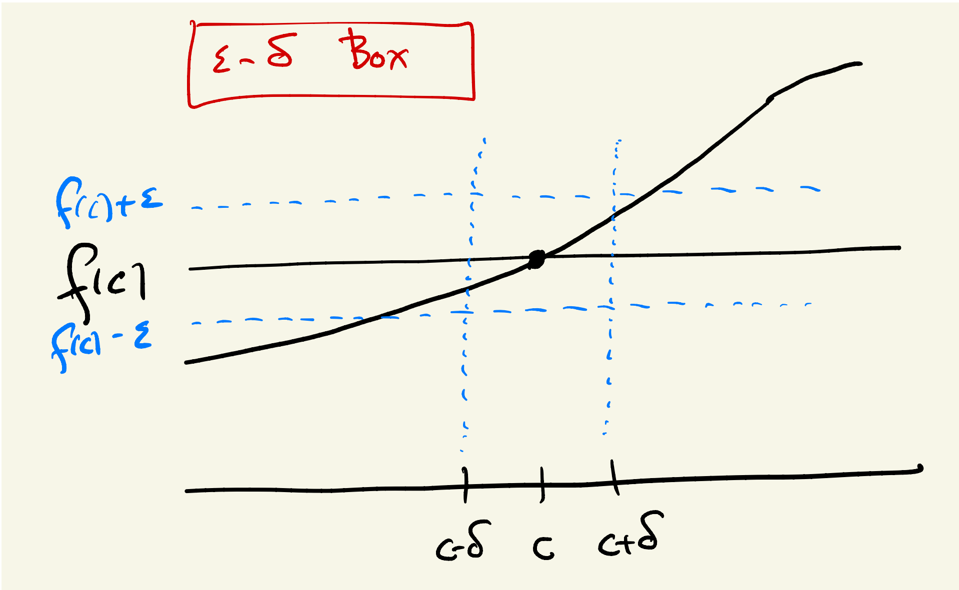
In the third situation $f(x)$ oscillates infinitely often over a wide range as x approaches 0.

We shall define what it means for a function to be continuous at a point in such a way as to rule out these problems. Then we say that a function is continuous on an interval if it is continuous at each point of the interval.

Definition (**Continuity**)**.** *A function* $f : I \to \mathbf{R}$ *defined on an interval* I *containing the number* c *is said to be* continuous *at* c *if for every* $\varepsilon > 0$*, there is a* $\delta > 0$ *so that if* $x \in I$ *and* $|x - c| < \delta$ *then*

$$|f(x) - f(c)| < \varepsilon.$$

The function is continuous on I *if it is continuous at each point of* I*.*



The picture shows the situation. The number ε specifies a band around the value $f(c)$. You want to be able to guarantee that $f(x)$ falls into that band as long as x is close enough to c.

Example (**Continuity of $x \mapsto x$**). *The function $x \mapsto x$ is continuous at every point of the real line.*

Proof If $f(x) = x$ then we can always take $\delta = \varepsilon$ since if $|x - c| < \varepsilon$ then

$$|f(x) - f(c)| = |x - c| < \varepsilon.$$

$\square$

As an exercise, you can check that constant functions are continuous.

Example (**A discontinuity**). *The function f given by*

$$f(x) = \begin{cases} 1 & \textit{if } x > 0 \\ 0 & \textit{if } x \leq 0 \end{cases}$$

is discontinuous at 0.

Proof We need to find a band around $f(0) = 0$ which we cannot guarantee to land in, merely by starting near 0. A band of width $\varepsilon = 1/2$ will do because if x is positive, however close to zero, $f(x) = 1$ and this is not within $1/2$ of $f(0)$. $\square$

We will do one more example which will also follow from our later arguments but will help to fix the ideas: the function $x \mapsto x^2$. The aim will be to show that if $x - c$ is small then so is $x^2 - c^2$. Now $x^2 - c^2 = (x - c)(x + c)$ so if $|x - c| < \delta$ then $|x^2 - c^2| < \delta|x + c|$ which looks good because the factor δ is small. But the second factor $|x + c|$ might be large (if $c = 1{,}000{,}000$ say). However, if x is close to c then $x + c$ is close to $2c$ which may be large but is a fixed number. So we shall cook up δ to compensate: roughly we want $\delta = \varepsilon/(2|c|)$. However we need to be a bit careful because $x + c$ isn't actually equal to $2c$.

Example (Continuity of $x \mapsto x^2$). *The function $x \mapsto x^2$ is continuous at every point of the real line.*

Proof Let $f(x) = x^2$, let $c \in \mathbf{R}$ and $\varepsilon > 0$. We want to guarantee that

$$|x^2 - c^2| < \varepsilon$$

by choosing x close enough to c. We know that $|x^2 - c^2| = |x - c|.|x + c|$.

If we choose $|x - c| < 1$ then $|x|$ cannot be larger than $|c| + 1$ and so $|x + c|$ cannot be larger than $2|c| + 1$. Then if we choose

$$|x - c| < \frac{\varepsilon}{2|c| + 1}$$

we will get

$$|x^2 - c^2| = |x - c|.|x + c| < \frac{\varepsilon}{2|c| + 1}(2|c| + 1) = \varepsilon$$

which is what we wanted.

We end up needing two conditions on $|x - c|$: namely $|x - c| < 1$ and $|x - c| < \frac{\varepsilon}{2|c|+1}$. But that is not a problem. If we choose δ to be the smaller of 1 and $\frac{\varepsilon}{2|c|+1}$ the two conditions will be satisfied simultaneously as long as $|x - c| < \delta$. $\square$

We wish to check the continuity of functions such as polynomials, more complicated than $x \mapsto x^2$. To do this we want a machine which allows us to build continuous functions from simpler ones. We want to know that when you add or multiply continuous functions the result is still continuous and also when you compose two continuous functions. There are a number of ways to do this but in view of what you already know about limits there is a particularly simple approach which depends upon rewriting the continuity property in terms of limits of sequences.

Theorem (**Sequential continuity**)**.** *Let $f : I \to \mathbf{R}$ be defined on the interval I and suppose that $c \in I$. Then f is continuous at c if and only if for every sequence (x_n) of points in I which converges to c,*

$$f(x_n) \to f(c) \quad \text{as } n \to \infty.$$

Proof Suppose first that f is continuous at c and $x_n \to c$. Then, given $\varepsilon > 0$ we can find $\delta > 0$ so that if $|x - c| < \delta$ then $|f(x) - f(c)| < \varepsilon$. Now choose N so that if $n > N$, $|x_n - c| < \delta$. Then if $n > N$ we have $|f(x_n) - f(c)| < \varepsilon$. So $f(x_n) \to f(c)$.

On the other hand suppose f is *not* continuous at c and choose $\varepsilon > 0$ with the property that whatever δ we pick there is a point x within δ of c for which

$$|f(x) - f(c)| \geq \varepsilon.$$

Now build a sequence as follows. For each n choose a point x_n with $|x_n - c| < 1/n$ but $|f(x_n) - f(c)| \geq \varepsilon$. Then $x_n \to c$ as $n \to \infty$ but $f(x_n)$ does not converge to $f(c)$. $\square$

The sequential continuity theorem immediately gives us the algebra of continuous functions.

Theorem (**Algebra of continuous functions**)**.** *Let $f, g : I \to \mathbf{R}$ be defined on the interval I and continuous at $c \in I$. Then*

1. *$f + g$ is continuous at c*
2. *$f.g$ is continuous at c*
3. *if $g \neq 0$ on I then f/g is continuous at c.*

Proof The proofs of all 3 are essentially the same. We shall do the first. We wish to show that if $x_n \to c$ then $(f + g)(x_n) = f(x_n) + g(x_n) \to f(c) + g(c)$ as $n \to \infty$. But we know that $f(x_n) \to f(c)$ and $g(x_n) \to g(c)$ so we can apply the properties of limits of sequences to conclude that $f(x_n) + g(x_n) \to f(c) + g(c)$. $\square$

From this we can immediately conclude that polynomials are continuous and that rational functions are continuous except where the denominator is zero.

Corollary (**Continuity of polynomials and rational functions**)**.** *If p is a polynomial then p is continuous at every point of $\mathbf{R}$. If $r = p/q$ is a ratio of two polynomials then it is continuous at every point of $\mathbf{R}$ where $q \neq 0$.*

Proof This will be on an exercise sheet. $\square$

To complete the continuity *machine* we need to know that we can compose continuous functions. The statement looks a bit complicated but the idea is as simple as for the algebraic properties.

Theorem (**Composition of continuous functions**)**.** *Let $f : I \to \mathbf{R}$ be defined on the interval I, $g : J \to I$ be defined on the interval J. If g is continuous at c and f is continuous at $g(c)$ then the composition $f \circ g$ is continuous at c.*

Proof Let (x_n) be a sequence in J converging to c. Then $g(x_n) \to g(c)$ in I and hence $f(g(x_n)) \to f(g(c))$. $\square$

The Intermediate Value Theorem

This section is devoted to a principle which has many uses and which you may have met at school. Roughly it says that if a continuous function crosses from below zero to above zero then there must be a point where it is equal to zero: it cannot *jump across* zero.

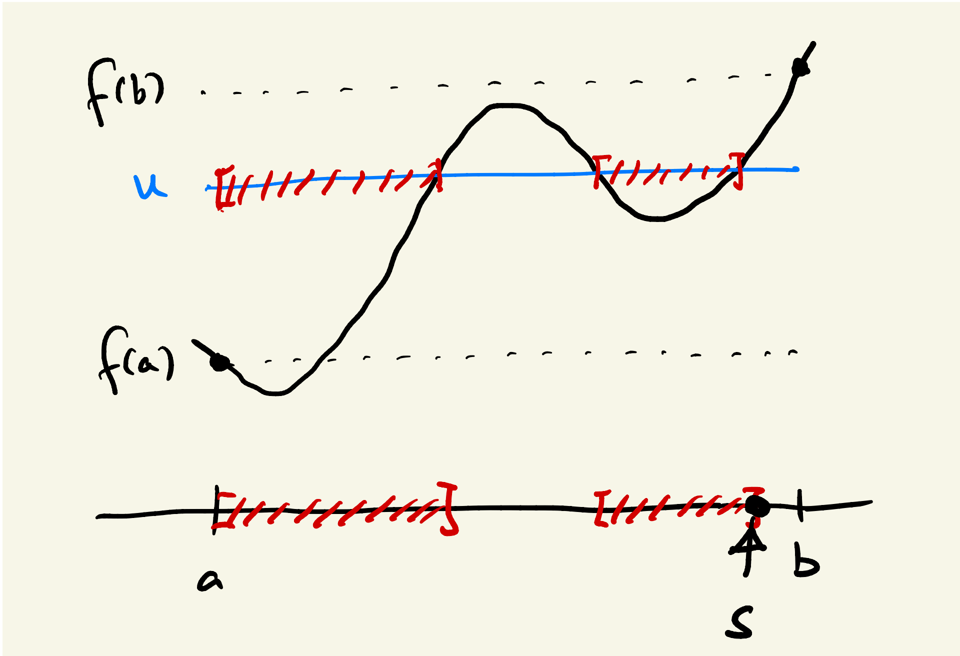
Our intuition about continuous functions is that we can draw them without taking pen from paper and therefore it should be "obvious" that a continuous function can't skip a value. However the proof is not quite so simple because we have defined continuity at each point individually. We therefore have to find the point where the function is supposed to take the correct value and then use continuity at that point. When we are trying to demonstrate the existence of a particular real number we usually have to invoke the completeness principle. The completeness principle is our way to finger a real number.

Theorem (**Intermediate Value Theorem**)**.** *Let $f : [a, b] \to \mathbf{R}$ be continuous and suppose that u lies between $f(a)$ and $f(b)$. Then there is a point c between a and b where $f(c) = u$.*

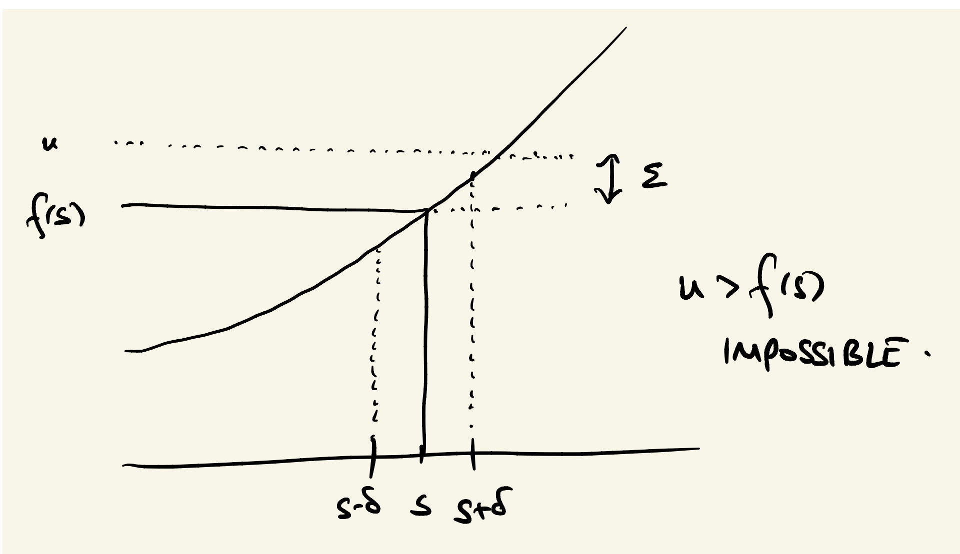
Proof Assume that $f(a) < u < f(b)$ and let A be the set

$$\{x \in [a, b] : f(x) \leq u\}.$$

This set is non-empty since it contains a and is bounded above by b. Let s be its least upper bound. The aim is to show that $f(s) = u$.



We shall eliminate the two other possibilities $f(s) < u$ and $f(s) > u$ separately. Suppose that we have $f(s) < u$.

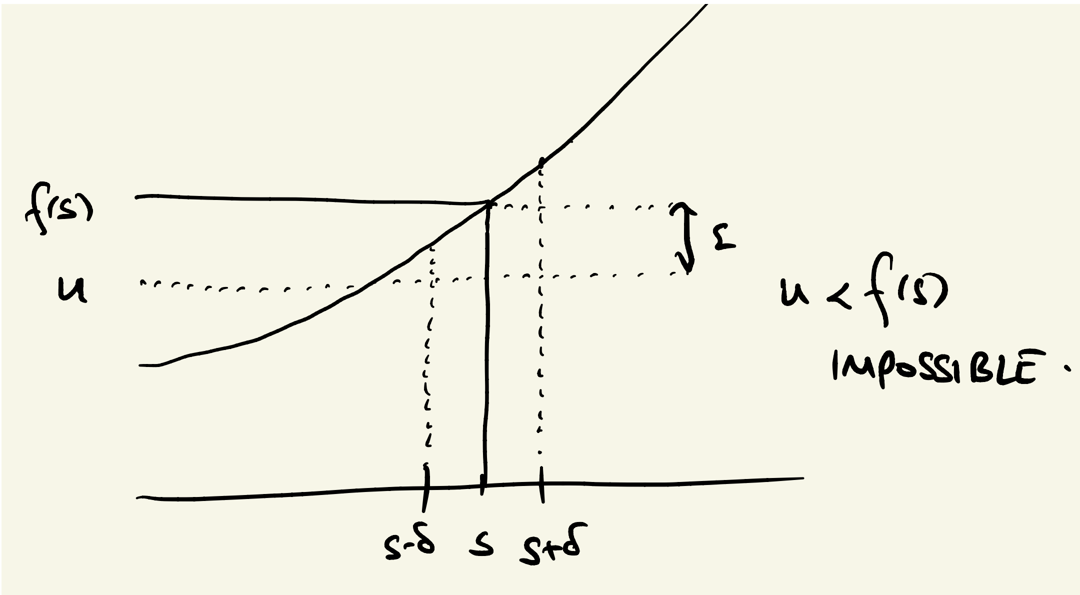


Note that that $s \neq b$ because we know $f(b) > u$. If we put $\epsilon = u - f(s)$ then for some $\delta > 0$

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

as long as $|x - s| < \delta$. So in particular if $x = s + \delta/2$ then $f(x) < f(s) + \epsilon = u$. This means that $s + \delta/2 \in A$ contradicting the fact that s is *an* upper bound for A.

Suppose on the other hand that $f(s) > u$.



Note that $s \neq a$ because $f(a) < u$. If we put $\epsilon = f(s) - u$ then for some $\delta > 0$

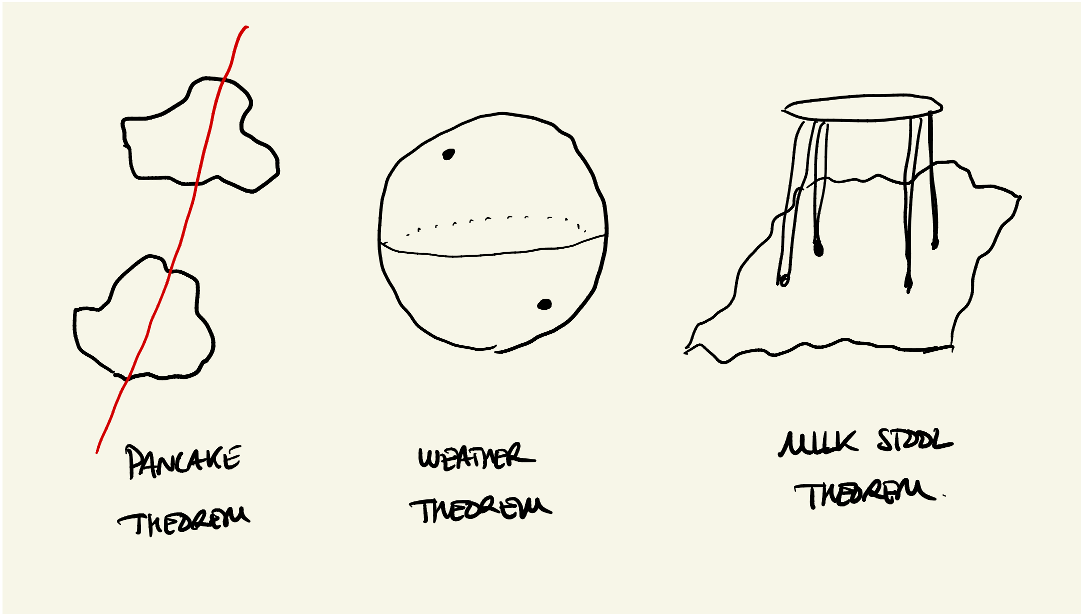
$$|f(x) - f(s)| < \epsilon$$

as long as $|x - s| < \delta$. Therefore if $s - \delta < x \leq s$, $f(x) > f(s) - \epsilon = u$ and hence x is *not* in A. This means that $s - \delta$ is an upper bound for A contradicting the fact that s is the *least* upper bound. □

Note that there were two things we had to check about the point s that we found: that $f(s)$ was not too small and that it was not too large. We used the two different properties of the least upper bound respectively for these two checks. This is a common feature of arguments using least upper bounds.

It is possible for a function to satisfy the intermediate value property without being continuous but it has to be a rather nasty function. Note that the IVT says that if u lies between $f(a)$ and $f(b)$ then there is a point c *between* a and b where $f(c) = u$. Try drawing a function with this property that is not continuous and you will be forced to wiggle a lot.

The IVT has many uses - in lectures I should have time to discuss some fo the following:



The simplest application is just for showing the existence of roots, as follows.

Example (**The existence of a solution to an equation**). *There is a solution of the equation* $x^3 + x - 1 = 0$ *between 0 and 1.*

Proof If $f(x) = x^3 + x - 1$ then $f(0) = -1 < 0$ while $f(1) = 1 > 0$. Apply IVT. $\square$

Whenever we have a continuous, strictly increasing function we expect to find an inverse for it: just as $y \mapsto \sqrt{y}$ is the inverse of the function $x \mapsto x^2$ on $[0, \infty)$. In order to make this precise we recall that the *range* of a function $f : A \to B$ is the set of values that f takes:

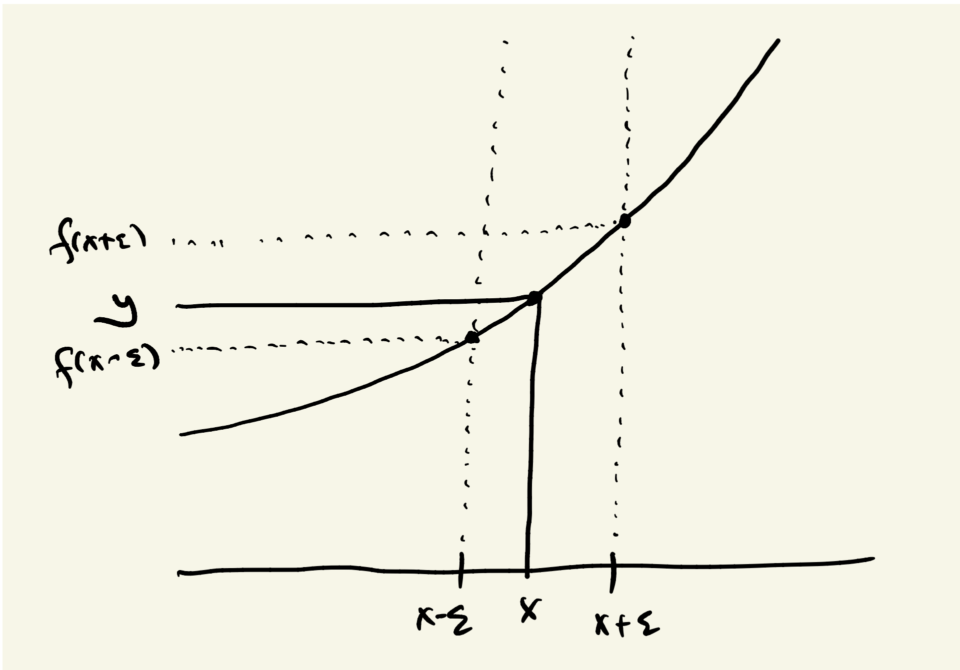
$$\{f(x) : x \in A\}.$$

Corollary (**Continuous image of an interval**). *If* $f : I \to \mathbf{R}$ *is continuous on the interval* I *then its range is an interval.*

Proof If x and y are in the range then by the IVT so is any point between x and y. So the range is an interval. $\square$

Theorem (**Existence of inverses**). *Let* $f : [a, b] \to \mathbf{R}$ *be continuous and strictly increasing. Then* f *has an inverse defined on its range and* f^{-1} *is continuous.*

The same holds if f is defined on an open interval or on the whole of $\mathbf{R}$. The proofs are essentially the same. The picture says it all. If you want to end up within ε of $x = g(y)$ then make sure you start between $f(x-\varepsilon)$ and $f(x+\varepsilon)$.



Proof Let $f(a) = c$ and $f(b) = d$. Since f is increasing all its values lie between c and d. So the range of f is exactly the interval $[c, d]$.

For each y there is an unique number x with $f(x) = y$ because f is strictly increasing, so we may define the inverse $g = f^{-1}$ by setting $g(y) = x$ in each case. Clearly g is strictly increasing. The final aim is to show that g is continuous. Suppose y lies in the open interval (c, d), $y = f(x)$ and $\varepsilon > 0$. We have $f(x-\varepsilon) < y < f(x+\varepsilon)$. Hence there is some $\delta > 0$ with the property that

$$f(x-\varepsilon) < y-\delta < y < y+\delta < f(x+\varepsilon).$$

For any z between $y-\delta$ and $y+\delta$ the fact that g is increasing ensures that

$$x-\varepsilon < g(z) < x+\varepsilon$$

and so $|g(z) - g(y)| < \varepsilon$.

If y is one of the end points then the argument is the same but with one of $f(x-\varepsilon)$ and $f(x+\varepsilon)$ replaced by the corresponding end point of $[c, d]$. $\square$

Corollary (**Existence of roots**)**.** *For each positive x and natural number n there is an unique positive n^{th} root $x^{1/n}$. The map*

$$x \mapsto x^{1/n}$$

is continuous.

Example (**A composition**)**.** *Let $f : \mathbf{R} \to \mathbf{R}$ be given by*

$$f(x) = \sqrt{1 + x^2}.$$

Then f is continuous on $\mathbf{R}$.

Boundedness and attainment of bounds

A continuous function on an open interval such as $(0, 1)$ may take arbitrarily large values. For example the function $x \mapsto 1/x$ approaches ∞ as x approaches 0. This particular problem can't occur if the function is continuous on the closed interval $[0, 1]$ since the function is supposed to approach $f(0)$ as x approaches 0. It turns out that indeed a continuous function on a closed interval must be bounded.

Theorem (**Boundedness of continuous functions**)**.** *Let $f : [a, b] \to \mathbf{R}$ be continuous. Then f is bounded.*

The proof uses a method that used to be called rather appealingly "the condensation of singularities". The idea is that if the function is unbounded we can find a sequence of points where it gets bigger and bigger. Using Bolzano-Weierstrass we find a subsequence of these points which converges. So a sequence of bad points converges to a point which we can conclude is so bad that it can't exist.

Proof Suppose f is unbounded. For each n choose $x_n \in [a, b]$ where $|f(x_n)| \geq n$. Now choose a subsequence x_{n_k} which converges to x say. Since the interval $[a, b]$ is closed we must have $x \in [a, b]$. Then we know that $f(x_{n_k}) \to f(x)$ but this can't happen since the values $f(x_{n_k})$ are becoming arbitrarily large. $\square$

At school you were often asked to find the maximum value of a function on some interval but probably never addressed the question of whether a nice function (a continuous one say) really must have a maximum. Obviously the previous example of $1/x$ shows that on an open interval a function need not have a maximum but this is also true even if the function is bounded. The function $x \mapsto x$ has no maximum on $(0, 1)$. Its least upper bound is 1 but there is no point in the interval where the function is equal to 1. Again, this situation can't happen on closed intervals.

Theorem (**Attainment of bounds**)**.** *Let $f : [a, b] \to \mathbf{R}$ be continuous. Then f has a maximum and a minimum attained in the interval.*

Proof Let M be the least upper bound of the set

$$\{f(x) : x \in [a, b]\}.$$

If there is no point in the interval where $f(c) = M$ then the function g given by $g(x) = M - f(x)$ is strictly positive and continuous on the interval. Hence by the algebra of continuous functions its reciprocal is continuous and therefore bounded. Let's say

$$\frac{1}{M - f(x)} \leq R$$

for all $x \in [a, b]$. Then $1/R \leq M - f(x)$ and hence $f(x) \leq M - 1/R$ for all x in the interval. This shows that $M - 1/R$ is an upper bound for the set of values and contradicts the fact that M is the *least* upper bound.

A similar argument works for the minimum. $\square$

This theorem will form the main tool in a proof of the *Mean Value Theorem* which plays a key role in many applications of the derivative and which is a highlight of Calculus 2.