ANALYSIS

REACHING FOR INFINITY

Steve Trettel

Table of contents

Preface

This is the beginnings of a textbook for a 1-year course on real analysis; the current version covers a semester and a half of material. This can be used as a 1-semester course by omitting some of the topics marked in the text as

- $\star$: optional topic independent of the main text (referenced only in later starred sections)
- ⧫: content used in main text, but only to prove some supporting or readily-believed fact: these arguments can be skipped or skimmed with little ill effect.
- $\oplus$: additional proof of a result which is proved by different (often cleaner) means elsewhere

The sections ‘Elementary Functions’ present across the second half of the text are self-contained and could be omitted from a course culminating with the Fundamental Theorem of Calculus, but will be an integral part of the eventual year-long course.

If you enrolled in my Spring 2025 course the homework assignments are available here.

Troubles with the Infinite

Real analysis is born out of our desire to understand infinite processes, and to overcome the difficulties raised by taking infinity seriously in this way. To appreciate this, we begin with an overview of some famous results from antiquity, as well as several paradoxes that arise from taking them seriously, if we are not careful.

The Diagonal of a Square

Around 3700 years ago, a babylonian student was assigned a homework problem, and their work (in clay) fortuitously survived until the modern day.

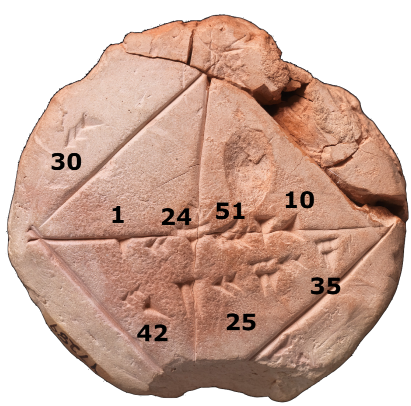


Figure 1.: Tablet YBC-7289

The problem involved measuring the length of the diagonal of a square of side length 1/2, which involves the square root of 2. The tablet records a babylonian approximation to $\sqrt{2}$ (Though it does so in base 60, where the 'decimal' expression is 1.(24)(51)(10))

$$\sqrt{2} \approx \frac{577}{408} \approx 1.414215686 \cdots$$

Definition 0.1 (Base Systems for Numerals)**.** If $b > 1$ is a positive integer, *base-b* refers to expressing a number in terms of powers of b. In base 10 we write 432 to mean $4 \cdot 10^2 + 3 \cdot 10^1 + 2 \cdot 10^0$, whereas in base 5 the string of digits 432 would denote $4 \cdot 5^2 + 3 \cdot 5^1 + 2 \cdot 5^0$.

Numbers between 0 and 1 can also be expressed in a base system, using *negative powers of the base.* In base 10, 0.231 means $2 \cdot 10^{-1} + 3 \cdot 10^{-2} + 1 \cdot 10^{-3}$, whereas in base 5 the same string of digits would denote $2 \cdot 10^{-1} + 3 \cdot 5^{-2} + 1 \cdot 5^{-3}$.

The babylonians used base 60, meaning all numbers were written as a series in 60^n for n ranging over the integers. This tablet records the approximate square root of 2 as

$$1.(24)(51)(10)$$

Which, in base 60 denotes

$$\begin{aligned}
\sqrt{2} &\approx 1 \cdot 60^0 + 24 \cdot 60^{-1} + 51 \cdot 60^{-2} + 10 \cdot 60^{-3} \\
&= 1 + \frac{24}{60} + \frac{51}{60^2} + \frac{10}{60^3} \\
&= 1 + \frac{24}{60} + \frac{17}{1200} + \frac{1}{21600} \\
&= \frac{577}{408}
\end{aligned}$$

Exercise 0.1. By inscribing a regular hexagon in a circle, the Babylonians approximated π to be $25/8$. Compute the base 60 'decimal' form of this number.

The tablet itself does not record how the babylonians came up with so accurate an approximation, but we have been able to reconstruct their reasoning in modern times

Example 0.1 (Babylonian Algorithm Computing $\sqrt{2}$)**.** Starting with a rectangle of area 2, call one of its sides x. If the rectangle is a square, then $x = \sqrt{2}$ exactly. And the closer our rectangle is to a square, the closer x is to $\sqrt{2}$. Thus, starting from this rectangle, we can build an *even better approximation* by making it more square. Precisely, the side lengths of this rectangle are x and $2/x$, and a rectangle with one side the *average* of these two numbers, will be closer to a square than this one.

Starting from a rectangle with side lengths 1 and 2, applying this procedure once improves our estimate from 1 to $3/2$, and then applying it again improves it to $577/408$. This Babylonian approximation is just the third element in an *infinite sequence* of approximations to $\sqrt{2}$

Exercise 0.2 (Babylonian Algorithm Computing $\sqrt{2}$)**.** Carry out this process, and show you get $577/408$ as the third approximation to $\sqrt{2}$. What's the next term in the sequence? How many decimal places is this accurate to in base 10? (Feel free to use a calculator of course!)

Exercise 0.3 (Computing Cube Roots)**.** Can you modify the babylonians procedure which found approximates of $\sqrt{2}$ to instead find rational approximates of $\sqrt[3]{2}$?

Here, instead of starting with a rectangle of sides x, y let's start with a three dimensional brick with a square base (sides x and x), height y, and area 2. Our goal is to find a "closer to cube" shaped brick than this one, and then to iterate. Propose a method of getting "closer to cube-shaped" and carry it out: what are the side lengths of the next shape in terms of x and y?

Start with a simple rectangular prism of volume 2 and iterate this procedure a couple times to get an approximate value of $\sqrt[3]{2}$. How close is your approximation?

It is clear from other Babylonian writings that they knew this was merely an approximation, but it took over a thousand years before we had more clarity on the nature of $\sqrt{2}$ itself.

Pythagoras

We often remember the Pythagoreans for the theorem bearing their name. But while they did prove this, the result (likely without proof) was known for millennia before them. The truly new, and shocking contribution to mathematics was the discovery that there must be numbers beyond the rationals, if we wish to do geometry.

Theorem 0.1 ($\sqrt{2}$ is irrational)**.** *There is no fraction p/q which squares to* 2*.*

To give a proof of this fact we need one elementary result of number theory, known as Euclid's Lemma (which says that if a prime p divides a product ab, then p must divide either a or b).

Proof. (Sketch) Assume p/q is in lowest terms, and squares to 2. Then $p^2/q^2 = 2$ so $p^2 = 2q^2$. Thus 2 divides p^2, so in fact 2 divides p (Euclid's lemma), meaning p is even.

Thus, we can write $p = 2k$ for some other integer k, which gives $(2k)^2 = 2q^2$, or $4k^2 = 2q^2$. Dividing out one factor of 2 yields $2k^2 = q^2$< so 2 divides q^2, and thus (Euclid's lemma, again) 2 divides q.

But now we've found that both p and q are divisible by 2, which means p/q is not in lowest terms after all, a contradiction! Thus there can not have been any fraction squaring to 2 in the first place. □

Exercise 0.4. Following analogous logic, prove that $\sqrt{3}$ is irrational. Generalize this to prove that $\sqrt{6}$ is irrational. But be careful! Make sure that your proof doesn't also apply to $\sqrt{9}$ (which of course, IS rational).

Knowing now that $\sqrt{2}$ is irrational, it is clear that the Babylonian procedure will never *exactly* return the correct answer, as if it starts with a rationally-sided rectangle, it'll always produce another with rational side lengths. But its a natural question to wonder just *how good* are the babylonian approximations?s

Definition 0.2 (The Babylonian Algorithm and Number Theory)**.** Because $\sqrt{2}$ is irrational, there is no pair of integers p, q with $p^2 = 2q^2$. Good rational approximations to $\sqrt{2}$ will *almost* satisfy this equation, and we will call an approximation *excellent* if it is only off by 1: that is p/q is an excellent approximation if

$$p^2 = 2q^2 + 1$$

Exercise 0.5 (The Babylonian Algorithm and Number Theory)**.** Prove that all approximations produced by the babylonian sequence starting from the rectangle with sides 1 and 2 are excellent, by induction.

To acomodate this discovery, the Greeks had to *add a new number to their number system* - in fact, after really absorbing the argument, they needed to add many. Things like $\sqrt{3}$, but also

$$\sqrt{1+\sqrt{3-\sqrt{2+\frac{\sqrt{3}+\sqrt{2}}{5}}}}$$

are called *constructible numbers*, as they were constructed by the greeks using a compass and straightedge, to extend the rational numbers.

Quadrature of the Parabola

The idea to compute some seemingly unreachable quantity by a succession of better and better approximations may have begun in babylon, but truly blossomed in the hands of Archimedes.

In his book *The Quadrature of the Parabola*, Archimedes relates the area of a parabolic segment to the area of the largest triangle that can be inscribed within.

Theorem 0.2. *The area of the segment bounded by a parabola and a chord is* $4/3^{rd}s$ *the area of the largest inscribed triangle.*

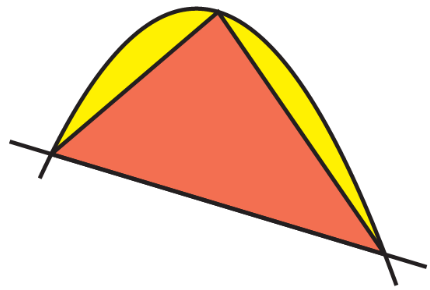


Figure 2.: A parabolic region and its largest inscribed triangle

After first describing how to find the largest inscribed triangle (using a calculation of the *tangent lines* to a parabola), Archimedes notes that this triangle divides the remaining region into two more parabolic regions. And, he could fill these with their largest triangles as well!

These two triangles then divide the remaining region of the parabola into *four new parabolic regions*, each of which has their own largest triangle, and so on.

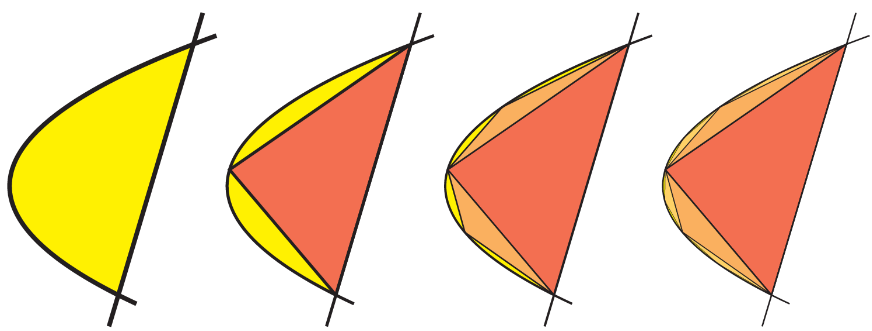


Figure 3.: Archimedes' infinite construction of the parabolic segment from triangles

Archimedes proves that in the limit, after doing this infinitely many times, the triangles *completely fill* the parabolic segment, with zero area left over. Thus, the only task remaining is to add up the area of these infinitely many triangles. And here, he discoveries an interesting pattern.

We will call the first triangle in the construction *stage 0* of the process. Then the two triangles we make next comprise *stage 1*, the ensuing four triangles *stage 2*, and the next eight *stage 3*.

Proposition 0.1 (Area of the n^{th} stage)**.** *The total area of the triangles in each stage is* $1/4$ *the total area of triangles in the previous stage.*

If A_n is the area in the n^{th} stage, Archimedes is saying that $A_{n+1} = \frac{1}{4}A_n$. Thus

$$A_0 = T \quad A_1 = \frac{1}{4}T \quad A_2 = \frac{1}{16}T \quad A_3 = \frac{1}{64}T \ldots$$

And the total area A is the infinite sum

$$\begin{aligned} A &= T + \frac{1}{4}T + \frac{1}{16}T + \frac{1}{64}T + \cdots \\ &= \left(1 + \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \cdots\right)T \end{aligned}$$

Now Archimedes only has to sum this series. For us moderns this is no trouble: we recognize this immediately as a geometric series

But why is it called *geometric*? Well (this is not the only reason, but...) Archimedes was the first human to sum such a series, and he did so completely geometrically. Ignoring the leading 1, we can interpret all the fractions as proportions of the area of a square. The first term 1/4 tells us to take a quarter of the square, the next term says to take a quarter of a quarter more, and so on. Repeating this process infinitely, Archimedes ends up with the following figure, where the highlighted squares on the diagonal represent the completed infinite sum.

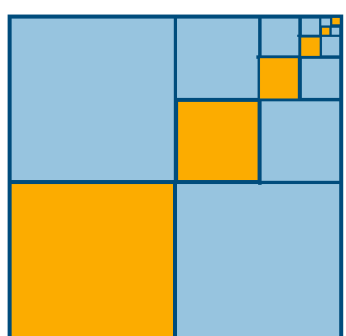


Figure 4.: The infinite process: $1/4 + 1/16 + 1/64 + \cdots$

He then notes that this is precisely one third the area of the bounding square, as two more identical copies of this sequence of squares fill it entirely (just slide our squares to the left, or down). Thus, this infinite sum is precisely 1/3, and so the total area is 1 plus this, or 4/3.

This tells us an important fact, beyond just the area of the parabola we sought! We were looking to compute the area of a *curved shape*, and the procedure we found could never give us the answer exactly, but only an infinite sequence of better approximations. Being acquainted with the work of Pythagoras and the Babylonians, this might have well led us to conjecture that the area of the parabola must be *irrationally* related to the area of the triangle. But Archimedes showed this is not the case; our infinite sum here evaluates to a *rational number*, $4/3$!

Infinite sequences of rational numbers can sometimes produce a wholly new number, and sometimes just converge to another rational.*

How can we tell? This is one motivating reason to develop a rigorous study of such objects. But it gets even more important, if we try to generalize Archimedes' argument.

Troubles with Geometric Series

Archimedes' quadrature of the parabola represents a monumental leap forward in human history. This is the first time in the mathematical literature where infinity is not treated as some distant ideal, but rather a real place that *can be reached.* And the argument itself is an absolute classic - involving the first occurrence of an infinite series in mathematics, and a wonderfully geometric summation method (hence the name *geometric series*, which survives until today). The elegance of Archimedes' calculation is almost dangerous - its easy to be blinded by its apparent simplicity, and – like Icarus – fly too close to the sun, falling from these heights of logic directly into contradiction.

Archimedes visualized his argument for the sum $\sum \frac{1}{4^n}$ as though it was occurring *inside* of a larger square, but there's another perspective we could take. Call the total sum S,

$$S = 1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \cdots$$

and note that multiplying S by $1/4$ is the same as removing the first term, as it shifts all the terms down by one space:

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

Thus, $\frac{1}{4}S = S - 1$, and we can *solve* this algebraic equation directly to find $S = 4/3$. The beauty of this argument is that unlike Archimedes' original, its not tied to the number $1/4$ at all! Imagine we took some number r, and we wanted to add up the infinite sum

$$1 + r + r^2 + r^3 + r^4 + r^5 + r^6 + r^7 + \cdots + r^n + \cdots$$

Call that sum S, and notice that we have the same property, multiplying the sum by r shifts every term down by one, so we get the same result as if we just removed the first term:

$$rS = S - 1$$

We can then solve this for S and get

$$S = \frac{1}{1-r}$$

This gives us what we expect when $r = 1/4$, and trying it for other fractions, like $r = 1/5$ or $r = 23/879$, we can confirm (with the help of a computer) that the infinite sum really does approach the value this formula gives!

Amazingly, it even works for negative numbers, after we think about what this means. If $r = \frac{-1}{2}$ then

$$1 + r + r^2 + r^3 + r^4 + r^5 + \cdots = 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \cdots$$

Using our formula above we see that this is supposed to converge to

$$S = \frac{1}{1 - \left(\frac{-1}{2}\right)} = \frac{1}{1 + \frac{1}{2}} = \frac{1}{\frac{3}{2}} = \frac{2}{3}$$

And, using a computer to add up the first 100 terms we see

$$S \approx 0.66666666666666666666666666666692962030174033726847057618$$

This is pretty incredible, as our original geometric reasoning doesn't make sense for $r = -1/2$, but the algebra works just fine! We may also wish to investigate what happens when $r = 1$, which would give

$$S = 1 + 1 + 1 + 1 + 1 + 1 + \cdots$$

This is going off to infinity, and our formula gives $S = 1/(1-1) = 1/0$, which could make sense: we could even take this as an indication that we should define $1/0 = \infty$. But things get more interesting with $r = -1$. Here the sum is

$$S = 1 - 1 + 1 - 1 + 1 - 1 + 1 - 1 + 1 - 1 + 1 - \cdots$$

As we add this up term by term, we first have 1, then 0, then 1 then 0, over and over agan as we repeatedly add a 1, and then immediately cancel it out. This isn't getting close to any number at all! But our formula gives

$$S = \frac{1}{1-(-1)} = \frac{1}{2}$$

Now we have a real question - did we just **discover a new, deep fact of mathematics** - that we can sensibly assign values to series like this, that we weren't originally concerned with, or did we **discover a limitation of our theorem**? This is an interesting, and important question to come out of our playing around!

Thus far, we haven't seen any cases where our theorem has output any 'obviously' wrong answers, so we may be inclined to trust it. But this does not hold up to further scrutiny: what about when $r = 2$? Here the sum is

$$1 + 2 + 4 + 8 + 16 + 32 + \cdots$$

which is clearly going to infinity. But our formula disagrees, as it would have you belive the sum is $S = 1/(1-2) = -1$. This raises the more general problem: **when working with infinity, sometimes a formula you derive works, and sometimes it doesn't. How can you tell when to trust it?**

Exercise 0.6. Explain what goes wrong with the argument when $r = 2$...

The Circle Constant

The curved shape that everyone was *really* interested in was not the parabola, but the circle. Archimedes tackles this in his paper *The Measurement of the Circle*, where he again constructs a finite sequence of approximations built from triangles, and then reasons about the circle *out at infinity*. First, we need a definition:

Definition 0.3 (π and τ)**.** The area of the unit circle is denoted by the constant π. The circumference of the unit circle is denoted by the constant τ.

Archimedes came up with a sequence of overestimates, and underestimates for π by inscribing and circumscribing regular polygons.

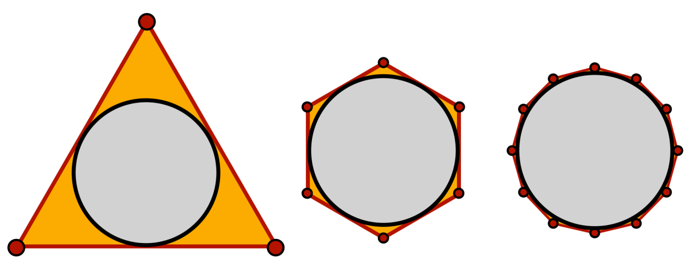


Figure 5.: Circumscribed polygons provide an overestimate of the area of the circle.

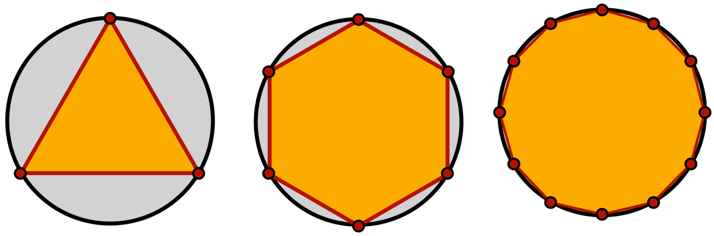


Figure 6.: Inscribed polygons provide an underestimate of the area of the circle.

Any polygon inside the unit circle gave an underestimate, and any polygon outside gave an overestimate. The more sides the polygon had, the better the approximations would be.

Calculating the area and perimeter of regular n-gons is (theoretically) straightforward, as they can be decomposed into $2n$ right triangles. Drawing a diagram, we find the relations below;

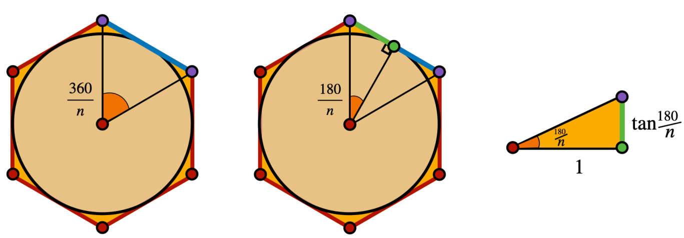


Figure 7.: Decomposing a circumscribed polygon into right triangles.

Proposition 0.2 (Area of a Circumscribed Polygon)**.** *The area of a regular n-gon circumscribing the unit circle is given by*

$$\begin{aligned} C_n &= 2n \cdot \left(\frac{1}{2} \cdot 1 \cdot \tan\frac{180}{n}\right) \\ &= n \tan\frac{180}{n} \end{aligned}$$

Proposition 0.3 (Perimeter of a Circumscribed Polygon)**.** *The perimeter of a regular n-gon circumscribing the unit circle is given by*

$$P_n = 2n \cdot \tan\frac{180}{n}$$

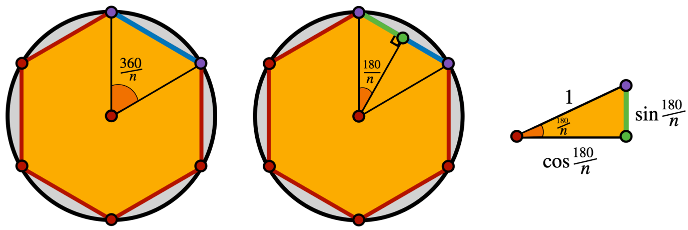


Figure 8.: Decomposing an inscribed polygon into right triangles.

Proposition 0.4 (Area of a Inscribed Polygon)**.** *The area of a regular n-gon inscribed in the unit circle is given by*

$$\begin{aligned} a_n &= 2n \cdot \left(\frac{1}{2} \cdot \cos\frac{180}{n} \cdot \sin\frac{180}{n}\right) \\ &= \frac{n}{2}\sin\frac{360}{n} \end{aligned}$$

Where we used the trigonometric identity $\sin(2x) = 2\sin x \cos x$ to simplify a_n above.

Proposition 0.5 (Perimeter of a Inscribed Polygon)**.** *The perimeter of a regular n-gon inscribed in the unit circle is given by*

$$p_n = 2n \cdot \sin\frac{180}{n}$$

Using these, Archimedes calculated away all the way to the 96-gon, which provided him with the estimates

$$\frac{223}{71} < \pi < \frac{22}{7}$$

This was the best estimate of π calculated during the classical period of the Greeks, but the same method was applied by Chinese mathematician Zu Chongzi in the 400s CE to much *much* larger polygons.
Working with the $24,576$-gon, he found

$$\frac{355}{113} < \pi < \frac{22}{7}$$

The lower bound here, $355/113$ is the best possible rational approximation of π with denominator less than four digits, and equals $3.14159292\cdots$, whereas $\pi = 3.14159265\cdots$. This was the most accurate approximate to π calculated anywhere in the world for over 800 years, and was only surpassed in the late 1300s by Indian mathematician Madhava, about whom we'll learn more soon.

Remark 0.1. The next best rational approximation is $\frac{52163}{16604}$, which is a significantly more complicated looking fraction!

Proving $\tau = 2\pi$

While impressive, Archimedes' main goal was not the *approximate* calculation above, but rather an *exact theorem.* He wanted to understand the true relationship between the area and perimeter of the circle, and wished to use these approximations as a guide to what is happening with the real circle, "out at infinity".

To understand this case, Archimedes argues that as n goes to infinity, the sequences of inscribed and circumscribed polygons approach the circle, and so *in the limit*, the sequences of areas must tend to the area of the circle (π) and the sequences of perimeters must tend to the perimeter of the circle (τ).

$$A_n \to \pi \qquad P_n \to \tau$$

But, now look carefully at the form of the expressions we derived for the circumscribing polygons in Proposition 0.2 and Proposition 0.3:

$$A_n = n \cdot \tan \frac{180}{n} \qquad P_n = 2n \cdot \tan \frac{180}{n}$$

Here, we do not need to worry about explicitly calculating A_n or P_n; all we need to notice is that the perimeter is *exactly* twice the area, $P_n = 2A_n$! This makes sense:

- Each polygon is built out of n triangles.
- The area of a triangle is half its base times its height
- The height of each triangle is 1 (the radius of the circle)
- Thus, the area the sum of half all the bases, or half the perimeter!

But since this exact relationship holds for every single value of n, Archimedes argued it must also be true in the limit, so the perimeter is twice the area:

Theorem 0.3 (Archimedes)**.**

$$\tau = 2\pi$$

Troubles With Limits

Archimedes again leaves us with an argument so elegant and deceptively simple that its easy to under-appreciate its subtlety and immediately fall prey to contradiction. What if we attempt to repeat Archimedes argument, but with a different sequence of polygons approaching the circle?

Remark 0.2. To be fair to the master, Archimedes is *much, much* more careful in his paper than I was above, so part of the apparent simplicity is a consequence of my omission.

For example, what if we start with a square circumscribing the circle, and then at each stage produce a new polygon with the following rule:

- At each corner of the polygon, find the largest square that fits within the polygon, and remains outside the circle. Then remove this square.

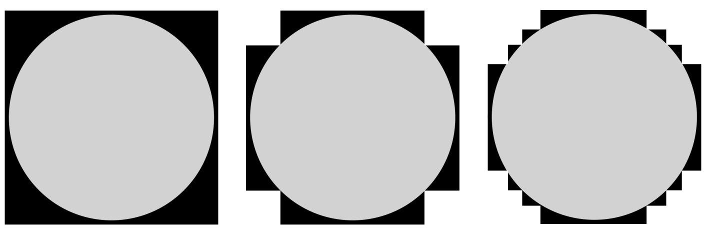


Figure 9.: Iteratively removing the largest square outside the circle at each vertex produces a sequence of right angled polygons which converges to the circle.

Exactly like in Archimedes' example this sequence of polygons approaches the circle as we repeat over and over. In fact, in the limit - this sequence *literally becomes the circle* (meaning that after infinitely many steps, there are no points of the resulting shape remaining outside the circle at all). Thus, just as for our original sequence of polygons, we expect that the areas and perimeters of these shapes approach the areas and perimeters of the circle itself. That is,

$$A_n \to \pi, \qquad P_n \to \tau$$

While the behavior of A_n takes a bit of work to understand, this sequence of polygons is constructed to make analyzing the perimeters particularly nice. Look what happens at each stage near a dent: two edges are turned inward to the circle, but do not change in length.

Figure 10.: Removing a square at a vertex does not change the perimeter of the polygon, as it replaces two segments with two other segments of the same length.

Since adding a dent does not change the length of the perimeter, each polygon in our sequence has *exactly the same perimeter* as the original! The original perimeter is easy to calculate, each side of the square is a diameter of the unit circle, so its total perimeter is 8. But since this both does not change *and* converges in the limit to the circles circumference, we have just derived the amazing fact that

$$\tau = 8$$

This is inconsistent with what we learn from Archimedes' argument which shows that $\pi < 22/7$ and $\tau = 2\pi$, so $\tau < 44/7 = 6.2857\ldots$. It appears that we have applied the same argument twice, and found a contradiction in comparing the results!

Exercise 0.7 (Convergence to the Diagonal)**.** We can run an argument analogous to the above which proves that $\sqrt{2} = 2$, by looking at a sequence of polygons that converge to a right triangle with legs of length 1. Let T_0 denote the unit square, and T_n

Prove that as n goes to infinity the area of the polygons T_n do converge to the area of the triangle (Hint: can you write down a formula for the *total error* between T_n and the triangle?) Also, prove that the length of the zig-zag diagonal side of the T_n has length 2 always, independent of n. Thus, the limit of the zigzag, which becomes the hypotenuse of the triangle, has length 2!

But the pythagorean theorem tells us that its length must be $\sqrt{1^2+1^2} = \sqrt{2}$, so in fact we have proven $\sqrt{2} = 2$, or $2 = 4$, a contradiction in mathematics.

Its quite difficult to pinpoint exactly what goes wrong here, and thus this presents a particularly strong argument for why we need analysis: without a rigorous understanding of infinite processes and limits, we can never be sure if our seemingly reasonable calculations give the right answers, or lies!

Estimating π

With our modern access to calculator technology, the trigonometric formulas above essentially solves the problem: for example, plug in $n = 96$ to a calculator (set to degrees!) to replicate the work of Archimedes in one click.

But this poses a historical problem: of course the ancients did not have a calculator, so how did they compute such accurate approximations millennia ago? And there's also a potential logical problem lurking in the background: inside our calculator there is some algorithm computing the trigonometric functions, and perhaps that algorithm depends on already knowing something about the value of π. If so, using this calculator to give a from-first-principles estimate of π would be circular!

To compute their estimates, both Archimedes and Zu Chongzi landed on an idea similar to the Babylonians and their computation of $\sqrt{2}$: they found an *iterative procedure* that starts with one polygon, and doubles its number of sides. With such a procedure in hand, they could start with any polygon and rapidly scale it up to better and better estimates. Beginning with an hexagon, Archimedes only needed to double four times:

$$6 \to 12 \to 24 \to 48 \to 96$$

Exercise 0.8 (The Doublings of Zu Chongzi)**.** How many times did Zu Chongzi double the sides of a hexagon to reach the 24,576 gon?

Following Archimedes, we'll look at the doubling procedure for the perimeter of inscribed polygons: given p_n we seek a method to compute p_{2n}. By the formula in Proposition 0.4, it is enough to be able to compute $\sin(360/(2n))$ in terms of $\sin(360/n)$, that is, we need to be able to compute the sine of half the angle. The half-angle identities from trigonometry prove helpful here:

Definition 0.4 (Half Angle Identities)**.**

$$\cos\left(\frac{\theta}{2}\right) = \sqrt{\frac{1+\cos\theta}{2}} \qquad \sin\left(\frac{\theta}{2}\right) = \sqrt{\frac{1-\cos\theta}{2}}$$

$$\tan\left(\frac{\theta}{2}\right) = \sqrt{\frac{1-\cos\theta}{1+\cos\theta}} = \frac{\sin\theta}{1-\cos\theta} = \frac{1-\cos\theta}{\sin\theta}$$

Also making use of the pythagorean identity $\sin^2\theta + \cos^2\theta = 1$, we can compute as follows:

$$\begin{aligned}
\sin\frac{\theta}{2} &= \sqrt{\frac{1-\cos\theta}{2}} \\
&= \sqrt{\frac{1-\sqrt{\cos^2\theta}}{2}} \\
&= \sqrt{\frac{1-\sqrt{1-\sin^2\theta}}{2}}
\end{aligned}$$

Lets write $s_n = \sin(180/n)$ for brevity. Then, the above formula tells us how to compute s_{2n} if we know s_n:

$$s_{2n} = \sqrt{\frac{1-\sqrt{1-s_n^2}}{2}}$$

This sort of relationship is called a *recurrence relation*, or a *recursively defined sequence* as it tells us how to compute the next term in the sequence if we have the previous one. Notice there are no more trigonometric formulas in the recurrence - so if we can find the value s_n for *any polygon*, we can start with that, and iteratively double.

Example 0.2 (A Recurrence for p_n)**.** By Proposition 0.5, we see that $p_n = 2ns_n$. Thus $p_{2n} = 2(2n)s_{2n} = 4s_{2n}$, and using the recurrence for s_{2n} we see

$$\begin{aligned}
p_{2n} &= 4ns_{2n} \\
&= 4n\sqrt{\frac{1-\sqrt{1-s_n^2}}{2}} \\
&= 2n\sqrt{2-2\sqrt{1-s_n^2}} \\
&= 2n\sqrt{2-\sqrt{4-4s_n^2}}
\end{aligned}$$

But, since $s_n = p_n/(2n)$, substituting this in gives a relation between p_{2n} and p_n directly:

$$\begin{aligned}
p_{2n} &= 2n\sqrt{2-\sqrt{4-4s_n^2}} \\
&= 2n\sqrt{2-\sqrt{4-\left(\frac{p_n}{n}\right)^2}}
\end{aligned}$$

The incredible fact: even though we used trigonometry to derive this recurrence, we do not need to know how to evaluate any trigonometric functions to actually use it! All we need to be able to do is find the perimeter of *some* inscribed n-gon, and then we can repeatedly double over and over!

But how can we get started? A beautiful observation of Archimedes was that a regular hexagon inscribed in the circle has perimeter *exactly equal to 6*, as it can be decomposed into six equilateral triangles, whose side length is the circle's radius. And with that, we are off!

Example 0.3 (The Perimeter of an Inscribed 96-gon)**.** Since $p_6 = 6$, we begin with a doubling to find p_{12} :

$$\begin{aligned} p_{12} &= 12\sqrt{2 - \sqrt{4 - \left(\frac{6}{6}\right)^2}} \\ &= 12\sqrt{2 - \sqrt{3}} \end{aligned}$$

Using this, we know $\frac{p_{12}}{12} = \sqrt{2 - \sqrt{3}}$, and we can double again:

$$\begin{aligned} p_{24} &= 24\sqrt{2 - \sqrt{4 - \left(2 - \sqrt{3}\right)}} \\ &= 24\sqrt{2 - \sqrt{2 + \sqrt{3}}} \end{aligned}$$

Now doubling to the 48 gon,

$$\begin{aligned} p_{48} &= 48\sqrt{2 - \sqrt{4 - (2 - \sqrt{2 + \sqrt{3}})}} \\ &= 48\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{3}}}} \end{aligned}$$

One more doubling brings us to the 96-gon,

$$p_{96} = 96\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{3}}}}}$$

Numerically approximating this gives 6.282063901781019276222, which is more recognizable to us if we compute the half perimeter:

$$\frac{p_{96}}{2} \approx 3.141031950890\ldots$$

Exercise 0.9. Find a recurrence relation for the area a_{2n} of the inscribed polygon, in terms of the area a_n of a polygon with half as many sides.

Exercise 0.10. Let $t_n = \tan(180/n)$. Show that t_n satisfies the recurrence relation

$$t_{2n} = \sqrt{1 + \frac{1}{t_n^2}} - \frac{1}{t_n}$$

Hint: you'll need some trig identities to write everything in terms of tangent! Use this to find a recurrence relation for P_n. Can you use this to find the circumference of an octagon circumscribing the unit circle?

After all of this are still left with a fundamental question: what sort of number *is* π? Archimedes' calculation out at infinity showed the area and circumference of a circle were related, but did not give us an exact value for either. These approximate calculations lead to some pretty scary looking numbers, but we know better than to trust that: we've already seen an infinite series of archimedes that summed to a nice rational number, and soon we will meet a nested sequence of square roots that collapses to a single root at infinity:

$$\sqrt{1 + \sqrt{1 + \sqrt{1 + \cdots}}} = \frac{1 + \sqrt{5}}{2}$$

Convergence, Concern and Contradiction

Madhava, Leibniz & $\pi/4$

Madhava was a Indian mathematician who discovered many infinite expressions for trigonometric functions in the 1300's, results which today are known as Taylor Series after Brook Taylor, who worked with them in 1715. In a particularly important example, Madhava found a formula to calculate the arc length along a circle, in terms of the tangent: or phrased more geometrically, the arc of a circle contained in a triangle with base of length 1.

> The first term is the product of the given sine and radius of the desired arc divided by the cosine of the arc. The succeeding terms are obtained by a process of iteration when the first term is repeatedly multiplied by the square of the sine and divided by the square of the cosine. All the terms are then divided by the odd numbers 1, 3, 5, ... The arc is obtained by adding and subtracting respectively the terms of odd rank and those of even rank.

As an equation, this gives

$$\begin{aligned}\theta &= \frac{\sin\theta}{\cos\theta} - \frac{1}{3}\frac{\sin^2\theta}{\cos^2\theta}\left(\frac{\sin\theta}{\cos\theta}\right) + \frac{1}{5}\frac{\sin^2\theta}{\cos^2\theta}\left(\frac{\sin^2\theta}{\cos^2\theta}\frac{\sin\theta}{\cos\theta}\right) + \cdots \\ &= \tan\theta - \frac{\tan^3\theta}{3} + \frac{\tan^5\theta}{5} - \frac{\tan^7\theta}{7} + \frac{\tan^9\theta}{9} - \cdots\end{aligned}$$

If we take the arclength $\pi/4$ (the diagonal of a square), then both the base and height of our triangle are equal to 1, and this series becomes

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

This result was also derived by Leibniz (one of the founders of modern calcuous), using a method close to something you might see in Calculus II these days. It goes as follows: we know (say from the last chapter) the sum of the geometric series

$$\sum_{n\geq 0} r^n = \frac{1}{1-r}$$

Thus, substituting in $r = -x^2$ gives

$$\sum_{n\geq 0}(-1)^n x^{2n} = \frac{1}{1+x^2}$$

and the right hand side of this is the derivative of arctangent! So, anti-differentiating both sides of the equation yields

$$\begin{aligned}\arctan x &= \int \sum_{n\geq 0}(-1)^n x^{2n}\,dx \\ &= \sum_{n\geq 0}\int (-1)^n x^{2n}\,dx \\ &= \sum_{n\geq 0}(-1)^n \frac{x^{2n+1}}{2n+1}\end{aligned}$$

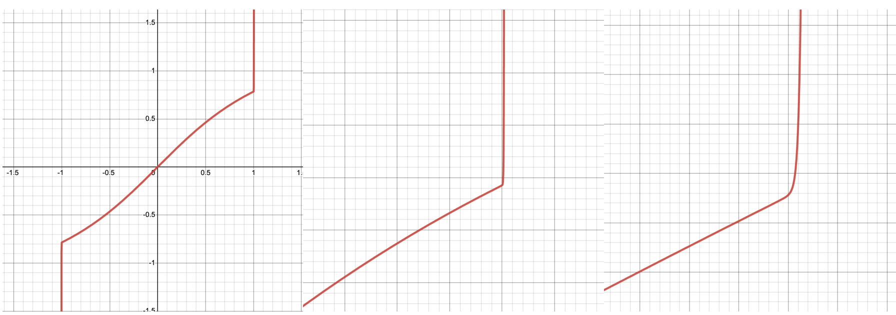
Finaly, we take this result and plug in $x = 1$: since $\arctan(1) = \pi/4$ this gives what we wanted:

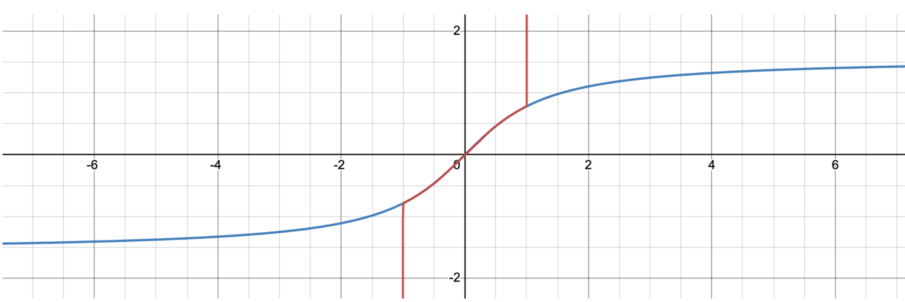
$$\frac{\pi}{4} = \sum_{n\geq 0}(-1)^n \frac{1}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

This argument is *completely full* of steps that should make us worried:

- **Why can we substitute a variable into an infinite expression and ensure it remains valid?**
- **Why is the derivative of arctan a rational function?**
- **Why can we integrate an infinite expression?**
- **Why can we switch the order of taking an infinte sum, and integration?**
- **How do we know which values of x the resulting equation is valid for?**

But beyond all of this, we should be even *more worried* if we try to plot the graphs of the partial sums of this supposed formula for the arctangent.





The infinite series we derived seems to match the arctangent *exactly* for a while, and then abruptly stop, and shoot off to infinity. Where does it stop? *Right at the point we are interested in, $\theta = \pi/4$, so $\tan(\theta) = 1$. So, even a study of which intervals a series converges in will not be enough here, we need a theory that is so precise, it can even tell us exactly what happens at the single point forming the boundary between order and chaos.

And perhaps, before thinking the eventual answer might simply say the series always converges at the endpoints, it turns out at the other endpoint $x = -1$, this series itself *diverges*! So whatever theory we build will have to account for such messy cases.

Dirichlet & $\log 2$

In 1827, Dirichlet was studying the sums of infinitely many terms, thinking about the *alternating harmonic series*

$$\sum_{n\geq 1}\frac{(-1)^n}{n+1}$$

Like the previous example, this series naturally emerges from manipulations in calculus: beginning once more with the geometric series $\sum_{n\geq 0} r^n = \frac{1}{1-r}$. We substitute $r = -x$ to get a series for $1/(1+x)$ and then integrate term by term to produce a series for the logarithm:

$$\log(1+x) = \int \frac{1}{1+x}dx = \int \sum_{n\geq 0}(-1)^n x^n$$

$$= \sum_{n\geq 0}(-1)^n\frac{x^{n+1}}{n+1} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

Finally, plugging in $x = 1$ yields the sum of interest. It turns out not to be difficult to prove that this series does indeed approach a finite value after the addition of infinitely many terms, and a quick check adding up the first thousand terms gives an approximate value of 0.6926474305598, which is very close to log(2) as expected..

$$\log(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10}\cdots$$

What happens if we multiply both sides of this equation by 2?

$$2\log(2) = 2 - 1 + \frac{2}{3} - \frac{1}{2} + \frac{2}{5} - \frac{1}{3} + \frac{2}{7} - \frac{1}{4} + \frac{2}{9} - \frac{1}{5}\cdots$$

We can simplify this expression a bit, by re-ordering the terms to combine similar ones:

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

After simplifying, we've returned to exactly the same series we started with! That is, we've shown $2\log(2) = \log(2)$, and dividing by $\log(2)$ (which is nonzero!) we see that $2 = 1$, a contradiction!

What does this tell us? Well, the only difference between the two equations is *the order in which we add the terms*. And, we get different results! This reveals perhaps the most shocking discovery of all, in our time spent doing dubious computations: **infinite addition is not always commutative, even though finite addition always is**.

Here's an even more dubious-looking example where we can prove that $0 = \log 2$. First, consider the infinite sum of zeroes:

$$0 = 0 + 0 + 0 + 0 + 0 + \cdots$$

Now, rewrite each of the zeroes as $x - x$ for some specially chosen xs:

$$0 = (1 - 1) + \left(\frac{1}{2} - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{3}\right) + \left(\frac{1}{} - \frac{1}{4}\right) + \cdots$$

Now, do some re-arranging to this:

$$\left(1 + \frac{1}{2} - 1\right) + \left(\frac{1}{3} + \frac{1}{4} - \frac{1}{2}\right) + \left(\frac{1}{5} + \frac{1}{6} - \frac{1}{3}\right) + \cdots$$

Make sure to convince yourselves that all the same terms appear here after the rearrangement!

Simplifying this a bit shows a pattern:

$$\left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{5} - \frac{1}{6}\right) + \cdots$$

Which, after removing the parentheses, is the familiar series $\sum \frac{(-1)^n}{n}$. But this series equals $\log(2)$ (or, was it $2 \log 2$?) So, if we are to believe that arithmetic with infinite sums is valid, we reach the contradiction

$$0 = \log 2$$

Infinite Expressions in Trigonometry

The sine function (along with the other trigonometric, exponential, and logarithmic functions) differs from the common functions of early mathematics (polynomials, rational functions and roots) in that it is defined not by a formula but *geometrically*.

Such a definition is difficult to work with if one actually wishes to *compute*: for example, Archimedes after much trouble managed to calculate the exact value of $\sin(\pi/96)$ using a recursive doubling procedure, but he would have failed to calculate $\sin(\pi/97)$ - 97 is not a multiple of a power of 2, so his procedure wouldn't apply! The search for a *general formula* that you could plug numbers into and compute their sine, was foundational to the arithmetization of geometry.

One big question about this procedure is **why in the world should this work**? We found a function that $\sin(x)$ satisfies, and then we plugged *something else into that function* and started iterating: what justification do we have that this should start to approach the sine? We can check after the fact that it (seems to have) worked, but this leaves us far from any understanding of what is actually going on. –>

Infinite Product of Euler

One famous infinite expression for the sine function arose from thinking about the behavior of polynomials, and the relation of their formulas to their roots. As an example consider a quartic polynomial $p(x)$ with roots at $x = a, b, c, d$. Then we can recover p up to a constant multiple as a product of linear factors with roots at a, b, c, d. If the $y-$intercept is $p(0) = k$, we can give a fully explicit description

$$p(x) = k\left(1 - \frac{x}{a}\right)\left(1 - \frac{x}{b}\right)\left(1 - \frac{x}{c}\right)\left(1 - \frac{x}{d}\right)$$

In 17334, Euler attempted to apply this same reasoning in the infinite case to the trigonometric function $\sin(x)$. This has roots at every integer multiple of π, and so following the finite logic, should factor as a product of linear factors, one for each root. There's a slight technical problem in directly applying the above argument, namely that $\sin(x)$ has a root at $x = 0$, so $k = 0$. One work-around is to consider the function $\frac{\sin x}{x}$. This is not actually defined at $x = 0$, but one can prove $\lim_{x \to 0} \frac{\sin x}{x} = 1$, and attempt to use $k = 1$

GRAPH

Its roots agree with that of $\sin(x)$ except there is no longer one at $x = 0$. That is, the roots are $\ldots, -3\pi, -2\pi, -\pi, \pi, 2\pi, 3\pi, \ldots$, and the resulting factorization is

$$\frac{\sin x}{x} = \cdots\left(1 + \frac{x}{3\pi}\right)\left(1 + \frac{x}{2\pi}\right)\left(1 + \frac{x}{\pi}\right)\left(1 - \frac{x}{\pi}\right)\left(1 - \frac{x}{2\pi}\right)\left(1 - \frac{x}{3\pi}\right)\cdots$$

Euler noticed all the factors come in pairs, each of which represented a difference of squares.

$$\left(1 - \frac{x}{n\pi}\right)\left(1 + \frac{x}{2n\pi}\right) = \left(1 - \frac{x^2}{n^2\pi^2}\right)$$

Not worrying about the fact that infinite multiplication may not be commutative (a worry we came to appreciate with Dirichlet, but this was after Euler's time!), we may re-group this product pairing off terms like this, to yield

$$\frac{\sin x}{x} = \left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{2^2\pi^2}\right)\left(1 - \frac{x^2}{3^2\pi^2}\right)\cdots$$

Finally, we may multiply back through by x and get an infinite product expression for the sine function:

Proposition 0.6 (Euler)**.**

$$\sin x = x\left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{4\pi^2}\right)\left(1 - \frac{x^2}{9\pi^2}\right)\cdots$$

This incredible identity is actually correct: there's only one problem - the argument itself is wrong!

Exercise 0.11. In his argument, Euler crucially uses that if we know

- all the zeroes of a function
- the value of that function is 1 at $x = 0$

then we can factor the function as an infinite polynomial in terms of its zeroes. This implies that a function is *completely determined* by its value at $x = 0$ and its zeroes (because after all, once you know that information you can just write down a formula like Euler did!) This is absolutely true for all finite polynomials, but it fails spectacularly in general.

Show that this is a serious flaw in Euler's reasoning by finding a *different* function that has all the same zeroes as $\sin(x)/x$ and is equal to 1 at zero (in the limit)!

Exercise 0.12 (The Wallis Product for π)**.** In 1656 John Wallis derived a remarkably beautiful formula for π (though his argumnet was not very rigorous).

$$\frac{\pi}{2} = \frac{2}{1}\frac{2}{3}\frac{4}{3}\frac{4}{5}\frac{6}{5}\frac{6}{7}\frac{8}{7}\frac{8}{9}\frac{10}{9}\frac{10}{11}\frac{12}{11}\frac{12}{13}\cdots$$

Using Euler's infinite product for $\sin(x)$ evaluated at $x = \pi/2$, give a derivation of Wallis' formula.

The Basel Problem

The Italian mathematician Pietro Mengoli proposed the following problem in 1650:

Definition 0.5 (The Basel Problem)**.** Find the exact value of the infinite sum

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

By directly computing the first several terms of this sum one can get an estimate of the value, for instance adding up the first 1,000 terms we find $1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots\frac{1}{1{,}000^2}=1.6439345\ldots$, and ading the first million terms gives

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{1,000^2}+\cdots+\frac{1}{1,000,000^2}=1.64492406\ldots$$

so we might feel rather confident that the final answer is somewhat close to 1.64. But the interesting math problem isn't to approximate the answer, but rather to figure out something exact, and knowing the first few decimals here isn't of much help.

This problem was attempted by famous mathematicians across Europe over the next 80 years, but all failed. All until a relatively unknown 28 year old Swiss mathematician named Leonhard Euler published a solution in 1734, and immediately shot to fame. (In fact, this problem is named the Basel problem after Euler's hometown.)

Proposition 0.7 (Euler)**.**

$$\sum_{n\geq 1}\frac{1}{n^2}=\frac{\pi^2}{6}$$

Euler's solution begins with two different expressions for the function $\sin(x)/x$, which he gets from the sine's series expansion, and his own work on the infinite product:

$$\begin{aligned}\frac{\sin x}{x}&=1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\frac{x^8}{9!}-\frac{x^{10}}{11!}+\cdots\\&=\left(1-\frac{x^2}{\pi^2}\right)\left(1-\frac{x^2}{2^2\pi^2}\right)\left(1-\frac{x^2}{3^2\pi^2}\right)\cdots\end{aligned}$$

Because two polynomials are the same if and only if the coefficients of all their terms are equal, Euler attempts to generalize this to infinite expressions, and equate the coefficients for sin. The constant coefficient is easy - we can read it off as 1 from both the series and the product, but the quadratic term already holds a deep and surprising truth.

From the series, we can again simply read off the coefficient as $-1/3!$. But from the product, we need to think - after multiplying everything out, what sort of products will lead to a term with x^2? Since each factor is *already quadratic* this is more straightforward than it sounds at first - the only way to get a quadratic term is to take one of the quadratic terms already present in a factor, and multiply it by 1 from another factor! Thus, the quadratic terms are $-\frac{x^2}{2^2\pi^2}-\frac{x^2}{3^2\pi^2}-\frac{x^2}{4^2\pi^2}-\cdots$. Setting the two coefficients equal (and dividing out the negative from each side) yields

$$\frac{1}{3!}=\frac{1}{\pi^2}+\frac{1}{2^2\pi^2}+\frac{1}{3^2\pi^2}+\cdots$$

Which quickly leads to a solution to the original problem, after multiplying by π^2:

$$\frac{\pi^2}{3!}=1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots$$

Euler had done it! There are of course many dubious steps taken along the way in this argument, but calculating the numerical value,

$$\frac{\pi^2}{3!}=1.64493406685\ldots$$

We find it to be exactly the number the series is heading towards. This gave Euler the confidence to publish, and the rest is history.

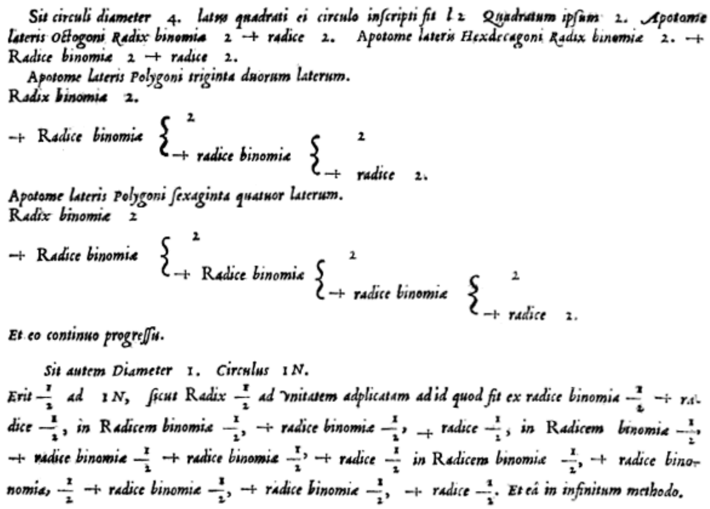
But we analysis students should be looking for potential troubles in this argument. What are some that you see?

Viète's Infinite Trigonometric Identity

Viete was a French mathematician in the mid 1500s, who wrote down for the first time in Europe, an exact expression for π in 1596.

Proposition 0.8 (Viète's formula for π)**.**

$$\frac{2}{\pi}=\frac{\sqrt{2}}{2}\,\frac{\sqrt{2+\sqrt{2}}}{2}\,\frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2}\,\frac{\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{2}}}}}{2}\cdots$$



Sit circuli diameter 4. latus quadrati ei circulo inscripti fit l 2 Quadratum ipsum 2. Apotome lateris Octogoni Radix binomia 2 + radice 2. Apotome lateris Hexdecagoni Radix binomia 2. + Radice binomia 2 + radice 2.

Apotome lateris Polygoni triginta duorum laterum.
Radix binomia 2.
+ Radice binomia { 2 + radice binomia { 2 + radice 2.

Apotome lateris Polygoni sexaginta quatuor laterum.
Radix binomia 2
+ Radice binomia { 2 + Radice binomia { 2 + radice binomia { 2 + radice 2.

Et eo continuo progressu.

Sit autem Diameter 1. Circulus 1 N.
Erit $\frac{1}{2}$ ad 1 N, sicut Radix $\frac{1}{2}$ ad vnitatem adplicatam ad id quod fit ex radice binomia $\frac{1}{2}$ + radice $\frac{1}{2}$, in Radicem binomia $\frac{1}{2}$, + radice binomia $\frac{1}{2}$, + radice $\frac{1}{2}$, in Radicem binomia $\frac{1}{2}$, + radice binomia $\frac{1}{2}$ + radice binomia $\frac{1}{2}$, + radice $\frac{1}{2}$ in Radicem binomia $\frac{1}{2}$, + radice binomia, $\frac{1}{2}$ + radice binomia $\frac{1}{2}$, + radice binomia $\frac{1}{2}$, + radice $\frac{1}{2}$. Et eâ in infinitum methodo.

Figure 11.: Viete's original publicaiton of this formula - it predates our modern notation for square roots!

How could one derive such an incredible looking expression? One approach uses trigonometric identities...an infinite number of times! Start with the familiar function $\sin(x)$. Then we may apply the double angle identity to rewrite this as

$$\sin(x) = 2\sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right)$$

Now we may apply the double angle identity *once again* to the term $\sin(x/2)$ to get

$$\begin{aligned}\sin(x) &= 2\sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right)\\ &= 4\sin\left(\frac{x}{4}\right)\cos\left(\frac{x}{4}\right)\cos\left(\frac{x}{2}\right)\end{aligned}$$

and again

$$\sin(x) = 8\sin\left(\frac{x}{8}\right)\cos\left(\frac{x}{8}\right)\cos\left(\frac{x}{4}\right)\cos\left(\frac{x}{2}\right)$$

and again

$$\sin(x) = 16\sin\left(\frac{x}{16}\right)\cos\left(\frac{x}{16}\right)\cos\left(\frac{x}{8}\right)\cos\left(\frac{x}{4}\right)\cos\left(\frac{x}{2}\right)$$

And so on....after the n^{th} stage of this process one can re-arrange the the above into the following (completely legitimate) identity:

$$\frac{\sin x}{2^n \sin \frac{x}{2^n}} = \cos\frac{x}{2}\cos\frac{x}{4}\cos\frac{x}{8}\cos\frac{x}{16}\cdots\cos\frac{x}{2^n}$$

Viete realized that as n gets really large, the function $2^n\sin(x/2^n)$ starts to look a lot like the function x...and making this replacement in the formula as we let n go to infinity yields

Proposition 0.9 (Viète's Trigonometric Identity)**.**

$$\frac{\sin x}{x} = \cos\frac{x}{2}\cos\frac{x}{4}\cos\frac{x}{8}\cos\frac{x}{16}\cdots$$

An incredible, infinite trigonometric identity! Of course, there's a huge question about its derivation: are we *absolutely sure* we are justified in making the denominator there equal to x? But carrying on without fear, we may attempt to plug in $x = \pi/2$ to both sides, yielding

$$\frac{2}{\pi} = \cos\frac{\pi}{4}\cos\frac{\pi}{8}\cos\frac{\pi}{16}\cos\frac{\pi}{32}\cdots$$

Now, we are left just to simplify the right hand side into something computable, using more trigonometric identities! We know $\cos\pi/4$ is $\frac{\sqrt{2}}{2}$, and we can evaluate the other terms iteratively using the half angle identity:

$$\cos\frac{\pi}{8} = \sqrt{\frac{1+\cos\frac{\pi}{4}}{2}} = \sqrt{\frac{1+\frac{\sqrt{2}}{2}}{2}} = \frac{\sqrt{2+\sqrt{2}}}{2}$$

$$\cos\frac{\pi}{16} = \sqrt{\frac{1+\cos\frac{\pi}{8}}{2}} = \sqrt{\frac{1+\frac{\sqrt{2+\sqrt{2}}}{2}}{2}} = \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2}$$

Substituting these all in gives the original product. And, while this derivation has a rather dubious step in it, the end result seems to be correct! Computing the first ten terms of this product on a computer yields $0.63662077105\ldots$, wheras $2/\pi = 0.636619772$. In fact, Viete used his own formula to compute an approximation of π to nine correct decimal digits. This leaves the obvious question, **Why does this argument work?**

The Infinitesimal Calculus

In trying to formalize many of the above arguments, mathematicians needed to put the calculus steps on a firm footing. And this comes with a whole collection of its own issues. Arguments trying to explain in clear terms what a derivative or integral was really supposed to be often led to nonsensical steps, that cast doubt on the entire procedure. Indeed, the history of calculus is itself so full of confusion that it alone is often taken as the motivation to develop a rigorous study of analysis. Because we have already seen so many other troubles that come from the infinite, we will content ourselves with just one example here: what is a derivative?

The derivative is meant to measure the slope of the tangent line to a function. In words, this is not hard to describe. But like the sine function, this does not provide a means of *computing*, and we are looking for a *formula*. Approximate formulas are not hard to create: if $f(x)$ is our function, and h is some small number the quantity

$$\frac{f(x+h)-f(x)}{h}$$

represents the slope of the secant line to f between x and h. For any finite size in h this is only an approximation, and so thinking of this like Archimedes did his polygons and the circle, we may decide to write down a *sequence of ever better approximations*:

$$D_n = \frac{f\left(x+\frac{1}{n}\right)-f(x)}{\frac{1}{n}}$$

and then define the derivative as the *infiniteth* term in this sequence. But this is just incoherent, taken at face value. If $1/n \to 0$ as $n \to \infty$ this would lead us to

$$\frac{f(x+0)-f(x)}{0} = \frac{0}{0}$$

So, something else must be going on. One way out of this would be if our sequence of approximates *did not actually converge to zero* - maybe there were infinitely small nonzero numbers out there waiting to be discovered. Such hypothetical numbers were called *infinitesimals*.

Definition 0.6 (Infinitesimal)**.** A positive number ϵ is infinitesimal if it is smaller than $1/n$ for all $n \in \mathbb{N}$.

This would resolve the problem as follows: if dx is some infinitesimal number, we could define the derivative as

$$D = \frac{f(x+dx)-f(x)}{dx}$$

But this leads to its own set of difficulties: its easy to see that if ϵ is an infinitesimal, then so is 2ϵ, or $k\epsilon$ for any rational number k.

Exercise 0.13. Prove this: if ϵ is infinitesimal and $k \in \mathbb{Q}$ show $k\epsilon$ is infinitesimal$.

So we can't just say define the derivative by saying "choose some infinitesimal dx" - there are many such infinitesimals and we should be worried about which one we pick! What actually happens if we try this calculation in practice, showcases this.

Let's attempt to differentiate x^2, using some infinitesimal dx. We get

$$\begin{aligned}(x^2)' &= \frac{(x+dx)^2 - x^2}{dx} = \frac{x^2 + 2xdx + dx^2 - x^2}{dx} \\ &= \frac{2xdx + dx^2}{dx} = 2x + dx\end{aligned}$$

Here we see the derivative is not what we expected, but rather is $2x$ plus an infinitesimal! How do we get rid of this? One approach (used very often in the foundational works of calculus) is simply to discard any infinitesimal that remains at the end of a computation. So here, because $2x$ is finite in size and dx is infinitesimal, we would just discard the dx and get $(x^2)' = 2x$ as desired.

But this is not very sensible: when exactly are we allowed to do this? If we can discard an infinitesimal whenever its added to a finite number, shouldn't we already have done so with the $(x + dx)$ that showed up in the numerator? This would have led to

$$\frac{(x+dx)^2 - x^2}{dx} = \frac{x^2 - x^2}{dx} = \frac{0}{dx} = 0$$

So, the when we throw away the infinitesimal matters deeply to the answer we get! This does not seem right. How can we fix this? One approach that was suggested was to say that we *cannot* throw away infinitesimals, but that the square of an infinitesimal is *so small that it is precisely zero*: that way, we keep every infinitesimal but discard any higher powers. A number satisfying this property was called *nilpotent* as *nil* was another word for zero, and *potency* was an old term for powers (x^2 would be the *second potency of x).

Definition 0.7. A number ϵ is nilpotent if $\epsilon \neq 0$ but $\epsilon^2 = 0$.

If our infinitesimals were nilpotent, that would solve the problem we ran into above. Now, the calculation for the derivative of x^2 would proceed as

$$\frac{(x+dx)^2-x^2}{dx}=\frac{x^2+2xdx+dx^2-x^2}{=}\frac{2xdx+0}{dx}=2x$$

But, in trying to justify *just this one calculation* we've had to invent two new types of numbers that had never occurred previously in math: we need positive numbers smaller than any rational, and we also need them (or at least some of those numbers) to square to precisely zero. **Do such numbers exist?**

Part I.

Numbers

- In Chapter 1 we begin axiomatizing the real numbers by axiomatizing their operations of addition and multiplication, leading to the field axioms.
- In Chapter 2 we define the notion of inequality in terms of the notion of *positivity* which we axiomatize, leading to the definition of an *ordered field*.
- In Chapter 3 we look to formalize the notion of limit used by the babylonians and archimedes, and end up with the *Nested Interval Property*. This leads us to introduce new concepts (infima and suprema) and a new axiom: completeness.
- In Chapter 4 we define the real numbers as the (unique) *complete, ordered field* and study its properties.
- In **?@sec-numbers-functions** we look at the modern defintion of real valued functions, and some of the monstrous objects this allows.

1. Operations

Highlights of this Chapter: We begin axiomatizing the real numbers by axiomatizing their operations of addition and multiplication, leading to the field axioms. We give careful definitions of various notations from arithmetic, and do several example calculations (including a proof that $2+2=4$ and $(a+b)^2=a^2+2ab+b^2$) to exhibit that all arithmetical facts are consequences of the field axioms.

The first step to axiomatizing numbers is to give a precise description of addition, subtraction, multiplication and division. These operations naturally group into two pairs (addition/subtraction as well as multiplication/division) of operation/inverse, so first we will formalize the notion of an *invertible operation*. Furthermore, the two operations are related to one another by the *distributive law*. Two invertible operations bonded together by the distributive law form a mathematical structure we call a *field*, which is what we axiomatize in this chapter.

1.1. Binary Operations

Definition 1.1 (Binary Operation)**.** A binary operation $\star$ on a set S is a rule that takes any two elements of S and combines them to make a new element of S.

Formally, this is a function $\star : S\times S \to S$. Whereas we often write functions $f : S\times S \to S$ as $f(a,b)$ for a binary operation we traditionally write the function name in the *middle* so $a\star b$ instead of $\star(a,b)$.

Addition is a binary operation on the natural numbers, integers, rationals, and real numbers. Subtraction is a binary operation on the integers, but not on the natural numbers, as $4-7=-3$ gives an element not in the original set.

Definition 1.2 (Commutativity & Associativity)**.** An operation $\star$ is commutative if the order the elements are combined does not affect the outcome: for all elements $a,b\in S$

$$a\star b=b\star a$$

An operation is *associative* if combinations of 3 or more terms can be re-grouped at will (not changing the order), without affecting the outcome: for all $a,b,c\in S$

$$(a\star b)\star c=a\star(b\star c)$$

The operation of addition is commutative and associative, but the operation of subtraction is neither. The operation of matrix multiplication is associative, but is not commutative in general.

An operation which is commutative but not associative is given by the children's game *rock paper scissors*: if $S = \{r, p, s\}$ we may define the operation $\star$ to select the *winning element* of any pair. Thus, because paper beats rock, we have $r \star p = p$. Explain why this is commutative, and find an example proving it is not associative.

Definition 1.3 (Identities & Inverses)**.** Let S be a set with binary operation $\star$. Then an element $e \in S$ is an *identity* for the operation if it does not change any elements under combination. Formally, for all $s \in S$

$$e \star s = s \star e = s$$

Given a binary operation $\star$ on a set S with identity $e \in S$, an element $x \in S$ is *invertible* if it can be combined with something to produce the identity. That is, if there exists a $y \in S$ with

$$x \star y = y \star x = e$$

This element y is called the *inverse* of x. An operation $\star$ is called *invertible* if every element of S has an inverse.

Zero is the identity of the operation of addition, 1 is the identity of multiplication (in any familiar number system you'd like to take as an example). The identity matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ is the identity of 2×2 matrix multiplication. Not all operations have an identity. Can you see why there is no identity operation for exponentiation x^y on the positive integers?

The operation of addition is invertible, and its inverse is *subtraction*. The operation of multiplication is not invertible, because the number 0 does not have an inverse (you can't divide by zero! We'll prove this soon)

Definition 1.4 (Group)**.** A *group* is a set G with an associative, invertible binary operation $\star$.

1.2. Fields

We've defined what a *nice* binary operation is. Numbers have two of these!

Definition 1.5 (Distributive Law)**.** Let S be a set with two commutative binary operations $+, \cdot$. Then $\cdot$ distributes over $+$ if for all $a, b, c \in S$ we have

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$

Definition 1.6 (Field). A Field is a set $\mathbb{F}$ with two binary operations denoted $+$ (addition) and $\cdot$ (multiplication) satisfying the following axioms.

- (Commutativity) If $a, b \in \mathbb{F}$ then $a + b = b + a$ and $a \cdot b = b \cdot a$.
- (Associativity) If $a, b, c \in \mathbb{F}$ then $(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
- (Identities) There are special elements denoted $0, 1 \in \mathbb{F}$ where for all $a \in \mathbb{F}$, $a + 0 = a$ and $1 \cdot a = a$.
- (Inverses) For every $a \in \mathbb{F}$ there is an element $-a$ such that $a + (-a) = 0$. If $a \neq 0$, then there is also an element a^{-1} such that $a \cdot a^{-1} = 1$.
- (Distributivity) If $a, b, c \in \mathbb{F}$ then $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$

Example 1.1. The rational numbers $\mathbb{Q}$ form a field, but the integers $\mathbb{Z}$ do not, as they do not contain multiplicative inverses.

1.2.1. Shorthand Notation

We will work with fields and their operations throughout the course, so it is useful to introduce some shorthand notation that is familiar to us from previous mathematics classes, and put it on rigorous foundations in terms of the field axioms.

- Since addition and multiplication are associative, we will drop parentheses when three or more terms are combined using the same operation. That is, we will write $a + b + c$ for both $(a + b) + c$ and $a + (b + c)$ when convenient.
- We will adopt the convention that multiplication *takes precedence* over addition; that is, we drop parentheses in $(a \cdot b) + c$ to allow ourselves to write $a \cdot b + c$; but we require parentheses to write $a \cdot (b + c)$.
- We will denote multiplication by simple juxtaposition when convenient, dropping the $\cdot$ symbol. That is, we will write ab for $a \cdot b$ and $a(b + c)$ for $a \cdot (b + c)$.
- We use a bar $\frac{a}{b}$ to denote multiplication by the inverse: that is $a(b^{-1})$.
- We denote repeated multiplication by powers: that is, for positive integers n we write x^n to mean the product of n copies of x.
- For $x \neq 0$ we define the symbol $x^0 = 1$ for convenience, and for negative n we define x^{-n} as $1/x^n$.

We also have a special shorthand for numerals, familiar to all

- The numerals 0 and 1 denote the special elements of any field guaranteed to exist by the axioms.
- We write 2 as a shorthand for $1 + 1$.
- We write 3 as a shorthand for $2 + 1$.
- We write 4 as a shorthand for $3 + 1$.

For large integers, we use the base 10 system unless otherwise specified. That is, we interpret 364 as $3 \cdot 10^2 + 6 \cdot 10^1 + 4 \cdot 10^0$.

Exercise 1.1 (2 + 2 and $2 \cdot 2$)**.** Prove, using only the field axioms and the definitions of the symbols $0, 1, 2, 3, 4$ that $2 + 2 = 4$ and $2 \cdot 2 = 4$.

Show that 2 is the only natural number where $x + x = x \cdot x$.

1.3. Elementary Computations

Example 1.2 (Multiplication by Zero)**.**

$$0x = 0$$

To prove this for an arbitrary $x \in \mathbb{F}$, recall that 0 is the additive identity so for any field element c, we have $0 + c = c$. Thus, when $c = 0$ we have $0 + 0 = 0$. We can use this together with the distributive property to get

$$\begin{aligned} 0x &= (0+0)x \\ &= 0x + 0x \end{aligned}$$

Now, we can take the additive inverse of $0x$ and add it to both sides:

$$0x + (-0x) = 0x + 0x + (-0x)$$

This gives the additive identity 0 by definition on the left side, and cancels one of the factors of $0x$ on the right, yielding

$$0 = 0x + 0$$

Finally we use again that 0 is the additive identity to see $0x + 0 = 0x$, which gives us what we want:

$$0x = 0$$

Example 1.3 (The Zero-Product Property)**.** Let a, b be elements of a field and assume that $ab = 0$. Then either $a = 0$ or $b = 0$.

We assume that both a and b are nonzero, and see that we reach a contradiction. Since they're nonzero, they have multiplicative inverses a^{-1} and b^{-1}, so we may multiply both sides of $ab = 0$ by these to get

$$b^{-1}a^{-1}ab = b^{-1}a^{-1}0$$

On the left this simplifies to $b^{-1}1b = b^{-1}b = 1$ by definition, and on the right this becomes $0(b^{-1}a^{-1}) = 0$ by the previous example. Thus, we've proven $0 = 1$! So this could not have been the case, and either a or b must have not been invertible to start with - they must have been zero.

Example 1.4 (Additive Inverses and Negatives).

$$-x = (-1)x$$

The definition of the symbol $-x$ is the element of $\mathbb{F}$ which, when added to x, gives 0. Thus, to prove that $-x = -1x$ we want to prove that if you add $(-1)x$ to x, you get 0. Since 1 is the additive identity, we know $1x = x$ so we may write

$$x + (-1x) = 1x + (-1x)$$

Using the fact that multiplication is commutative and the distributive law, we may factor out the x:

$$1x + (-1)x = (1 + (-1))x$$

Now, by definition $1 + (-1)$ is the additive identity 0, so this is just equal to $0x$. But by Example 1.2 we know $0x = 0$! Thus

$$x + (-1x) = 0$$

And so $-1x$ is the additive inverse of x as claimed. Thus we may write $-x = (-1)x$

Example 1.5 (Negative times a Negative).

$$(-1)(-1) = 1$$

This is an immediate corollary of the above: we know that $(-1)x$ is the additive inverse of x, and so $(-1)(-1)$ is the additive inverse of -1. But this is just 1 itself, by definition!

Exercise 1.2 (Negative of a Negative). For any $x \in \mathbb{F}$ we have

$$-(-x) = x$$

All of the standard arithmetic "rules" learned in grade school are consequences of the field axioms, and so you are welcome to use all of them in this course, without comment. To feel justified in doing this, its good to prove a couple of them yourself, to convince yourself that you could in fact trace and any all such manipulations back to the rigorous axioms we laid down.

Exercise 1.3 (The difference of squares). Prove that for any $a, b \in \mathbb{F}$

$$(a + b)(a - b) = a^2 - b^2$$

In your proof you may use the field axioms, the notational shorthands, and any of the example properties proved above in the notes. Anything else you need, you should prove from this.

Exercise 1.4. Prove, using the field axioms and our notational shorthands, for any a, b and $c \neq 0$

$$\frac{a+b}{c} = \frac{a}{c} + \frac{b}{c}$$

Exercise 1.5. Prove that fraction addition works by finding a common denominator: for any a, c and nonzero b, d

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$$

In your proof you may use the field axioms, the notational shorthands, and any of the example properties proved above in the notes. Anything else you need, you should prove from this.

Exercise 1.6. Fix some number $r \neq 1$ in a field, and prove by induction that

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

2. Order

Highlights of this Chapter: We define the notion of inequality in terms of the notion of *positivity* which we axiomatize, leading to the definition of an *ordered field*. We prove this new axiom is required as not all fields can be ordered (by looking at the complex numbers), and then we investigate several important properties and definitions related to order that are essential to real analysis:

- We define absolute value, and give several characterizations
- We prove the triangle inequality
- We define square roots, and n^{th} roots

2.1. Defining Inequality

How are we supposed to make sense of $a < b$? One approach is to start by thinking about a simpler case: can formalize the idea $a > 0$? We will give *axioms* for how the set of positive numbers should behave:

Definition 2.1 (Positivity)**.** A subset $P \subset \mathbb{F}$ is called the *positive elements* if

- (Trichotomy) For every $a \in \mathbb{F}$ exactly one of the following is true: $a = 0$, $a \in P$ $-a \in P$.
- (Closure) If $a, b \in P$ then $a + b \in P$ and $ab \in P$.

Given these, we can *define* inequality in terms of positivity!

Definition 2.2 (Inequality)**.** Let F be an ordered field and P a set of positive elements for F. If $a, b \in \mathbb{F}$, we write $a < b$ as a shorthand for the statement that $b - a \in P$, and we write $a \leq b$ if either $a < b$ or $a = 0$.

Analogously, we write $a > b$ if $a - b \in P$ and $a \geq b$ if either $a > b$ or $a = b$.

Definition 2.3 (Ordered Field)**.** An ordered field is a field F together with a fixed choice of positive elements P (which then gives a precise definition of inequality).

2.1.1. Properties of Ordered Fields

Proposition 2.1 (1 is a Positive Number)**.** *If (F, P) is any ordered field, then $1 \in P$.*

Proof. Since $1 \neq 0$ we know that either $1 \in P$ or $-1 \in P$. So, to show $1 \in P$ its enough to see $-1 \in P$ leads to contradiction.

If $-1 \in P$ then by closure, $(-1)(-1) = 1 \in P$: so now we have *both* 1 and -1 in P, contradicting trichotomy. □

Exercise 2.1 (Squares are Positive)**.** Let F be an ordered field and $x \neq 0$ an element. Then $x^2 > 0$.

Proposition 2.2 ($\mathbb{C}$ is not ordered)**.** *The complex numbers cannot be made into an ordered field: there is no subset $P \subset \mathbb{C}$ such that P is a positive cone for $\mathbb{C}$.*

Proof. The complex numbers contain an element i with the property that $i^2 = -1$. If they were ordered, since $i \neq 0$ we know either $i \in P$ or $-i \in P$, but both of these lead to contradiction.

If $i \in P$ then $i^2 = -1 \in P$ contradicting the previous theorem that $1 \in P$ always. And $-i \in P$ leads to the same problem: $(-i)^2 = (-i)(-i) = --(i^2) = i^2 = -1$, so $-1 \in P$ again. □

This may seem like a strange example to start with, as the course is about *real* analysis. But its actually quite important: every time we introduce a new concept to the foundations of our theory we should ask ourselves, *is this an axiom, or a theorem*? We don't want to add as axioms things that we can already prove from the existing axioms, as that is redundant! So before adding a new axiom, we should convince ourself that its *necessary*: that it is impossible to prove the existence of this new structure given the previous. And that's what this example does. By exhibiting something that satisfies all the field axioms but *cannot be ordered*, we see that it is logically impossible to prove the existence of an order from the field axioms alone, and thus we must take Definition 2.3 as a new axiom.

Theorem 2.1 (The Rationals are an Ordered Field)**.**

In fact the rationals are *uniquely ordered*: we know that $1 > 0$ and this, together with the behavior of inequality, determines exactly when one rational number is greater than any other.

Exercise 2.2. Prove that $\frac{a}{b} < \frac{c}{d}$ if and only if $ad < bc$.

2.1.2. Definitions Requiring an Order

Definition 2.4 (Intervals)**.** Let F be an ordered field. We write $[a, b]$ for the set $\{x \mid a \leq x \leq b\}$, and call this set a *closed interval* in F. Similarly we write (a, b) for the set $\{x \mid a < x < b\}$, which we call an *open interval.* Mixed intervals are also possible, such as $[a, b) = \{x \mid a \leq x < b\}$.

An *unbounded interval*, or a *ray* is a set of the form $\{x \mid x > a\}$ or $\{x \mid x \geq a\}$. We call the first an *open ray* and the latter a *closed ray*, and often denote them (a, ∞) or $[a\infty)$ as a shorthand. Similarly with $(-\infty, a)$ and $(-\infty, a]$.

Definition 2.5 (Absolute Value)**.** Let $\mathbb{F}$ be an ordered field. Then the *absolute value* is a function $|\cdot| \colon \mathbb{F} \to \mathbb{F}$ defined by

$$|x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}$$

Definition 2.6 (The $\sqrt{\cdot}$ symbol)**.** Let $\mathbb{F}$ be an ordered field, and $x \in \mathbb{F}$. If there exists a $y \geq 0$ in $\mathbb{F}$ such that $y^2 = x$, we call y the *square root of* x and denote $\sqrt{x}$.

We generalize this by defining $\sqrt[q]{x}$ to be the number y with $y^p = x$, when such a number exists.

Exercise 2.3 (No Square Roots of Negatives)**.** Let F be any ordered field, and let $x < 0$. Prove that x does not have a square root in F.

Definition 2.7 (Rational Powers)**.** Let $a \in \mathbb{F}$ and $p/q \in \mathbb{Q}$.
Then if the element $a^p \in \mathbb{F}$ has a q^{th} root, we define the fractional power $a^{p/q}$ as

$$a^{p/q} = \sqrt[q]{a^p}$$

2.2. Working with Inequalities

All the standard properties of inequalities from arithmetic hold in an ordered field, and so you will be able to use them without comment throughout the course. However, its good to derive a few of these for yourselves from the definitions at first, to see how it goes.

Example 2.1 (Inequality is antisymmetric)**.** By trichotomy we see that for every $x \neq y$ we have either $x < y$ or $y < x$ (as, $x - y \neq 0$ implies either $x - y \in P$, so $x - y > 0$ and $x > y$ or the reverse).

Proposition 2.3 (Inequality is transitive)**.** *Let F be an ordered field and a, b, c in F. If $a < b$ and $b < c$, then $a < c$.*

Proof. If $a < b$ then $b - a \in P$. Similarly, $b < c$ implies $c - b \in P$. Closure then tells us their sum, $(c - b) + (b - a) \in P$, and so after simplifying,

$$c + (-b + b) - a = c + 0 - a = c - a \in P$$

This is the definition of $c > a$. □

Exercise 2.4 (Adding to an Inequality). Let F be an ordered field and $a, b, c \in F$ with $a < b$. Then

$$a + c < b + c$$

Proposition 2.4 (Multiplying an Inequality). *Let F be an ordered field and $a, b, c \in F$ with $a < b$. Then if $c > 0$, it follows that $ca < cb$, and if $c < 0$ we have instead $ca > cb$.*

Proof. First treat the case $c > 0$. Since $a < b$ we know $b-a \in P$, and $c \in P$ so $c(b-a) \in P$ by the closure axiom. Distributing gives $cb - ca \in P$ which is the definition of $cb > ca$.

Now, if $c < 0$, we know $c \notin P$, so $-c \in P$. Closure then gives $(-c)(b - a) \in P$, and simplifying yields $-cb + ca \in P$ or $ca - cb \in P$, the definition of $ca > cb$. □

2.2.1. Powers and Roots

Some basic inequalities for powers and roots that will prove useful: like other basic properties of inequalities, you do not need to prove or cite these when you use them in this course, but it is good to have a reference seeing *why* they are true from our axioms.

Example 2.2 ($x \mapsto x^2$ is increasing). If F is an ordered field and $a, b \in F$ are elements with $0 < a < b$ then $a^2 < b^2$.

To prove this, we use both Proposition 2.3 and Proposition 2.4. Since $a < b$ and $a > 0$ we see $a^2 < ab$. But since $a < b$ and $b > 0$, we see $ab < b^2$. Putting these together yields $a^2 < ab < b^2$, so $a^2 < b^2$.

Its necessary to assume a, b are positive in the theorem above: for example $-3 < 1$ but $(-3)^2 = 9$ is not less than $1^2 = 1$. In fact this proof works in reverse as well (check this!) to provide the following useful fact:

Proposition 2.5. *If $a, b \in \mathbb{F}$ are positive elements of an ordered field, then*

$$a < b \iff a^2 < b^2$$

This generalizes to arbitrary powers:

Exercise 2.5 ($x \mapsto x^n$ is increasing)**.** Prove that if F is an ordered field containing positive elements a, b, then for all $n \in \mathbb{N}$, $a < b$ if and only if $a^n < b^n$.

In fact, when n is odd, you may wish to prove that you can remove the assumption that $a, b > 0$.

Here's a quick fact about inequalities that will prove useful to us later on in the course:

Exercise 2.6 (Bernoulli's inequality)**.** Let $\mathbb{F}$ be an ordered field and $x > 0$ be a positive element. Prove by induction that for all natural numbers $\mathbb{N}$

$$(1+x)^n \geq 1 + nx$$

Exercise 2.7 ($\sqrt{\cdot}$ is increasing)**.** Prove that if $0 < x < y$ in an ordered field F, and F contains the square roots $\sqrt{x}, \sqrt{y}$, then $\sqrt{x} < \sqrt{y}$.

Proposition 2.6. *If $r \in \mathbb{Q}$, $r > 0$ is a positive rational number and $x, y \in F$ are positive field elements*

$$x < y \implies x^r < y^r$$

Proof. Use that $x^r = x^{p/q} = (\sqrt[q]{x^p})$ to break this into two problems: first $x < y$ implies $x^p < y^p$. Now, if $u = x^p$ and $v = y^p$ we have $u < v \implies \sqrt[q]{u} < \sqrt[q]{v}$, completing the proof. ☐

2.3. Working with Absolute Values

Proposition 2.7 (Absolute Values and Maxima)**.** *For all x in an ordered field,*

$$|x| = \max\{x, -x\}$$

Corollary 2.1. *If x, a are in an ordered field, the conditions $-x < a$ and $x < a$ are equivalent to*

$$|x| < a$$

Proof. If $-x < a$ and $x < a$ then $\max\{x, -x\} < a$, so by Proposition 2.7, $|x| < a$. Conversely, if $|x| < a$ then $\max\{x, -x\} < a$ so both $x < a$ and $-x < a$. ☐

Corollary 2.2 (Defining Feature of the Absolute Value)**.** *Let F be an ordered field: then $|x| < a$ if and only if $-a < x < a$.*

Proof. By the above $|x| < a$ means $x < a$ and $-x < a$. Multiplying the second inequality by -1 yields $x > -a$, and stringing them together results in $-a < x < a$. □

Finally, we can get a *formula* for the absolute value in terms of squaring and roots.

Example 2.3. For all x in an ordered field $|x| = \sqrt{x^2}$.

Example 2.4 (Multiplication and the Absolute Value)**.**

$$|xy| = |x||y|$$

$$\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$$

The interaction of the absolute value with *addition* is more subtle, but crucial. One of the most important inequalities in all of analysis is the *triangle inequality* of the absolute value:

Proposition 2.8 (The Triangle Inequality)**.** *For any x, y in an ordered field*

$$|x + y| \leq |x| + |y|$$

Proof. It suffices to prove that we have both

$$x + y \leq |x| + |y| \qquad -(x + y) \leq |x| + |y|$$

For the first, note that as $x \leq |x|$ and $y \leq |y|$,

$$x + y \leq |x| + y \leq |x| + |y|$$

Similar reasoning succeeds for the second as $-x \leq |x|$ and $-y \leq |y|$:

$$-x - y \leq |x| + (-y) \leq |x| + |y|$$

□

Exercise 2.8. Let $a_1 + a_2 + \cdots + a_n$ be any finite sum. Prove that

$$\left|\sum_{i=1}^{n} a_i\right| \leq \sum_{i=1}^{n} |a_i|$$

The reverse triangle inequality is another very useful property of absolute values, logically equivalent to the usual triangle inequality, but giving a lower bound for $|a - b|$ instead of an upper bound for $|a + b|$.

Exercise 2.9 (Reverse Triangle Inequality)**.** Prove that for all a, b in an ordered field F

$$||a| - |b|| \leq |a - b|$$

Finally, two corollaries of the triangle inequality and its reverse, by replacing y with $-y$.

Corollary 2.3 (Corollaries of the Triangle Inequality)**.** *For all x, y in an ordered field,*

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

$$|x + y| \geq ||x| - |y||$$

2.4. Further Topics

2.4.1. ⋆ Topology

A final familiar property that arises from ordering a field is the notion of *open sets* and *closed sets*. This in turn is the foundations of the subject of *topology* or the abstract study of shape, which becomes quite important in advanced applications of analysis.

We will not require any deep theory in this course, and stop pause briefly to give a definition of openness and closedness.

Definition 2.8 (Open Set)**.** A set of the form $(a, b) = \{x \mid a < x < b\}$ is called an *open interval*. A set $U \subset \mathbb{F}$ is called open, if for every point $u \in U$ there is some open interval I containing u which is fully contained in U:

$$u \in I \subset U$$

One notable property of this definition: the empty set $\emptyset = \{\}$ is open, as this condition is *vacuously true*: there are no points of $\emptyset$ so this condition doesn't pose any restriction!

Exercise 2.10. Explain why the set $U = \{x \mid x > 0 \text{ and } x \neq 2\}$ is an open set.

Exercise 2.11. Let $\{U_n\}$ be *any* collection of open sets. Prove that the union $\bigcup_n U_n$ is also open.

Hint: his collection doesn't have to be finite, so induction won't help us here. Can you supply a direct proof, using the definition of union and open?

Definition 2.9 (Closed Set)**.** A set is $K \subset \mathbb{F}$ is *closed*, if its complement is an open set.

Exercise 2.12. Show that intervals of the form $[a, b] = \{x \mid a \leq x \leq b\}$ are closed sets. This is why we call them *closed intervals* in calculus courses.

This terminology is rather unfortunate when first learning the subject, as while open and closed are antonyms in english, they are not in mathematics! Being open is a special property that most sets do not have, and so being closed (which is defined relative to an open set) is also a special property. Most sets are neither open nor closed!

Example 2.5 (A set that is neither open nor closed)**.** The set $S = [1, 2)$ is neither open nor closed. Its not open because the point $1 \in S$, but there is no open interval containing 1 which is fully contained in S (every open interval containing 1 contains numbers *smaller than 1* as well).

To see its not closed, we need to show that its complement is not open. Its complement is the set

$$S^c = \{x \mid x < 1\} \cup \{x \mid x \geq 2\}$$

Here we have the same problem at the number two: $2 \in S^c$ but there is no open interval containing 2 which is fully inside S^c, as any such interval would contain points less than 2, and these are not in S^c.

Thus, $[1, 2)$ is neither open nor closed.

But perhaps even stranger, not only can sets be *neither* open nor closed, but they can also be *both* open and closed! Such sets are called *clopen*.

Example 2.6 (A set that is both open and closed)**.** If $\mathbb{F}$ is the entire ordered field, $\mathbb{F}$ is both open and closed.

To see it is open, note for any x we can form the interval $(x - 1, x + 1)$ and this lies inside of $\mathbb{F}$. To see its closed, note that its complement is the empty set and this is vacuously open as commented above.

3. Completeness

Highlights of this Chapter: We look to formalize the notion of limit used by the babylonians and archimedes, and come to the Nested Interval Property. We see that this property does not hold in $\mathbb{Q}$, so we must seek another axiom which implies us. This leads us to bounds, infima, and suprema. We study the properties of this new definition, use it to define completeness, and show completeness does indeed imply the nested interval property, as we wished.

Now that we have axiomatized the notion of a 'number line' as an ordered field, it's time to try and figure out how to describe "completed" infinite processes in a formal way. This is an inherently slippery notion, as it runs into the difficulty of "talking about infinity, without saying infinity" that lies at the heart of analysis.

So, before introducing the abstract tools that end up best suited for this task (the infimum and supremum), we'll begin with some motivational exploration, and think about what sort of theorems we would *want to be true* in a number system that allows one to do infinite constructions.

3.1. Dreaming of Infinity

Archimedes idea for calculating π was to give an upper bound and a lower bound for the area of a circle, in terms of the area a_n of an inscribed polygon and a circumscribed polygon A_n. This provided an interval that archimedes hoped to trap π inside of, each time n grows, a_n grows and A_n shrinks - so the *confidence interval* of Archimedes shrinks!

$$\cdots [a_6, A_6] \supset [a_{12}, A_{12}] \supset [a_{24}, A_{24}] \supset \cdots$$

A collection of intervals like this is called nested:

Definition 3.1 (Nested Intervals). A sequence of intervals $I_1, I_2, I_3, \ldots$ in an ordered field is called *nested* if for all n, $I_{n+1} \subseteq I_n$.

As these nested intervals shrink in size, the hope is that they zero in on π exactly: mathematically we might express this with an *intersection* over all intervals (where the question mark over the equals means we have not proven this, but hope its true)

$$\bigcap_n [a_n, A_n] \stackrel{?}{=} \{\pi\}$$

The babylonian process approximating $\sqrt{2}$ can also be recast in terms of a sequence of nested intervals: where we take the two sides w_n, h_n (width and height) of each approximating rectangle as a confidence interval around $\sqrt{2}$. We of course want, that in the limit this zeroes in directly on the square root,

$$\bigcap_n [w_n, h_n] \stackrel{?}{=} \{\sqrt{2}\}$$

In formulating any of these processes (pre-rigorously, say, in antiquity) mathematicians always assumed without proof that if you had a collection of shrinking intervals, they were shrinking around *some number* that could be captured after infinitely many steps. We capture this unstated assumption rigorously below, and title it the *Approximation Property* based on its use approximating numbers by intervals:

Definition 3.2 (Approximation Property)**.** A number system has the approximation property if the intersection of any sequence of nested intervals whose lengths go to zero contains a single element.

How do we tell if our current axioms imply that our number system has the approximation property? In a situation like this, mathematicians may try to ask *what sort of things satisfy the current axioms* and look at these for inspiration. Here - the rational numbers satisfy the axioms of an ordered field, and this provides a big hint: Pythagoras proved that there is no rational square root of 2, which implies the Babylonian process does not zero in on any number at all, but rather at infinity reaches nothing!

$$\bigcap_n [w_n, h_n] = \emptyset$$

Because there is at least one ordered field (the rationals) that does not satisfy the dream theorem, we know that these axioms are not enough.

Theorem 3.1. *The axioms of an ordered field are not enough to deal with completed infinity: there are ordered fields in which* do not have *the approximation property.*

This tells us we must look to *extend our axiom system* and search out a new axiom that will help our number system capture the slippery notion of infinite processes. One might be tempted to just take the approximation property itself as an axiom(!); but this comes with its own challenges. The property is rather specific (about certain collections nested intervals), whereas we want axioms to be as general and simple-to-state as possible, and worse, it contains a currently undefined term *lengths tending to zero* which we would have to first make rigorous.

Happily, it turns out a productive approach to this grows naturally out of our discussion of nested intervals. But, to decrease the complexity instead of focusing on the entire interval $[\ell_n, u_n]$, we will look separately at the sequence of *lower bounds* ℓ_n and *upper bounds* u_n. Understanding the behavior of either of these will turn out to be enough to extend our axiom system appropriately.

3.2. Suprema and Infima

A confidence interval like $[\mathrm{width}_n, \mathrm{height}_n]$ or $[\mathrm{inscribed}_n, \mathrm{circumscribed}_n]$ gives us for each n both an *upper bound* for the number we are after, and a *lower bound.*
It will be useful to describe these concepts more precisely.

Definition 3.3 (Bounds)**.** Let S be a nonempty subset of an ordered field. An *upper bound* for S is an element $u \in \mathbb{F}$ greater than or equal to all the elements of S:

$$\forall s \in S\, s \leq u$$

A *lower bound* for S is an element $\ell \in \mathbb{F}$ which is less than or equal to all the elements of S:

$$\forall s \in S\, \ell \leq u$$

S is said to be *bounded above* if there exists an upper bound, and to be bounded below if there exists a lower bound. If S is both bounded above and below, then S is said to be *bounded.*

Definition 3.4 (Maximum & Minimum)**.** Let S be a nonempty subset of an ordered field. Then S has a $maximum* if there is an element of $M \in S$ that is also an upper bound for S, and a *minimum* if some element m is also a lower bound for S.

The maximum and minimum elements of a set are the *best possible* upper and lower bounds when they exist: after all, you couldn't hope to find a *smaller* lower bound than the maximum, as the maximum would be greater than it, so it couldn't be an upper bound! While maxima and minima always exist for *finite sets* things get trickier with infinity. For example, the open interval $(0, 1)$ of rational numbers does not have any maximum element.

The correct generalization of *maximum* to cases like this is called the *supremum*: the best possible upper bound.

Definition 3.5 (Supremum)**.** Let S be a set which is bounded above. The *least upper bound* for S is a number σ such that

- σ is an upper bound for S
- If u is any upper bound, then $\sigma \leq u$.

When such a least upper bound exists, we call it the *supremum of* S and denote it $\sigma = \sup S$.

This notion of best possible upper bound allows us to rigorously capture the notion of *endpoint* even for infinite sets that do not have a maximum.

Example 3.1 (A set with no maximum)**.** The set $(0, 1) = \{x \in \mathbb{Q} \mid 0 < x < 1\}$ has no maximal element, but it does have a supremum in $\mathbb{Q}$, namely $1 = \sup S$.

Definition 3.6 (Infimum)**.** The infimum of a set S is the *least upper bound*: that is, an element λ where

- λ is a lower bound for S.
- If ℓ is any other lower bound for S, then $\ell \leq \lambda$.

If such an element exists it is denoted $\lambda = \inf S$.

Example 3.2.

- The set $\mathbb{N}$ has no upper bounds at all, so $\sup \mathbb{N}$ does not exist. It has many lower bounds (like 0, and -14), and its infimum is $\inf \mathbb{N} = 1$.
- The rational numbers themselves have no upper nor lower bound, so $\sup \mathbb{Q}$ and $\inf \mathbb{Q}$ do not exist.

3.3. Completeness

Because infima and suprema are such a useful tool to precisely describe the final state of certain infinite processes, they are a natural choice of object to concentrate on when looking for an additional axiom for our number system. Indeed - after some thought you can convince yourself that the statement *every infinite process that should end in some number, does end in some number* is equivalent to the following definition of *completeness.*

Definition 3.7 (Completeness)**.** An ordered set is *complete* if every nonempty subset S that is bounded above has a supremum.

Remark 3.1. One question you might ask yourself is why we chose *supremum* here, and not *infimum* - or better, why not both?! It turns out that all of these options are *logically equivalent*, as you can prove in some exercises below. So, any one of them suffices

We can formalize Pythagoras' observation about the irrationality of $\sqrt{2}$ in this language

Theorem 3.2 ($\mathbb{Q}$ is not complete)**.** *The set $S = \{s \in \mathbb{Q} \mid s^2 < 2\}$ does not have a supremum in $\mathbb{Q}$.*

Sketch. A rigorous proof can be given by contradiction: assume that a supremum $\sigma = \sup S$ exists, and then show that we must have $\sigma^2 = 2$ by ruling out the possibilities $\sigma^2 < 2$ and $\sigma^2 > 2$. The calculations required for these steps are more relevant to the next chapter, so we postpone until then (specifically, Example 4.1 and Exercise 4.2).

Once its known that the supremum must satisfy $\sigma^2 = 2$, we apply Pythagoras' observation (Theorem 0.1) that there are no rational solutions to this equation, to reach a contradiction. □

Thus, asking a field to be complete is a constraint above and beyond being an ordered field. So, this is a good candidate for an additional axiom! But before we too hastily accept it, we should check that it actually solves our problem:

Theorem 3.3 (Nested Interval Property)**.** *Let $\mathbb{F}$ be an ordered field* ***which is also complete****, and $I_0, I_1, I_2, \ldots, I_n, \ldots$ be a collection of nested closed intervals. Then their intersection is nonempty:*

$$\bigcap_{n\geq 0} I_n \neq \emptyset$$

Proof. Let $I_n = [a_n, b_n]$. We need to use the fact that $\mathbb{F}$ is complete to help us find a number which lies in I_n for every n. One idea - consider the set of lower endpoints

$$A = \{a_1, a_2, a_3, \ldots, a_n, \ldots\}$$

This set is nonempty, and because the intervals are nested any one of the b_n's serves as an upper bound for A.

By completeness the supremum must exist: lets call this $\alpha = \sup A$. Now we just need to see that $\alpha \in I_n = [a_n, b_n]$ for all n. Fix some n: then as $a_n \in A$ and α is an upper bound, we know that $a_n \leq \alpha$. But b_n is an upper bound for A so the *least upper bound* must satisfy $\alpha \leq b_n$. Putting these together

$$a_n \leq \alpha \leq b_n \implies \alpha \in I_n$$

And, since this holds for all natural numbers n, we actually have $\alpha \in \bigcap_n I_n$, so the intersection is nonempty. □

Confirm that the property the ancients *assumed* held of the number line is now a *theorem* of our formal system!

Exercise 3.1 (The Approximation Property). Let $I_n = [a_n, b_n]$ be a nested sequence of intervals and assume $\sup\{a_n\} = \inf\{b_n\}$. Then show $\bigcap_n I_n$ contains exactly one point.

3.4. Working with inf and sup

Proposition 3.1 (Uniqueness of Supremum). *If the supremum of a set exists, it is unique.*

Proof. Let A be a set. To show uniqueness, we will assume that there are two numbers x and y which both satisfy the definition of the supremum of A, and then we will show $x = y$. Thus, any two possibilities for the supremum are equal, so if theres a supremum at all there can only be one.

To prove $x = y$, we will prove $x \leq y$ and $y \leq x$. Once we have these two, we can immediately conclude that since we can't simultaneously have $x < y$ and $y < x$ (what axiom of an ordered field would this violate?) we must have $x = y$.

If x and y both are least upper bounds for A, then they are both in particular upper bounds. So, x is an upper bound and y is a *least upper bound* implies $y \leq x$. But similarly, y being an upper bound while x is a *least upper bound* implies $x \leq y$. Thus $x = y$ and so the supremum is unique. □

This uses two important proof techniques in analysis.
First, one way to show that something is *unique* is to show that if you had two of them, they have to be equal. Second, to show $x = y$ it is often useful to show both $x \leq y$ and $y \leq x$.

Exercise 3.2. Prove the infimum of a set is unique when it exists.

Proposition 3.2. *Let A be a set which is bounded above. An upper bound α for A is actually the supremum if for every positive $\epsilon > 0$, there exists some element of A greater than $\alpha - \epsilon$.*

Proof. Let's prove the *contrapositive*, meaning we assume the *conclusion* is false and prove the *premise* is false. The conclusion would be false if there were *some* positive ϵ where no element of a is larger than $\alpha - \epsilon$. But this means that $\alpha - \epsilon \geq a$ for all $a \in A$, or that $\alpha - \epsilon$ is an upper bound for A. Since this is less than α (remember, ϵ is positive), we found a smaller upper bound, so α cannot be the least upper bound: thus its *false* that $\alpha = \sup A$.

Since anytime our proposed condition doesn't hold, α isnt the supremum, this means if α were the supremum, the condition must hold! And this is what we sought to prove. $\square$

Remark 3.2. The contrapositive is a very useful proof style, especially in situations where the premise is something short, and the conclusion is something complicated. By taking a look at the contrapositive, you get to assume the negation of the conclusion, meaning you get to assume the complicated thing, and then use it to prove the simple thing (the negation of the premise)

Exercise 3.3. Prove the corresponding characterization of infima: a lower bound ℓ for a set A is the infimum if for every positive $\epsilon > 0$ there is some element of A less than $\ell + \epsilon$.

Exercise 3.4. Let A, B be nonempty bounded subsets of a complete field, and suppose $A \subset B$. Prove that $\sup A \leq \sup B$.

Example 3.3. Let A be a bounded set with supremum $\sup A$ and c an element of the field. Define the set $S = \{a + c \mid a \in A\}$. Then

$$\sup S = c + \sup A$$

To prove this, we need to show two things: (1) that $c + \sup A$ is an upper bound for S, and (2) that its in fact the least upper bound.

First, we consider (1). Since $\sup A$ is an upper bound for A, we know $\forall a \in A, a \leq \sup A$. Adding c to both sides, we also have $c + a \leq c + \sup A$ for all a, which implies $c + \sup A$ is an upper bound.

Now, (2). Let u be *any upper bound* for S. This means that $u \geq c + a$ for all $a \in A$, so subtracting c from both sides, that $u - c \geq a$. Thus, $u - c$ is an upper bound for A, and this is real progress because we know $\sup A$ is the *least upper bound.* That implies $\sup A \leq u - c$ and so adding c to both sides, $c + \sup A \leq u$. Putting this all together, we assumed u was *any upper bound* and we proved $c + \sup A$ was a smaller one.

Thus, $c + \sup A$ is the least upper bound to S, and so by definition we have $\sup S = c + \sup A$ as required.

Exercise 3.5. Let $c > 0$ and A be a bounded set with supremum $\sup A$. Define the set $S = \{ca \mid a \in A\}$. Then $\sup S$ exists and

$$\sup S = c \sup A$$

Exercise 3.6. Let A, B be two bounded nonempty sets. Assuming that the suprema and infima of A and B both exist, prove they do for $A \cup B$ as well and

$$\sup A \cup B = \max\{\sup A, \sup B\}$$
$$\inf A \cup B = \min\{\inf A, \inf B\}$$

Exercise 3.7 (Sup and Inf of Intervals)**.** Let A, B be two open intervals in $\mathbb{R}$, and assume that $\sup A = \inf B$.

True or false: it is possible to add a single point to $A \cup B$ so the entire set is an interval. (Explain your reasoning, but you don't have to write a rigorous proof).

3.5. Problems

Exercise 3.8. Let A, B be subsets of a complete ordered field with $\sup A < \sup B$.

- Prove that there is an element $b \in B$ which is an upper bound for A.
- Give an example to show this is not necessarily true if we only assume $\sup A \leq \sup B$.

Exercise 3.9. Consider the following subsets of the rational numbers. State whether or not they have infima or suprema; when they do, give the inf and sup.

- $[1, 3]$
- $[1, 3)$
- $\{x \mid x^2 < 1\}$
- $\{x \mid x^3 < 1\}$
- $\{x \mid 1 + \frac{1}{n}, n \in \mathbb{N}\}$
- $\{x \mid 1 + \frac{(-1)^n}{n}, n \in \mathbb{N}\}$

Exercise 3.10. For each item, compute the supremum and infimum, or explain why they does not exist. (You should explain your answers but you do not need to give a rigorous proof)

- $A = \left\{\frac{(-1)^n}{n} \mid n \in \mathbb{N}\right\}$$
- Fix $\beta \in (0, 1)$, and define $B = \{\beta^n \mid n \in \mathbb{N}\}$
- Fix $\gamma \in (1, \infty)$ and define $C = \{\gamma^n \mid n \in \mathbb{N}\}$.

Exercise 3.11. The proof of the nested interval theorem used the *endpoints* of the intervals crucially in the proof. One might wonder if the same theorem holds for *open intervals* (even though the proof would have to change).

Show the analogous theorem for open intervals is false by finding a counter example: can you find a collection of nested open intervals whose intersection is empty?

Exercise 3.12. Either give an example of each (explaining why your example works) or provide an argument (it doesn't have to be a formal proof) why no such example should exist:

- A sequence of nested closed intervals, whose intersection contains exactly n points, for some finite $n > 1$.
- A sequence of nested closed rays whose intersection is empty. (A closed ray has the form $[a, \infty)$ or $(-\infty, a]$ as in **?@def-intervals**).

3.5.1. ⋆ Equivalents to Completeness

Here we tackle the natural questions about why we chose *suprema* to codify completeness in a series of exercises.
Our goal at the end of these is to show that the following three possible completeness axioms are all logically equivalent:

- (1) Any nonempty set thats bounded above has a supremum.
- (2) Any nonempty set thats bounded below has an infimum.
- (3) Any nonempty set thats bounded has a supremum and infimum.

Exercise 3.13. For a set A let $-A$ denote the set of additive inverses: $-A = \{-a \mid a \in A\}$.
Prove that in a complete field if A is nonempty and bounded below then

$$\sup(-A) = -\inf(A)$$

Thus, assuming that suprema exist forces infima to exist, so in our list above, (1) implies (2).

Exercise 3.14. Prove the converse of the above: if we instead assume that the infimum of every nonempty set thats bounded below exists, show that the supremum of every nonempty set thats bounded above exists.

This shows (2) implies (1), so all together we know that (1) and (2) are equivalent.
But since (3) is just the conditions (1) and (2) together, we can derive (3) from either as

$$(1) \implies (1) \text{ and } (2) = (3)$$
$$(2) \implies (2) \text{ and } (1) = (3)$$

Thus both (1) and (2) imply (3). But since (1) and (2) are themselves special cases of (3), we already know (3) implies each of them! So, both of (1) and (2) are equivalent to (3), and all three conditions are logically equivalent to one another.

4. The Real Numbers

We have now carefully axiomatized the properties that are used in classical mathematics when dealing with the number line, defining a the structure of a *complete ordered field*.

Definition 4.1. A complete ordered field is an ordered field that satisfies the completeness axiom. Precisely, it is a set $\mathbb{F}$ with the following properties

- **Addition:** A commutative associative operation $+$, with identity 0, where ever element has an additive inverse.
- **Multiplication:** A commutative associative operation $\cdot$ with identity $1 \neq 0$, where every nonzero element has a multiplicative inverse.
- **Distributivity:** For all $a, b, c \in \mathbb{F}$ we have $a(b + c) = ab + ac$
- **Order:** A subset $P \subset \mathbb{F}$ called *the positives* containing exactly one of $x, -x$ for every nonzero $x \in \mathbb{F}$, which is closed under addition and multiplication: if $a, b \in P$ then $a + b \in P$ and $ab \in P$.
- **Completeness:** Every nonempty subset $A \subset \mathbb{F}$ which is bounded above has a least upper bound.

The subject of *real analysis* is the study of complete ordered fields and their properties, so everything that follows in this course logically follows from this set of axioms, *and nothing more.* The success and importance of the above definition is best exemplified by the following theorem:

Remark 4.1. This was very important work at the turn of the previous century; as neither step is a priori obvious. It's easy to write down axiom systems that don't describe *anything* because they're inconsistent (for example, add to ordered field axioms that all polynomials have at least one zero, and there is no longer such a structure), and its also common that axioms don't uniquely pick out a single object but rather describe an entire class (the axioms of a group define a whole subject, not a single example).

Theorem 4.1 (Uniqueness of the Reals)**.** *There exists a complete ordered field, and it is unique. We call this field* the real numbers *and denote it by* $\mathbb{R}$.

This theorem represents the culmination of much work at the end of the 19th and beginning of the 20th century to fully understand the real number line.

While not necessarily beyond our abilities, proving existence of a structure satisfying these axioms is a job for the set theorists and logicians that we will not tackle here.

Beyond providing justification for our usual way of speaking, the uniqueness of the reals is an important result to the history of mathematics. Its statement and proof in 1903 by Huntington marked the end of the era of searching for the fundamental principles behind the real numbers, and the beginning of the modern point of view, completely specifying their structure axiomatically.

Remark 4.2. The completeness axiom is what sets analysis apart from algebra, as it does not tell us how elements behave with respect to a given operation, but rather tells us about the *existence* of new elements. Indeed, this assertive ability of the completeness axiom is more radical than it seems at first, and can even be captured by mathematical logic: the other axioms are all *first order axioms*, whereas the completeness axiom is *second order.*

We will spend the majority of this course working out the properties of the real number line from these axioms, but its important to not loose sight of the bigger picture, *why* we are doing this. The real numbers provide a foundation for many objects in modern math:

- Complex numbers can be defined as pairs of real numbers $(x, y) := x + iy$ with component-wise addition and a new rule for multiplication
- Real and complex vector spaces can be constructed from n-tuples of real numbers, which lie at the foundation of much of modern mathematics, computer science, and physics.
- Manifolds are spaces which look locally like real vector spaces, and underly the modern fields of topology and differential geometry.

4.1. Dubious Numbers

Proposition 4.1 (Fields have no Nilpotent Numbers)**.** *Let \$$\mathbb{F}$ be any field, and ϵ some number where $\epsilon^2 = 0$. By the zero-product-property (Example 1.3), this implies $\epsilon = 0$. Thus there are no nonzero elements that square to zero.*

Theorem 4.2 (Infinite Numbers Do Not Exist)**.** *There are no infinite elements of* $\mathbb{R}$.

Proof. Assume for the sake of contradiction that there is some infinite number: without loss of generality (perhaps after multiplying by -1) we may assume its positive. Thus, this number is greater than every natural number, and so the natural numbers are bounded above.

Thus, by the completeness axiom, we find that the natural numbers must have a supremum. Denote this by $X = \sup \mathbb{N}$. So far, everything seems fine. But consider the number $X - 1$. This is smaller than X, and since X is the *least upper bound*, $X - 1$ cannot be an upper bound to $\mathbb{N}$. This means there must be some element $n \in \mathbb{N}$ with $n > X - 1$. But this means $X < n + 1$, and as $n + 1$ is a natural number whenever n is, we've run headfirst into a contradiction: X is not an upper bound at all! $\square$

It is an immediate corollary of this that infinitesimals also do not exist (but, because this is such an important result, we call it a theorem on its own.)

Theorem 4.3 (Infinitesimals Do Not Exist)**.** *There are no infinitesimal elements of* $\mathbb{R}$.

Proof. Let x be a positive element of $\mathbb{R}$, and consider its reciprocal $1/x$. By Theorem 4.2 $1/x$ is finite, so there's some $n \in \mathbb{N}$ with $n > 1/x$. Re-arranging the inequality shows $x > 1/n$ as required, so x is not infinitesimal. □

This argument shows that for a field, containing infinite elements and infinitesimal elements are *logically equivalent*: thanks to division, you can't have one without the other.

4.2. The Archimedean Property

A useful way to repackage the nonexistence of infinite numbers and infinitesimals into a *usable statement* known as the *Archimedean property*, as Archimedes took it as an axiom describing the number system in his paper *The Sphere and the Cylinder.* It also appears (earlier) as a definition in Euclid's elements: Book V Definition 4:

> Magnitudes are said to have a ratio to one another which can, when multiplied, exceed one another.

We rephrase this in precise modern terminology below:

Definition 4.2 (Archimedean Field)**.** A field $\mathbb{F}$ is *archimedean* if for every positive $a, b \in \mathbb{F}$ there is a natural number n with

$$na > b$$

Remark 4.3. While Archimedes himself attributes this to Eudoxus of Cnidus, it was named after Archimedes in the 1880s.

The important applications of this property all come from the case where b is really large, and a is really small. In an archimedean field, no matter how small a is you can always collect enough of them $na = a + a + a + \cdots + a$ to surpass b. A common way to remember this property is to poetically rephrase it as *you can empty the ocean with a teaspoon.*

Its possible to give an elementary proof (directly from the definition of rational numbers as fradtions p/q for $p, q \in \mathbb{Z}$, $q \neq 0$) that $\mathbb{Q}$ is an archimedean field:

Exercise 4.1 (The Rationals are Archimedean)**.** Prove the rationals are an archimedean field. *Hint: write a and b as fractions, can you figure out from the inequality you want, what n can be?*

Such a proof is not possible for $\mathbb{R}$ as we don't have an explicit description of its elements! All we know is its axiomatic properties. However, a proof is immediate using Theorem 4.2:

Theorem 4.4 (The Reals are Archimedean)**.** *Complete ordered fields satisfy the Archimedean property.*

Proof. Let a, b be positive real numbers. Since $b/a \in \mathbb{R}$ it is finite (by Theorem 4.2), so there is some $n \in \mathbb{N}$ with $n > \frac{b}{a}$, and thus $na > b$. □

Its also a short proof to show that archimedean fields cannot contain infinite elements (and thus also cannot contain infinitesimals), providing a useful equivalence:

Theorem 4.5. *The following three conditions are equivalent, for an ordered field* $\mathbb{F}$*:*

- $\mathbb{F}$ *is archimedean.*
- $\mathbb{F}$ *contains no infinite elements.*
- $\mathbb{F}$ *contains no infinitesimal elements.*

Proof. We already know the existence of infinite elements and infinitesimal elements are equivalent, so all we need to show is that $\mathbb{F}$ is archimedean if and only if all elements are finite.

But the proof of Theorem 4.4 already provides an argument that a field with only finite elements is necessarily archimedean, so we seek only the converse.

If $\mathbb{F}$ is archimedean, then for any positive $b \in \mathbb{F}$ we may take $a = 1$ and apply the archimedean property to get an $n \in \mathbb{N}$ with $n \cdot 1 > b$. For negative b, applying the ame to $-b$ results in a $n \in \mathbb{N}$ where $-n < b$, and together these imply all elements of $\mathbb{F}$ are finite. □

Remark 4.4. In fact one can be more precise than this: it turns out that the real numbers are the *largest possible archimedean field* - and every archimedean field fits somewhere between the rationals and the reals.

4.3. Irrationals

Definition 4.3 (Irrational Numbers)**.** A number $x \in \mathbb{R}$ is irrational if it is not rational.

4.3.1. Existence of $\sqrt{2}$

Our first goal is to prove that irrational numbers exist, by exhibiting one. We will use the example of the square root of two, and rigorously prove that $\sqrt{2}$ is a real number. (Just so you don't brush this off as trivial, its not immediately obvious: after all, $\sqrt{-2}$ is not a real number!)

Theorem 4.6. *Let* $\mathbb{F}$ *be archimedean, and consider the set*

$$S = \{r \in \mathbb{F} \mid r^2 < 2\}$$

Then if $\sigma = \sup S$ *exists,* $\sigma^2 = 2$.

We prove this rather indirectly, showing that both $\sigma^2 > 2$ and $\sigma^2 < 2$ are impossible, so the only remaining option is $\sigma^2 = 2$.

Example 4.1 ($\sigma^2 > 2$ is impossible.)**.** To show this is impossible, we will show if you have *any upper bound* $b \in \mathbb{F}$ with $b^2 > 2$, it's not the *least upper bound*, as we can make a smaller one.

Let b be any upper bound with $b^2 > 2$. To find a smaller upper bound, one idea is to try and find a natural number n where $\beta = b - 1/n$ works. That is,

$$\left(b - \frac{1}{n}\right)^2 > 2$$

Expanding this out, we see $b^2 - 2b/n + 1/n^2 > 2$, or after moving terms around, $b^2 - 2 > 2b/n - 1/n^2$. Now we need a little ingenuity: notice that $2b/n - 1/n^2$ is less than $2b/n$ (because we're subtracting something) so in fact, if we can find an n where $2b/n < b^2 - 2$ we're already good. Re-arranging this equation, we need to find n with

$$(b^2 - 2)n > 2b$$

But this is possible using the Archimedean property! Since $A = b^2 - 2$ and $B = 2b$ are both positive numbers, we can always find an $n \in \mathbb{N}$ where $nA > B$. Thus, we may choose this value of n, and note that $\beta = b - \frac{1}{n}$ is an upper bound for S that is smaller than b. Thus b was not the least upper bound!

Exercise 4.2 ($\sigma^2 < 2$ is impossible.)**.** Can you preform an argument similar to Example 4.1, to prove that $\sigma^2 < 2$ also leads to contradiction?

Since both the real numbers and the rationals are archimedean, the above applies to a consideration of either field

However applying the same knowledge to the reals yields the opposite conclusion, by virtue of the completeness axiom.

Theorem 4.7 ($\sqrt{2}$ is a Real Number)**.** *There exists a positive real number which squares to* 2.

Proof. Let $S = \{r \in \mathbb{R} \mid r^2 < 2\}$. Then, S is nonempty, as $0 \in S$ since $O^2 = 0$ and $0 < 2$. Next, we show that S is bounded above by 10:

Let $r \in S$ is arbitrary. Without loss of generality we may assume $r > 0$ as if $r < 0$ then certainly $r < 10$. By the definition of S, we know $r^2 < 2$ and thus clearly $r^2 < 100$. But recall Proposition 2.5: for positive a, b if $a^2 < b^2$ then $a < b$, so from $r^2 < 100$ we may conclude $r < 10$.

Knowing that S is both nonempty and bounded above, the completeness axiom applies to furnish us with a least upper bound $\sigma = \sup S$. And knowing its existence, Theorem 4.6 immediately implies that $\sigma^2 = 2$, so σ is by definition a square root of 2. □

Theorem 4.8 (The Rationals are Incomplete)**.** *Within the field of rational numbers, the set* $S = \{r \in \mathbb{Q} \mid r^2 < 2\}$ *is bounded above and nonempty, but does not have a supremum.*

Proof. The argument that S is nonempty and bounded above is identical to that in Theorem 4.7. And, Theorem 4.6 implies that *if* the supremum exists it must square to 2. But we know by Theorem 0.1 that there is no such rational number. Thus, the supremum must not exist, and so $\mathbb{Q}$ fails the completeness axiom. □

There is nothing special about 2 in the above argument, other than it is easy for us to work with. We could stop right now to prove the more general statement that all square roots exist:

Theorem 4.9 (Square Roots Exist)**.** *If* $x \in \mathbb{R}$ *is positive, then* $\sqrt{x}$ *is a real number.*

Though to not be *too* repetitive, we will hold off and prove this a different way, to illustrate more powerful tools in CITE.

Exercise 4.3. Prove that the product of a nonzero rational and an irrational number is irrational.

Exercise 4.4. The sum of two irrational numbers need not be irrational, as the example $\sqrt{2} - \sqrt{2} = 0$ shows. Prove or disprove: the sum of two *positive* irrational numbers is irrational.

4.3.2. Density

Definition 4.4 (Density)**.** Let S be a subset of an ordered field $\mathbb{F}$. Then S is *dense* in $\mathbb{F}$ if between any two elements $a, b \in \mathbb{F}$ with $a < b$ there is some $s \in S$ with

$$a < s < b$$

Theorem 4.10 (Density of the Rationals)**.** *The rational numbers are dense in the real numbers.*

Proof. We need to start with two arbitrary real numbers $a < b$, and find a rational number r between them. Let's do some scratch work: if $r = m/n$ and we want $a < m/n < b$ then it suffices to find an integer m between na and nb. This sounds doable!

Precisely, since $b - a > 0$, we can use the archimedean property to find some $n \in \mathbb{N}$ with $n(b - a) > 1$. Now since $nb - na > 1$, we just need to prove there's an integer m between them, and this'll be the number we want!

To rigorously prove this m exists, we can reason as follows: we know there are integers greater than na (since $\mathbb{R}$ has no infinite elements), so let m be the smallest such. Then by definition $m > na$, so all we need to show is $m < nb$. Since m is the smallest integer greater than na, we know $m - 1 < na$, or $m < na + 1$. But $na + 1 < nb$ so $m < nb$ as required.

Now we have a natural number n and an integer m with $na < m < nb$. Dividing through by n gives

$$a < \frac{m}{n} < b$$

□

As we have gotten used to being very careful in our arguments, you may think while working out the above argument to fill in a little lemma showing that every set of integers bounded below has a *minimum*. And, you could indeed do so by induction (try it - but fair warning, the argument is a little tricky! It's easiest with "strong induction" - what are we inducting over?). However this fact is actually *logically equivalent* to the principle of induction, and in foundations of arithmetic things are often reversed: we take this as an axiom, and prove induction from it! The statement is called the *well ordering principle.*

Definition 4.5 (The Well Ordering Principle)**.** Every nonempty subset of $\mathbb{N}$ has a least element.

Exercise 4.5 (Density of the Irrationals)**.** Use **?@thm-rationals-are-dense** above to prove that the irrationals are also dense in the reals.

Exercise 4.6. The *dyadic* rationals are the subset of $\mathbb{Q}$ which have denominators that are a power of 2 when written in lowest terms.

Prove the dyadic rationals are dense in $\mathbb{R}$.

4.4. Uncountability

We can use this to prove the uncountability of the reals using Cantor's original argument. (We will give the better known Cantor diagonalization argument later, once we've introduced decimals)

Theorem 4.11 ($\mathbb{R}$ is Uncountable)**.** *There is no bijection between* $\mathbb{N}$ *and* $\mathbb{R}$

Proof. Let $f : \mathbb{N} \to [0,1]$ be any function whatsoever. We can use this function to produce a sequence of points as follows:

$$f(1) = x_1, f(2) = x_2, f(3) = x_3 \dots$$

From this we can construct a set of nested intervals.
Let $I_1 \subset [0,1]$ be any closed interval that doesn't contain x_1. Then let $I_2 \subset I_1$ be a closed interval which does not contain x_2 (if x_2 was outside I_1, you could just take I_1 again, otherwise if its inside I_1 just take an interval on one side or the other of it). Continuing, we can easily choose an interval $I_{n+1} \subset I_n$ which doesn't contain x_{n+1}.

This gives us an infinite sequence of closed nested intervals inside a complete ordered field, so Theorem 3.3 tells us that their intersection must be nonempty. That is, there is some point $y \in [0,1]$ where $y \in I_n$ for all n.

What does this mean? Well, since $y \in I_1$ we know $y \neq x_1$ since I_1 was purpose-built to exclude x_1. Similarly $y \in I_2$ guarantees $y \neq x_2$, and so on...$y \in I_n$ means $y \neq x_n$. Thus, y is some point in $[0,1]$ which is not in our list!

Since $y \neq f(n)$ for any n, we see that our original (arbitrary) function cannot have been surjective. And, since bijections are both injective and surjective, this proves there *is no bijection* from $\mathbb{N}$ to $[0,1]$, so $[0,1]$ is uncountable! Then, as $[0,1] \subset \mathbb{R}$ we see $\mathbb{R}$ is uncountable as well. ☐

This has some pretty wild corollaries if you have studied countable sets before. Here's a couple examples

Corollary 4.1 (Transcendental Numbers)**.** *There exist real numbers which are not the solution of any algebraic equation with rational coefficients.*

Corollary 4.2 (Uncomputable Numbers)**.** *There exist real numbers which cannot be computed by any computer program.*

These are additional motivation for why we really need a precise theory of the real numbers: with very little work we've already *proven* that there is no way to study this number system with algebra alone - or even with the most powerful computer you could imagine.

4.5. ⋆ Infinity

The real numbers do not contain any infinite numbers, but dealing with the infinite is a key component of a real analysis course. To help us conduct such discussions rigorously we make clear what is meant (and what is *not* meant) by the infinity symbol familiar from prevous mathematics courses.

Definition 4.6. The symbol ∞ is a *formal symbol*: that is, a symbol that we agree to write, but do not attach any specific *value* to.

By default, any expression involving the symbol ∞ is considered *undefined.* We will use define certain contexts where the symbol ∞ is meaningful below.

Our first use of the symbol ∞ is to expand *interval notation* of the real numbers. Right now, using the order $<$ we have rigorously defined intervals such as (a, b), $[a, b)$ and $[a, b]$ for $a, b \in \mathbb{R}$.

Definition 4.7. For any real number a, we define the following intervals with $\pm\infty$ as an endpoint:

$$(-\infty, a) = \{x \in \mathbb{R} \mid x < a\}$$

$$(-\infty, a] = \{x \in \mathbb{R} \mid x \leq a\}$$

$$(a, \infty) = \{x \in \mathbb{R} \mid x > a\}$$

$$[a, \infty) = \{x \in \mathbb{R} \mid x \geq a\}$$

But we can take this farther, by actually adding the formal symbols $\pm\infty$ to our number system, to create a set called the *extended reals.*

Definition 4.8 (The Extended Reals)**.** The *extended real number line* is the set

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}.$$

Definition 4.9 (Ordering on $\overline{\mathbb{R}}$)**.** The order $<$ on $\mathbb{R}$ can be extended to $\overline{\mathbb{R}}$ by the following two rules:

$$\forall x \in \mathbb{R},\ x < \infty \qquad \forall x \in \mathbb{R},\ -\infty < x$$

This allows for interval notation on $\overline{\mathbb{R}}$ where, we may may write intervals such as $[-\infty 1]$ to mean the points $\{x \mid \overline{R} \mid x \leq 1\}$ etc.

In $\overline{\mathbb{R}}$ then, ∞ is an upper bound for every set, and $-\infty$ is a lower bound for every set. On the real numbers alone, the completeness axiom tells us that the supremum of *bounded* nonempty sets exist, but unbounded sets do not have a supremum. In the extended reals, we see that $\pm\infty$ naturally satisfy the definitions of

Proposition 4.2 (Unbounded Above means $\sup = \infty$)**.** *Let A be a nonempty subset of $\mathbb{R}$ which is not bounded above. Then as a subset of of the extended reals,* $\sup A = \infty$*.*

Proof. By the definition of ∞, we see that ∞ is an upper bound for A always, so we need only show it is the supremum. Let $x \in \overline{\mathbb{R}}$ be any element less than ∞. Then x must be an element of $\mathbb{R}$, and since A is not bounded above in $\mathbb{R}$, there is some $a \in A$ with $a > x$. Thus x is not an upper bound, and so every element less than ∞ fails to be an upper bound: that is, ∞ is the *least upper bound* as claimed. □

Exercise 4.7 (Unbounded Below means $\inf = -\infty$)**.**

Corollary 4.3 (Sup and Inf in the Extended Reals)**.** *Every nonempty subset of the extended real line has both an infimum and a supremum.*

Proof. Let A be a nonempty subset of $\overline{\mathbb{R}}$. First, if A contains ∞, then $\sup A = \infty$ as it is the maximum. So, we can consider the case that $\infty \notin A$. If A is bounded above by a real number, then $\sup A$ is also a real number by completeness, and if A is not bounded above, then $\sup A = \infty$ by Proposition 4.2.

The same logic applies to lower bounds: after taking care of the case where $\inf A = \min A = -\infty$, if A is bounded below completeness furnishes a real infimum, and if it is not, Exercise 4.7 shows the infimum to be $-\infty$. □

In the extended reals, it is still common to take the infimum and supremum of the empty set to be undefined. But there is also another option: one can assign $\inf \emptyset = \infty$ and $\sup \emptyset = -\infty$: if we do this then every set in the extended reals has an infimum and supremum!

4.6. ⋆ Topology

One final basic property of $\mathbb{R}$ that we will show follows from completeness is that its "connected" - it really does form a continuous line.

Definition 4.10 (Connected)**.** Let S be a subset of a topological space. Then a *separation* of S is a pair of disjoint open sets U, V whose union is S.

A subset is called *disconnected* if there is a separation, and *connected* if there is no way to make a separation.

Example 4.2 (A disconnected set)**.** Let $S = \{x \in \mathbb{R} \mid x > 0, x < 2, \text{ and } x \neq 1\}$. Then S is disconnected as we can write

$$S = (0,1) \cup (1,2)$$

And note these two intervals are both open, and dont share any points in common (so they are disjoint).

It's harder to imagine doing this for the interval $(0,2)$ however: if you try to imagine cutting it into two disjoint intervals at some point x, you're going to end up with $(0,x) \cup [x,2)$ or $(0,x] \cup (x,2)$. In either case, these intervals are not both open! To make them both open you could try $(0,x) \cup (x,2)$ but now they miss the point x (so their union isnt the whole space) or $(0,x+0.01)\cup(x-0.01,2)$ but now they overlap and aren't disjoint. Intuitively there's no way to do it - the interval $(0,2)$ is connected!

Theorem 4.12 (The Real Line is Connected)**.**

Proof. Assume for the sake of contradiction that $U \cup V$ is a separation of $\mathbb{R}$ (so, U, V are nonempty open sets and every point of $\mathbb{R}$ is in exactly one of them).

Choose some $x \in U$ and $y \in V$ - we can do this because they're nonempty - and without loss of generality assume that $x < y$. Considering the interval $[x,y]$ we know the left side is in U and the right in V, so we can define the

$$Z = \{z \in [x,y] \mid [x,z] \subset U\}$$

This set is nonempty (as $x \in Z$) and its bounded above (by y), so by completeness it has some supremum $\zeta = \sup Z$. Now the question is, which set is ζ in, U or V?

If $\zeta \in V$ then we know that since V is open ther's some small interval $(\zeta - \epsilon, \zeta + \epsilon)$ fully contained in V. But this means there's a number *smaller* than ζ contained in V, which means the interval $[0,\zeta]$ isnt fully contained in U, a contradiction!

If $\zeta \in U$ then we know since U is open, that there must be some tiny open interval $(\zeta - \epsilon, \zeta + \epsilon)$ around ζ contained U. This means there's a number *larger than ζ (for example, $\zeta + \epsilon/2$) where $[x, \zeta + \epsilon/2]$ is contained in U. So, ζ can't even be an upper bound to the set of all such numbers, a contradiction!

Both cases lead to contradiction, so there must be no such ζ, and hence no such separation. $\square$

Exercise 4.8 (Open Intervals of $\mathbb{R}$ are connected)**.** Prove that every open interval $(a, b) \subset \mathbb{R}$ is connected, mimicking the proof style above.

This *fails* for the rational numbers - they are not connected!

Theorem 4.13 (The Rationals are Not Connected)**.** *Consider the following two subsets of the rational numbers:*

$$A = \{x > 0 \mid x^2 > 2\}$$

$$B = \{x \in \mathbb{Q} \mid x \notin A\}$$

Then A and B form a separation of $\mathbb{Q}$.

Proof. A and B are open intervals in $\mathbb{Q}$ (they're the rational points of the open intervals $(\sqrt{2}, \infty)$ and $(-\infty, \sqrt{2})$). By definition every point of $\mathbb{Q}$ is in either A or B and they're disjoint. Since we just showed they are open, they form a separation, so $\mathbb{Q}$ is disconnected. $\square$

In fact, $\mathbb{Q}$ is *extremely disconnected* - this same argument applies at every irrational number of $\mathbb{R}$.

5. Functions

Highlights of this Chapter: we briefly explore the evolution of the modern conception of a function, and give foundational definitions for reference.

5.1. Freedom from Formulas

The term *function* was first introduced to mathematics by Leibniz during his development of the Calculus in the 1670s (he also introduced the idea of *parameters* and *constants* familiar in calculus courses to this day). In the first centuries of its mathematical life, the term function usually denoted what we would think of today as a *formula* or *algebraic expression*. For example, Euler's definition of function from his 1748 book *Introductio in analysin infinitorum* embodies the sentiment:

> A function of a variable quantity is *an analytic expression* composed in any way whatsoever of the variable quantity and numbers or constant quantities.

As a first step to adding functions to our theory of real analysis, we would somehow like to make this definition rigorous. But upon closer inspection, this concept, of "something expressible by a (single) analytic expression" is actually logically incoherent! For example, say that we decide, after looking at the definition of $|x|$, that it cannot be a function as it is not expressed as a single formula:

$$|x| = \begin{cases} -x & x \leq 0 \\ x & x > 0 \end{cases}$$

But we also agree that x^2 and $\sqrt{x}$ are both (obviously!) functions as they are given by nice algebraic expressions. What are we then to make of the fact that for all real numbers x,

$$\sqrt{x^2} = |x|$$

It seems we have found a perfectly good "single algebraic expression" for the absolute value after all! This even happens for functions with infinitely many pieces (which surely would have been horrible back then)

$$f(x) = \begin{cases} \vdots & \vdots \\ 3 + \sin(x) & x \in (0, \pi] \\ 1 + \sin(x) & x \in (\pi, 2\pi] \\ 3 + \sin(x) & x \in (2\pi, 3\pi] \\ \vdots & \vdots \end{cases}$$

This can be written as a composition involving just one piecewise function

$$f(x) = |1 + \sin x| + 2$$

Which can, by the earlier trick, be reduced to a function with no "pieces" at all:

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

So the idea of "different pieces" or different rules, seemingly so clear to us, is not a good mathematical notion at all! We are *forced* by logic to include such things, whether we aimed to or not. This became clear rather quickly, as even Euler had altered a bit his notion of functions by 1755:

> When certain quantities depend on others in such a way that they undergo a change when the latter change, then the first are called functions of the second. This name has an extremely broad character; it encompasses all the ways in which one quantity can be determined in terms of others.

The modern approach is to be much more open minded about functions, and define a function as *any rule whatsoever* which uniquely specifies an output given an input. This seems to have first been clearly articulated by Lobachevsky (of hyperbolic geometry fame) in 1834, and independently by Dirichlet in 1837

> The general concept of a function requires that a function of x be defined as a number given for each x and varying gradually with x. The value of the function can be given either by an analytic expression, or by a condition that provides a means of examining all numbers and choosing one of them; or finally the dependence may exist but remain unknown. (Lobachevsky)

> If now a unique finite y corresponding to each x, and moreover in such a way that when x ranges continuously over the interval from a\$ to b, $y = f(x)$ also varies continuously, then y is called a *continuous function of* x for this interval. It is not at all necessary here that y be given in

> terms of x by one and the same law throughout the entire interval, and it is not necessary that it be regarded as a dependence expressed using mathematical operations. (Dirichlet)

Through this definitions added generality comes *simplicity*: we are not trying to police what sort of rules can be used to define a function, and so the notion can be efficiently captured in the language of sets and logic.

Definition 5.1. A function from a set X to a set Y is an assignment to each element of X a *unique* element of Y. If we call the function f, we write the unique element of Y assigned to $x \in X$ as $y = f(x)$, and the entire function as

$$f : X \to Y$$

The definition of a function comes with three parts, so its good to have precise names for all of these.

Definition 5.2. If f is a function, its input set X is called the *domain*, and the set of possible outputs Y is called the *codomain*. The set of *actual outputs*, that is $R = \{f(x) \mid x \in X\}$ is called the *range*.

If the codomain of a function f is the real numbers, we call f a **real-valued function**. We will be most interested in real valued function throughout this course.

5.2. Composition and Inverses

Likely familiar from previous math classes, but it is good to get rigorous definitions down on paper when we are starting anew.

Definition 5.3 (Composition)**.** If $f : X \to Y$ and $g : Y \to Z$ then we may use f to send an element of X into Y, and follow it by g to get an element of Z. The result is a function from X to Z, known as the *composition*

$$g \circ f : X \to Z \qquad g \circ f(x) := g(f(x))$$

Every set has a particularly simple function defined on it known as the *identity function*: $\mathrm{id}_X : X \to X$ is the function that takes each element $x \in X$ and *does nothing*: $\mathrm{id}_X(x) = x$. These play a role in concisely defining inverse functions below:

Definition 5.4 (Inverse Functions)**.** If $f : X \to Y$ is a function, and $g : Y \to X$ is another function such that

$$g \circ f = \mathrm{id}_X \qquad f \circ g = \mathrm{id}_Y$$

Then f and g are called *inverse functions* of one another, and we write $g = f^{-1}$ if we wish to think of g as inverting f, or $f = g^{-1}$ rather we started with g, and think of f as undoing it.

Example 5.1. The function $f(x) = 2x$ and $g(x) = x/2$ are inverses of one another as functions $\mathbb{R} \to \mathbb{R}$.

The squaring function $s : \mathbb{R} \to \mathbb{R}$ defined by $s(x) = x^2$ has the square root as an inverse, only if the domain and codomain are restricted to the nonnegative reals. Otherwise, we see that $s(-2) = 4$ and $\sqrt{4} = 2$ so $\sqrt{\ } \circ s$ is not the identity: it takes -2 to 2!

5.3. Useful Terminology

Definition 5.5 (Restricting the Domain)**.** Given a function f with domain D, the restriction to a subset $S \subset D$ is denoted $f|_S$.

Definition 5.6. Given a function with a domain D, an *extension* of f to a set $X \supset D$ is a function $\tilde{f} : X \to \mathbb{R}$ such that $\tilde{f}|_D = f$.

Definition 5.7 (Increasing / Decreasing)**.** A function f is (monotone) increasing if for all $x \leq y$ we have $f(x) \leq f(y)$. It's monotone decreasing if instead $x \leq y$ implies $f(x) \geq f(y)$. A function is *strictly* increasing if $x < y$ impleis $f(x) < f(y)$, and analogously for strictly decreasing.

Exercise 5.1. If f is a strictly increasing function, then it is one-to-one: every output y is achieved by a unique input x.

This exercise implies that strict monotone functions are invertible, as the inverse of any one-to-one function is defined by sending a given y to the unique x that maps to it.

Definition 5.8 (Convexity)**.** Let f be a function defined on some interval (possibly all of $\mathbb{R}$). Then f is convex if for any interval $[x, y] \subset \mathrm{dom} f$, the value of f at the midpoint exceeds the average value of f at the endpoints:

$$\forall x, y \quad f\left(\frac{x+y}{2}\right) \geq \frac{f(x) + f(y)}{2}$$

A function f is said to be *monotone* or *convex* (etc) on a set S if the restriction of f to S is monotone / convex.

Definition 5.9 (Local Extrema)**.**

- Increasing Decreasing
- Convex
- Local Extrema

5.4. A Zoo of Examples

Example 5.2 (Polynomial Functions)**.** A polynomial function is an assignment $p:\ \mathbb{R} \to \mathbb{R}$ which takes each x to a linear combination of powers of x:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$$

The highest power of x appearing in p is called the *degree* of the polynomial.

The idea of a *function defined by a formula* can be extended even farther by allowing the field operation of division; though this time we must be careful about the inputs.

Example 5.3 (Rational Functions)**.** A rational function is a an assignment

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

where p and q are polynomials. Rational functions are real-valued, but their domain is not all of $\mathbb{R}$: at any zero of q the formula above is undefined, a rational function is only defined on the set of points where q is nonzero.

We already saw that piecewise formulas count in our modern definition, but perhaps didn't fully think through the implications: they can be *very, very piecewise*

Example 5.4 (The Characteristic Function of $\mathbb{Q}$)**.** The function $f:\ \mathbb{R} \to \mathbb{R}$ defined as follows

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

Here's another monstrous piecewise function we will encounter again soon:

Example 5.5 (Thomae's Function)**.** This is the function $\tau:\ \mathbb{R} \to \mathbb{R}$ defined by

$$\tau(x) = \begin{cases} \frac{1}{q} & x \in \mathbb{Q} \text{ and } \frac{p}{q} \text{ is lowest terms.} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

We've stressed that functions don't need to be given by explicit formulas, so we should give an example of that: here's a function that is defined at each point as a different limit (using the completeness axiom)

Example 5.6. A function may be defined for each $x \in \mathbb{R}$ as the limit of a sequence, such as

$$E(x) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!}$$

A function can also be defined by a less explicit limit procedure, like the limits defining powers: where we've previously seen that any sequence $r_n \to x$ of rationals converging to x produces the same limiting value of a^{r_n}.

Example 5.7 (Exponential as Powers). For any $x \in \mathbb{R}$ and $a \geq 0$ the function $f(x) = a^x$ is defined by

$$a^x = \lim_n a^{r_n}$$

for r_n a sequence of rational numbers converging to x.

A function can also be defined by an *existence proof* telling us that a certain relationship determines a function, without giving us any hint on how to compute its value:

Example 5.8 ($\sqrt{\cdot}$ defined by an existence theorem). We proved that for every $x \geq 0$ that there exists some number $y > 0$ with $y^2 = x$, back in our original study of completeness (Theorem 4.9).

We can easily see that such a number is *unique*: if $y_1 \neq y_2$ then by the order axioms one is greater: without loss of generality $0 < y_1 < y_2$. Thus $y_1^2 < y_2^2$, so we can't have both $y_1^2 = x$ and $y_2^2 = x$, and $x \to y = \sqrt{x}$ is a function.

Alright - that's plenty of examples to get ourselves in the right mindset. Let's give a non-example, to remind us that while there need not be formulas, the modern notion of function is not '*anything goes*'!

Example 5.9. The assignment taking an integer to one of its prime factors does *not* define a function. This would take the integer 6 to both 2 and 3, and part of the definition of a function is that the output is *unique* for a given input.

Exercise 5.2 (Invertibility implies Monotonicity). Let f be an invertible function. Prove that f is (strictly) monotone increasing, or (strictly) monotone decreasing.

Part II.

Sequences

This part of the text covers the elementary theory of sequences:

- In Chapter 6 we define sequences and convergence, and see how to prove $\lim s_n = L$ directly from the definition.
- In Chapter 7 we study the arithemetic of convergent sequences, and prove the limit laws familiar from an introductory calculus course.
- In Chapter 8 we prove the *monotone convergence theorem* which gives simple conditions that ensure the convergence of a sequence, and use this to study infinite processes and the square root calculating algorithm of the babylonians.
- In Chapter 9 we extend the reach of our theory to cover *non-monotone* sequences, by decomposing them into subsequences and investigating the resulting limits.
- In Chapter 10 we define the notion of a *Cauchy sequence*, and prove it is equivalent to convergence. This lets us study all sorts of new convergent sequences, such as contraction maps.
- In **?@sec-sequences-iterated** we get a first look at the complications that arise when multiple limits interact in a single expression. Such limits underlie many interesting situations in analysis, from the theory of power series, to the commutativity of partial derivatives and the ability to differentiate under the integral.

6. Convergence

Having formalized the number line, we can now get to work. If we want to rigorously understand any of the approximation efforts of the ancients, we must think about *sequences.*

Definition 6.1 (Sequence)**.** A sequence is an infinite ordered list of numbers

$$(s_1, s_2, s_3, \ldots, s_n, \ldots)$$

Each individual element is a *term* of the sequence, with an subscript (the *index*) denoting its position in the list.

Most often, we take the set of indices to be $1, 2, 3, \ldots$, but any infinite subset of the integers will do. For example, the sequence p_n of perimeters of inscribed n-gons starts with index 3 (the triangle), as this is the smallest polygon. And, the subsequence Archimedes used to calculate π started with the hexagon and then iterated doubling: $P_6, P_{12}, P_{24}, \ldots$ so has index set

$$\{6, 12, 24, 48, 96, 192, 384, \ldots\}$$

Formally, we note all of this is captured using functions, though we will not need this perspective during our day-to-day usage of sequences.

Remark 6.1. Let $I \subset \mathbb{Z}$ be any infinite set of indices. Then a sequence is a *function* $s : I \to \mathbb{R}$.

While sequence itself is just an infinite ordered list of numbers, to work with such an object we often require a way to compute its terms. Sometimes this is hard! For example, the sequence

$$\pi_n = \text{ the number of prime numbers} \leq n$$

Is called the *prime counting function*, and being able to compute its exact values efficiently would be monumental progress in number theory. In practice, sequences that we *can* compute with efficiently are often presented to us in one of two ways:

- **Closed Formula** For each n, we are given some formula of the type familiar from high school mathematics, and plugging n into this formula yields the n^{th} term of the sequence. Some examples are

$$a_n = \frac{n^2+1}{3n-2}, \qquad b_n = \sin\left(\frac{1}{n}\right) \qquad c_n = \sqrt{1+\frac{\sqrt{n}}{n+1}}$$

- **Recursive Definition** For each n, we are not given a formula to compute s_n directly, but rather we are given a formula to compute it from the previous value s_{n-1}.

Here's some example sequences that are important both to us, and the history of analysis:

Example 6.1 (Babylonians and $\sqrt{2}$)**.** Starting from rectangle of width and height w, h, the Babylonians created a new rectangle whose width was the average of these, and whose height was whatever is required to keep the area 2:

$$w_{\text{new}} = \frac{w+h}{2} \qquad h_{\text{new}} = \frac{2}{w_{\text{new}}}$$

This because we can solve for h in terms of w, this induces a recursive sequence for the widths. Starting from some (w_n, h_n) we have

$$w_{n+1} = \frac{w_n+h_n}{2} = \frac{w_n + \frac{2}{w_n}}{2} = \frac{w_n}{2} + \frac{1}{w_n}$$

Thus, in modern terminology the babylonian procedure defines a recursive sequence, given any starting rectangle. If we begin with the rectangle of wdith 2 and height 1, we get

$$w_0 = 2, \qquad w_{n+1} = \frac{w_n}{2} + \frac{1}{w_n}$$

Exercise 6.1 (Babylonians and $\sqrt{2}$)**.** Following the same type of reasoning as for width, use the babylonian procedure to produce a recursive formula for the sequence of heights h_n, for a rectangle starting with $h = 1$.

Example 6.2. An *infinite sum* is a type of recursively defined sequence, built from another sequence called its *terms*. Assume that a_n is any sequence. Then we build a sequence s_n by

$$s_0 = a_0 \qquad s_{n+1} = s_{n-1} + a_n$$

Unpacking this, we see that $s_1 = s_0 + a_1 = a_0 + a_1$, and thus $s_2 = s_1 + a_2 = a_0 + a_1 + a_2$ etc.

6.1. Convergence

The reason to define a sequence precisely is that we are interested in making rigorous the idea of *infinitely many steps*, the way the Babylonians may have pictured running their procedure an infinite number of times to produce a perfect square, or Archimedes who ran his side-doubling procedure infinitely many times to produce a circle.

In both cases, there was some number L *out there at infinity* that they were probing with a sequence. We call such a number L the *limit of the sequence*.

Definition 6.2 (Convergent Sequence)**.** A sequence s_n *converges* to a limit L if for all $\epsilon > 0$ there is some threshold N past which every further term of the sequence is within ϵ of L. Formally, this is the logic expression

$$\forall \epsilon > 0\, \exists N\, \forall n > N\ |s_n - L| < \epsilon$$

When a sequene converges to L we write

$$\lim s_n = L \qquad \text{or} \qquad s_n \to L$$

A sequence is *divergent* if its not convergent. The definition of convergence formalizes the idea the ancients sought **if you keep calculating terms, you'll get as close as you like to the number you seek**

That is, the definition sets up a challenge between you (the computer of the sequence) and the error tolerance. Once you set a certain amount of acceptable error ϵ, the definition furnishes an N and guarantees that if you compute the sequence out until N you'll be within the tolerated error - and if you keep computing more terms, the approximation will never get worse. Its good to look at some specific examples, while getting comfortable with this:

Exercise 6.2 (Understanding Convergence)**.** Consider the sequence $a_n = \frac{1}{n^2+13}$. Feel free to use a calculator (even just the google search bar) to experiment and answer these questions.

- What value L do you think this sequence converges to?
- If $\epsilon = 1/10$, what value of N ensures that a_n is always within ϵ of L for, $n > N$?
- If $\epsilon = 1/100$, what value of N ensures that a_n is always within ϵ of L for, $n > N$?

Exercise 6.3 (Convergence and $\sqrt{2}$)**.** This problem concerns the babylonian sequence for $\sqrt{2}$ in Example 6.1. Again, use a calculator to play around and answer the following

- For which value of N are we guaranteed that w_n calculates the first **two** decimal places $\sqrt{2}$ correctly, when $n > N$?
- For which value of N are we guaranteed that w_n calculates the first **eight** decimal places $\sqrt{2}$ correctly, when $n > N$?

6.1.1. The $\epsilon - N$ Game

To prove a sequence converges, we need to work through the string of quantifiers $\forall\epsilon\exists N\forall n$... This sets up a sort of imagined *battle* between an imagined foe setting a value of ϵ, and you needing to come up with an N such that you can get the sequence within ϵ of the limit.

Here's one incredibly useful example, that will serve as the basis of many future calculations.

Proposition 6.1 ($1/n$ converges to 0.)**.** *Prove that the sequence $s_n = 1/n$ of reciprocals of the natural numbers converges to* 0*.*

Proof. Let $\epsilon > 0$. Then set $N = 1/\epsilon$, and choose arbitrary $n > N$. Since $n > 1/\epsilon$ it follows that $1/n < \epsilon$, and hence that

$$\left|\frac{1}{n} - 0\right| < \epsilon$$

Since $n > N$ was arbitrary, this holds for all such n, and we have proved for this ϵ, theres an N with $n > N$ implying the sequence $1/n$ is within ϵ of the proposed limit 0. Since ϵ was also arbitrary, we have in fact proved this for *all positive epsilon*, and thus we conclude

$$\frac{1}{n} \to 0$$

$\square$

Often when working out such a computation, the scratch work is backwards of the final proof. In a proof, you need to fix an arbitrary epsilon, then

Exercise 6.4 ($\frac{n}{n+1}$ converges to 1)**.**

Sometimes the scratch work takes a bit more thinking or algebraic manipulation. Its OK if the scratch work isn't fully rigorous or perfectly written, as long as the eventual proof is! Here's an example of some scratch work taking a naive approach (just "solve for n") that arrives at an easy bound, and a nice formal proof verifying it.

Example 6.3 ($\frac{n}{n^2+1}$ converges to 0.)**.**

Scratch. We want $\frac{n}{n^2+1} < \epsilon$, and we attempt to solve this inequality for n by multiplying through and using the quadratic formula:

$$n < \epsilon n^2 + \epsilon \implies 0 < \epsilon n^2 - n + \epsilon$$

The zeroes satisfy

$$n = \frac{1 \pm \sqrt{1-4\epsilon^2}}{2\epsilon}$$

The larger of these is the one with the + sign, so as long ad n is bigger than this the proof will work. This term as written is rather annoying to deal with as we have $1-4\epsilon^2$ under the square root, and to do a formal proof using it as is, we'd need to ensure this wasn't negative. But since we are only looking for a bound, we can use that

$$\frac{1+\sqrt{1-4\epsilon^2}}{2\epsilon} < \frac{1+1}{2\epsilon} = \frac{1}{\epsilon}$$

and just require $n > \frac{1}{\epsilon}$. □

Formal. Let $\epsilon > 0$ and set $N = \frac{1}{\epsilon}$. For any $n > N$ we see that $\frac{1}{n^2} < \epsilon^2$, and $\frac{1}{n^2+1} < \frac{1}{n^2}$ so $\frac{1}{n^2+1} < \epsilon^2$. Multiplying by n gives

$$\frac{n}{n^2+1} < n\epsilon^2 < \frac{1}{\epsilon}\epsilon^2 = \epsilon$$

Thus for any $n > N$ we have $\left|\frac{n}{n^2+1} - 0\right| < \epsilon$ so the sequence converges to 0 by definition. □

Example 6.4 ($\frac{1}{2^n} \to 0$)**.** Here's a sketch of an argument: you should fill in the details. Let $\epsilon > 0$. Then we want to find an N where $n > N$ implies $1/2^n < \epsilon$. First, we prove by induction that $2^n \geq n$ for all n. Thus, $1/2^n < 1/n$, and so it suffices to find N where $1/n < \epsilon$. But this is exactly what we did above in the proof that $1/n \to 0$. So this is possible, and hence $1/2^n \to 0$.

Exercise 6.5. Give an example of the following, or explain why no such example can exist.

- A sequence with infinitely many terms equal to zero, but does not converge to zero.
- A sequence with infinitely many terms equal to zero, which converges to a nonzero number.
- A sequence of irrational numbers that converges to a rational number.
- A convergent sequence where every term is an integer.

Exercise 6.6. Prove, directly from the definition of convergence, that

$$\frac{2n-2}{5n+1} \to \frac{2}{5}$$

6.1.2. Divergence

The definition of convergence picks out a very nice class of sequences: those that *get arbitrarily close to a fixed value, as their index grows.* The rest of sequences - anything that does not have this nice property, are all lumped into the category of *divergent.*

Definition 6.3 (Divergence)**.** A sequence diverges if its *not true* that for any ϵ you can find an N where beyond that, all terms of the sequence differ from some constant (the limit) less than ϵ.

Phrasing this positively: a sequence a_n diverges if for *every value of* a, there exists some $\epsilon > 0$ where no matter which N you pick, there's always some $n > N$ where $|a_n - a| > \epsilon$. There's a lot of quantifiers here! Written out in first order logic:

$$\forall a \in \mathbb{R}\ \exists \epsilon > 0\ \forall N\ \exists n > N\ \ |a_n - a| > \epsilon$$

Again, its easiest to illustrate with an example:

Example 6.5 ($(-1)^n$ Diverges)**.** Here's the idea: The sequence $s_n = (-1)^n$ alternates back and forth from 1 to -1 forever. Assume for the sake of contradiction that it in fact converges to some real number L. Then (by definition) eventually all terms must get within ϵ of L, but this is impossible if ϵ is small as every term differs from its successor by 2.

Proof. Note that for all n, $|s_n - s_{n+1}| = 2$ as when n is even this is $|1 - (-1)| = |2|$ and when n is odd its $|-1-(1)| = |-2|$. Assume for the sake of contradiction that $s_n \to L$ for some $L \in \mathbb{R}$, and set $\epsilon = \frac{1}{2}$. This implies there exists an N such that for all $n > N$ we have $|s_n - L| < \frac{1}{2}$. Choosing some $n > N$ we use the triangle inequality to see

$$2 = |s_n - s_{n+1}| = |s_n - L + L - s_{n+1}| \leq |s_n - L| + |L - s + n + 1| \leq \epsilon + \epsilon = 1$$

Thus we've proven $2 < 1$ which is a contradiction, so it must not be true that $s_n \to L$ for any L: the sequence diverges. ☐

Definition 6.4 (Diverging to $\pm\infty$)**.** A sequence s_n diverges to ∞ if for all $M > 0$ there exists an threshold past which the sequence is always larger than M. As a logic statement,

$$\forall M > 0\, \exists N\, \forall n > N\ s_n > M$$

Exercise 6.7 (n^2 diverges to ∞.)**.**

Exercise 6.8.

- Give an example of two divergent sequences a_n, b_n where $a_n + b_n$ is convergent.
- Give an example of two divergent sequences a_n, b_n where $a_n b_n$ is convergent.

6.2. Uniqueness

Theorem 6.1 (Limits are unique). *Let a_n be a convergent sequence. Then there exists a unique $a \in \mathbb{R}$ with $a_n \to a$.*

Here's a sketch of the idea, which uses several big ideas that can be recycled in similar arguments:

- We prove uniqueness by showing that if x and y were both limits, then $x = y$.
- We prove $x = y$ by showing that for every $\epsilon > 0$ the difference $|x - y| < \epsilon$.
- We prove $|x - y| < \epsilon$ by an $\epsilon/2$ argument:
  - We add zero in a clever way: $|x - y| = |x - a_n + a_n - y|$
  - We use the triangle inequality $|x - a_n + a_n - y| \leq |x - a_n| + |a_n - y|$
  - We use the fact that $a_n \to x$ and $a_n \to y$ to make each of $|a_n - x|$ and $|a_n - y|$ less than $\epsilon/2$.

Proof. Assume that a sequence a_n converges to two limits $a_n \to x$ and $a_n \to y$. Then for any ϵ we can find an N_1 where $n > N_1$ implies $|a_n - x| < \epsilon/2$ and an N_2 where $n > N_2$ implies $|a_n - y| < \epsilon/2$. Setting $N = \max\{N_1, N_2\}$ we see for any $n > N$ that

$$|x - y| = |x - a_n + a_n - y| \leq |x - a_n| + |a_n - y| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus for *any* positive ϵ we know $|x - y| < \epsilon$, so in particular $|x - y| \neq \epsilon$, and $|x - y|$ can't be positive. Since absolute values are always nonnegative, the only remaining option is that $|x - y| = 0$. But this means $x - y = 0$ and hence $x = y$: the two limits are equal. □

There's one more *uniqueness-type* theorem about limits that's useful to get a handle on. We just saw that the limit is uniquely determined by the sequence, but we can say something slightly stronger. Its uniquely determined by the *end of the sequence*: if you throw away the first finitely many terms, it won't change the limit.

Definition 6.5. A *shifted sequence* the result of shifting the indices by a constant k, deleting the first k terms. Precisely, given a sequence a_n and some $k \in \mathbb{N}$, the sequence $s_n = a_{n+k}$ is the result of shifting a by k.

$$s_0 = a_k, s_1 = a_{k+1}, s_2 = a_{k+2}, \ldots$$

Proposition 6.2. *Shifting a convergent sequence does not change its limit.*

Scratch Work. Assume that a_n converges to a, and define the sequence s_n by deleting the first k terms of a_n, that is, $s_n = a_{n+k}$. We claim that $s_n \to a$.

Let $\epsilon > 0$ and choose an N such that if $n > N$ we know that $|a_n - a| < \epsilon$ (we know such an N exists by the assumption $a_n \to a$). Now consider $|s_n - a|$. Since $s_n = a_{n+k}$, we know $|s_n - a| < \epsilon$ because we already knew $|a_{n+k} - a| < \epsilon$: we knew this for *every single index bigger than N*.

Thus, for all $n > N$ we have $|s_n - a| < \epsilon$, which is the definition of $s_n \to a$. □

This can be generalized, to show that any two sequences which are eventually the same have the same limit. Since the first finite part of any sequence is irrelevant to its limiting behavior, its nice to have a word for "the rest of the sequence, after throwing away an unspecified amount at the beginning". This is called the *tail*.

Definition 6.6 (Tail of a Sequence)**.** The *tail of a sequence* is what remains after chopping off an arbitrary (finite) number of terms from the beginning of the sequence. Two sequences *have the same tail* if they agree after some point: more precisely, a_n and b_n have the same tail if there is an N_a and N_b such that for all $k \in \mathbb{N}$

$$a_{N_a+k} = b_{N_b+k}$$

Example 6.6 (Tail of a Sequence)**.** The following two sequences have the same tail:

$$a_n = 1, 1, 4, 3, 1, 5, 1, 3, 1, 4, 7, 8, 9, 10, 11, 12, 13, 14, \ldots$$

$$b_n = -4, 3, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 \ldots$$

We can see this because $a_{13} = b_3 = 9$, and $a_{14} = b_4 = 10$, and $a_{15} = b_5 = 11$...for every k we have that $a_{13+k} = b_{3+k}$ so they agree after chopping the first 12 terms off of a_n and the first two terms off of b_n.

Exercise 6.9 (Convergence only depends on the tail)**.** If two sequences have the same tail, then they either both converge or both diverge, and if they converge, they have the same limit.

6.3. Important Sequences

We will soon develop several theorems that let us calculate many limits without tediously chasing down an N for every ϵ. But there are still several 'basic limits' that we will need to know, that will prove useful as building blocks of more complicated limits, as well as foundations to further theory in analysis. We compute several of them here: you should not worry too hard about committing these to memory; but rather read the proofs as examples of how to play the $\epsilon - N$ game in tricky situations.

The first and most important is familiar from above:

Example 6.7. As $n \to \infty$ the sequence $1/n$ converges to 0.

The next is useful in developing the theory of power series, and other things. We'll prove it using Bernoulli's inequality:

Example 6.8. Let $|a| < 1$, then the sequence a^n of repeated powers of a converges to 0.

Proof. First note that if $a = 0$ then $a^n = 0^n = 0$, which clearly converges to zero so we can assume $a \neq 0$ in what follows.

Let $\epsilon > 0$. We wish to find an N such that $n > N$ implies $|a^n - 0| = |a^n| = |a|^n < \epsilon$. Since $|a| < 1$ by assumption, we know $1/|a| > 1$ and so we may write $1/|a| = 1 + x$ for some $x > 0$. This is helpful because Bernouli's inequality tell us that

$$(1+x)^n > 1 + nx > nx$$

and so, we know

$$|a|^n = \frac{1}{(1+x)^n} \leq \frac{1}{1+nx} \leq \frac{1}{nx}$$

For this to be less than ϵ, we need $1/nx < \epsilon$, or $n > \frac{1}{x\epsilon}$. Thus setting $N = \frac{1}{x\epsilon}$ will do it. □

This proof is OK as written because it produces a value of N in terms of epsilon, and also shows (in the previous line) for larger n that the quantity we are trying to bound is even smaller. But it might be more *readable* to rewrite it in "reverse": first proposing the value $N = 1/x\epsilon$ we discovered, and then showing it works (do this as an exercise). As another exercise, its useful to look at what happens for $a > 1$.

Exercise 6.10. If $a > 1$ then a^n diverges to infinity.

This next is an essential building block of the theory of exponential functions: we again make use of Bernoulli's inequality in the proof below, but suggest an alternative (second) proof as an exercise

Example 6.9. Let $a > 0$. Then the sequence $a^{1/n}$ converges to 1.

Proof. We proceed in two cases, starting first with $a > 1$. Fix an $\epsilon > 0$; we want to find an N where $n > N$ implies $|a^{1/n} - 1| < \epsilon$. Since $a < 1$ is positive, it follows that $a^{1/n} > 1^{1/n} = 1$. Call $b_n = a^{1/n} - 1$ the quantity we are trying to show is small. Then we can apply Bernoulli's inequality to see

$a = \left(a^{1/n}\right)^n = (1 + b_n)^n \geq 1 + nb_n \geq nb_n$\$

Thus, for all n, we see $0 \leq b_n \leq \frac{a}{n}$, and for this to be less than ϵ, we need only assure $a/n < \epsilon$ so $n > a/\epsilon$: choosing $N = a/\epsilon$ will do.

$$n > N = \frac{a}{\epsilon} \implies b_n = |a^{1/n} - 1| < \epsilon$$

Now we turn to the second case, which is $0 < a < 1$. Note that for any a in this case we know $1/a > 1$, so the work we just did applies directly to $1/a$, and $\left(\frac{1}{a}\right)^{1/n} \to 1$. Unpacking this, that means for any $\epsilon > 0$ there is some N where $n > N$ implies

$$\left|\frac{1}{a^n} - 1\right| < \epsilon$$

We can simplify the fraction inside the aboslute value and then multiply the entire inequality through by $a^{1/n}$ to see

$$\left|\frac{1 - a^{1/n}}{a^{1/n}}\right| < \epsilon \implies |1 - a^{1/n}| < \epsilon a^{1/n}$$

But since $0 < a < 1$ we know $a^{1/n} < 1^{1/n} = 1$ so the right hand side is *already* less than ϵ, and we are done. □

Exercise 6.11. In this problem you give an *altenative* proof that $a^{1/n} \to 1$ for all $a > 0$, using the geometric sum. Recall, this stated that for any $|r| < 1$ and $n \in \mathbb{N}$,

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

- First consider only the case that $a > 1$. Show that the geometric sum can be rewritten $r^n - 1 = (r - 1)(1 + r + r^2 + \cdots + r^{n-1})$, and use this to prove that $a - 1 \geq n(a^{1/n} - 1)$. *Hint: apply it to* $r = a^{1/n}$ *and do some estimating.*
- Use this to show that $0 \leq a^{1/n} - 1 \leq \frac{a-1}{n}$ for all n, and then prove (either directly, or using the squeeze theorem) that this implies $a^{1/n} \to 1$.
- Now consider the case $0 < a < 1$ and show the same. *Hint: if* $0 < a < 1$ *then* $1/a > 1$, *and perhaps you can do a similar trick to the textbook?*

As a (challenging!) exercise, one might consider what happens if instead of taking the n^{th} root of a fixed constant, we take the n^{th} root of n itself. This

Exercise 6.12. The sequence $n^{1/n}$ converges to 1.

6.4. ⋆ Topology

With an eye to topology, everything about sequences and convergence can be rephrased in terms of open sets, instead of with talk about ϵ and inequalities.

Definition 6.7 (Neighborhoods)**.** A *neighborhood* of a point x is any open set U containing x. The ϵ-neighborhood of x is the neighborhood $U_\epsilon = (x - \epsilon, x + \epsilon)$

Definition 6.8 (Convergence and ϵ-Neighborhoods)**.** A sequence a_n converges to a if every ϵ neighborhood contains all but finitely many terms of the sequence.

That this is equivalent to Definition 6.2, because the definition of epsilon neighborhood exactly captures the interval discussed in the original definition of convergence.

Exercise 6.13 (Convergence and ϵ-Neighborhoods)**.** The definition of convergence in terms of epsilon neighborhoods is equivalent to the usual definition in terms of absolute values and inequalities.

The definition of an epsilon neighborhood makes sense only somewhere like the real line, where we can talk about intervals. So, the general topological definition must dispense with this notion and talk just about open sets:

Definition 6.9. A sequence a_n converges to a if every neighborhood contains all but finitely many terms of the sequence.

Exercise 6.14 (Convergence and Neighborhoods)**.** Prove this is equivalent to convergence using ϵ neighborhoods. *Hint: show that every neighborhood contains some epsilon neighborhood. Can you show that is enough?*

6.5. Problems

Exercise 6.15. Come up with a recursive sequence that could be used to formally understand the infinite expression below:

$$\sqrt{1+\sqrt{1+\sqrt{1+\sqrt{1+\sqrt{1+\sqrt{1+\cdots}}}}}}$$

Exercise 6.16. Given two sequences x_n, y_n that converge to a, show the interleaved sequence $x_1, y_1, x_2, y_2, \ldots$ converges to a.

Exercise 6.17. Let $a_n \to 0$ and let b_n be a sequence such that for all n you know $|b_n - L| < a_n$. Prove that $\lim b_n = L$.

7. Limit Laws

Highlights of this Chapter: We develop techniques for bounding limits by inequalities, and computing limits using the field axioms. We use these techniques to prove two interesting results:

- The Babylonian sequence approximating $\sqrt{2}$ truly does converge to this value.
- Given any real number, there exists a sequence of rational numbers converging to it.

Now that we have a handle on the definition of convergence and divergence, our goal is to develop techniques to avoid using the definition directly, wherever possible (finding values of N for an arbitrary ϵ is difficult, and not very enlightening!)

The natural first set of questions to investigate then are how our new definition interacts with the ordered field axioms: can we learn anything about limits and inequalities, or limits and field operations? We tackle both of these in turn below.

7.1. Limits and Inequalities

Proposition 7.1 (Limits of nonnegative sequences)**.** *Let a_n be a convergent sequence of nonnegative numbers. Then* $\lim a_n$ *is nonnegative.*

Proof. Assume for the sake of contradiction that $a_n \to L$ but $L < 0$. Since L is negative, we can find a small enough epsilon (say, $\epsilon = |L|/2$) such that the entire interval $(L - \epsilon, L + \epsilon)$ consists of negative numbers.

The definition of convergence says for this ϵ, there must be an N where for all $n > N$ we know a_n lies in this interval. Thus, we've concluded that for large enough n, that a_n *must be negative*! This is a contradiction, as a_n is a nonnegative sequence. $\square$

Exercise 7.1. If a_n is a convergent $a_n \geq L$ for all n, then $\lim a_n \geq L$. Similarly prove if a_n is a convergent $a_n \leq U$ for all n, then $\lim a_n \leq U$.

This exercise provides the following useful corollary, telling you that if you can bound a sequence, you can bound its limit.

Corollary 7.1 (Inequalities and Convergence). *If a_n is a convergent sequence with $L \leq a_n \leq U$ for all n, then*

$$L \leq \lim a_n \leq U$$

In fact, a kind of converse of this is true as well: if a sequence converges, then we know the limit 'is bounded' (as it exists, as a real number, and those can't be infinite). But this is enough to conclude that the entire sequence is bounded!

Proposition 7.2 (Convergent Sequences are Bounded). *Let s_n be a convergent sequence. Then there exists a B such that $|s_n| < B$ for all $n \in \mathbb{N}$.*

Proof. Let $s_n \to L$ be a convergent sequence. Then we know for any $\epsilon > 0$ eventually the sequence stays within ϵ of L. So for example, choosing $\epsilon = 1$, this means there is some N where for $n > N$ we are assured $|s_n - L| < 1$, or equivalently $-1 < s_n - L < 1$. Adding L,

$$L - 1 < s_n < L + 1$$

Thus, we have both upper and lower bounds for the sequence after N and all we are left to worry about is the finitely many terms before this. For an upper bound on these we can just take the max of $s_1, \ldots, s_N$ and for a lower bound we can take the min.

Thus, to get an overall upper bound, we can take

$$M = \max\{s_1, s_2, \ldots, s_N, L + 1\}$$

and for an overall lower bound we can take

$$m = \min\{s_1, s_2, \ldots, s_N, L - 1\}$$

Then for all n we have $m \leq s_n \leq M$ so the sequence s_n is bounded. □

Theorem 7.1 (The Squeeze Theorem). *Let a_n, b_n and c_n be sequences with $a_n \leq b_n \leq c_n$ for all n. Then if a_n and c_n are convergent, with $\lim a_n = \lim c_n = L$, then b_n is also convergent, and*

$$\lim b_n = L$$

Proof. Choose $\epsilon > 0$. Since both $a_n \to L$, we can choose N_a such that $n > N_a$ implies $|a_n - L| < \epsilon$, and similarly as $c_n \to L$ there's an N_c with $n > N_c$ implying \$|c_n-L|<⊠. Set $N = \max\{N_a, N_b\}$ and note that for any $n > N$ this means $-\epsilon < a_n - L < \epsilon$ and $-\epsilon < c_n - L < \epsilon$. Since $a_n \leq c_n$ by assumption, we can string these inequalities together to get

$$-\epsilon < a_n - L \leq c_n - L < \epsilon$$

But we know more than this: in fact, $a_n \leq b_n \leq c_n$ and subtracting L allows us to *squeeze* this directly into the one given above:

$$-\epsilon < a_n - L \leq b_n - L \leq c_n - L < \epsilon$$

Ignoring the terms with a_n and c_n, this says $-\epsilon < b_n - L < \epsilon$, or $|b_n - L| < \epsilon$. Thus $b_n \to L$ as claimed. □

7.1.1. Example Computations

The squeeze theorem is *incredibly useful* in practice as it allows us to prove the convergence of complicated looking sequences by replacing them with two (hopefully simpler) sequences, an upper and lower bound. To illustrate, let's look back at **?@exr-another-seq-converges**, and re-prove its convergence.

Example 7.1 ($\frac{n}{n^2+1}$ converges to 0.)**.** Since we are trying to converge to zero, we want to bound this sequence above and below by sequences that converge to zero. Since n is always positive, a natural lower bound is the constant sequence $0, 0, 0, \ldots$.

One first thought for an upper bound may be $\frac{n}{n+1}$: its easy to prove that $\frac{n}{n^2+1} < \frac{n}{n+1}$ (as we've made the denominator smaller), and so we have bounded our sequence $0 < a_n < \frac{n}{n+1}$. Unfortunately this does not help us, as $\lim \frac{n}{n+1} = 1$ (Exercise 6.4) so the two bounds do not squeeze a_n to zero!

Another attempt at an upper bound may be $1/n$: we know this goes to zero (Proposition 6.1) and it is also an upper bound:

$$\frac{n}{n^2+1} < \frac{n}{n^2} = \frac{1}{n}$$

Thus since $\lim 0 = 0$ and $\lim \frac{1}{n} = 0$, we can conclude via squeezing that $\lim \frac{n}{n^2+1} = 0$ as well.

This theorem is particularly useful for calculating limits involving functions whose values are difficult to compute. While we haven't formally introduced the sine function yet in this class, we know (and will later confirm) that $-1 \leq \sin(x) \leq 1$ for all $x \in \mathbb{R}$. We can use this to compute many otherwise difficult limits:

Example 7.2 ($s_n = \frac{\sin n}{n}$ converges to 0.). Since $-1 \leq \sin(x) \leq 1$ we know $0 \leq |\sin x| \leq 1$ for all x, and thus

$$0 \leq \frac{\sin n}{n} \leq \frac{1}{n}$$

Since both of these bounding sequences converge to zero, we know the original does as well, by the squeeze theorem.

This sort of estimation can be applied to even quite complicated looking limits:

Example 7.3. Compute the following limit:

$$\lim \left(\frac{n^2 \sin(n^3 - 2n + 1)}{n^3 + n^2 + n + 1} \right)^n$$

Lets begin by estimating as much as we can: we know $|\sin(x)| \leq 1$, so we can see that

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

Next, we see that by shrinking the denominator we can produce yet another over estimate:

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

Bringing back the n^{th} power

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

And, unpacking the definition of absolute value:

$$-\frac{1}{n^n} < \left(\frac{n^2 \sin(n^3 - 2n + 1)}{n^3 + n^2 + n + 1} \right)^n < \frac{1}{n^n}$$

It now suffices to prove that $1/n^n$ converges to zero, as we ve squeezed our sequence with it. But this is easiest to do with *another squeeze*: namely, since $n^n > 2^n$ we see $0 < 1/n^n < 1/2^n$, and we already proved that $1/2^n \to 0$, so we're done!

$$\lim \left(\frac{n^2 \sin(n^3 - 2n + 1)}{n^3 + n^2 + n + 1} \right)^n = 0$$

Exercise 7.2. Use the squeeze theorem to prove that

$$\lim \left(\frac{n^3 - 2 - \frac{1}{n^3}}{3n^3 + 5} \right)^{2n+7} = 0$$

A nice corollary of the squeeze theorem tells us when a sequence converges by estimating its difference from the proposed limit:

Exercise 7.3. Let a_n be a sequence, and L be a real number. If there exists a sequence α_n where $|a_n - L| \leq \alpha_n$ for all n, and $\alpha_n \to 0$, then $\lim a_n = L$.

This is useful as unpacking the definition of absolute value (Definition 2.5), a sequence α_n with

$$-\alpha_n \leq a_n - L \leq \alpha_n$$

can be thought of as giving "error bounds" on the difference of a_n from L. In this language, the proposition says if we can bound the error between a_n and L by a sequence going to zero, then a_n must actually go to L.

7.2. Limits and Field Operations

Just like inequalities, the field operations themselves play nicely with limits.

Theorem 7.2 (Constant Multiples)**.** *Let s_n be a convergent sequence, and k a real number. Then the sequence ks_n is convergent, and*

$$\lim ks_n = k \lim s_n$$

Proof. We distinguish two cases, depending on k. If $k = 0$, then ks_n is just the constant sequence $0, 0, 0 \ldots$ and $k \lim s_n = 0$ as well, so the theorem is true.

If $k \neq 0$, we proceed as follows. Denote the limit of s_n by L, and let $\epsilon > 0$. Choose N such that $n > N$ implies $|s_n - L| < \frac{\epsilon}{|k|}$ (we can do so, as $s_n \to L$). Now, for this same value of N, choose arbitrary $n > N$ and consider the difference $|ks_n - kL|$:

$$|ks_n - kL| = |k(s_n - L)| = |k||s_n - L| < |k|\frac{\epsilon}{|k|} = \epsilon$$

Thus, $ks_n \to kL$ as claimed! $\square$

To do a similar calculation for the sum of sequences requires an $\epsilon/2$ type argument:

Theorem 7.3 (Limit of a Sum)**.** *Let s_n, t_n be convergent sequences. Then the sequence of term-wise sums $s_n + t_n$ is convergent, with*

$$\lim(s_n + t_n) = \lim s_n + \lim t_n$$

This is a great example of a classic proof technique known as *an $\epsilon/2$ argument* that we will use many times.

Proof. Let $\epsilon > 0$ be arbitrary. Since we know that both a_n and b_n converge, we can provide notation for their limits - specifically, $\lim a_n = A$, and $\lim b_n = B$. Since $a_n \to A$, there exists some N_a such that for any $n > N_a$, $|a_n - A| < \frac{\epsilon}{2}$. Similarly, since $b_n \to B$, there exists some N_b so that for any $n > N_b$, $|b_n - B| < \frac{\epsilon}{2}$. Lets set N equal to the maximum of the set $\{N_a, N_b\}$. This means that if $n > N$,

$$|a_n - A| + |b_n - B| < \epsilon$$

According to the triangle inequality, we also know that

$$|(a_n - A) + (b_n - B)| \leq |a_n - A| + |b_n - B|$$

so by combining the previous two inequalities we know that

$$\begin{aligned} &|(a_n - A) + (b_n - B)| < \epsilon \\ \implies &|(a_n + b_n) - (A + B)| < \epsilon \end{aligned}$$

This is equivalent to the convergence definition saying that $\lim (a_n + b_n) = A + B = \lim a_n + \lim b_n$. □

Corollary 7.2 (Limit of a Difference)**.** *Let s_n, t_n be convergent sequences. Then $s_n - t_n$ is convergent and*

$$\lim(s_n - t_n) = \lim s_n - \lim t_n$$

Proof. Rewrite $s_n - t_n$ as $s_n + (-t_n)$. Note that since t_n is convergent we know the multiple $-t_n$ is convergent, with $\lim t_n = t$ implying $\lim(-t_n) = -t$ by Theorem 7.2. Now using the limit of sums (Theorem 7.3) we see since s_n and $-t_n$ are convergent so is $s_n + (-t_n)$, and

$$\lim(s_n + (-t_n)) = \lim s_n + \lim(-t_n) = \lim s_n + (-1)\lim t_n = \lim s_n - \lim t_n$$

□

The case of products is a little more annoying to prove, but the end result is the same - the limit of a product is the product of the limits.

Theorem 7.4 (Limit of a Product)**.** *Let s_n, t_n be convergent sequences. Then the sequence of term-wise products s_nt_n is convergent, with*

$$\lim(s_nt_n) = (\lim s_n)(\lim t_n)$$

Sketch. Let $s_n \to S$ and $t_n \to T$ be two convergent sequences and choose $\epsilon > 0$. We wish to find an N beyond which we know s_nt_n lies within ϵ of \$ST.

To start, we consider the difference $|s_nt_n - ST|$ and we add zero in a clever way:

$$|s_nt_n - ST| = |s_nt_n - s_nT + s_nT - ST| = |(s_nt_n - s_nT) + (s_nT - ST)|$$

applying the triangle inequality we can break this apart

$$|s_nt_n - ST| \leq |s_nt_n - s_nT| + |s_nT - ST| = |s_n||t_n - T| + |s_n - S||T|$$

The second term here is easy to bound: if $T = 0$ then its just literally zero, and if $T \neq 0$ then we can make it as small as we want: we know $s_n \to S$ so we can make $|s_n - S|$ smaller than anything we need (like ϵ/T, or even $\epsilon/2T$ if necessary).

For the first term we see it includes a term of the form $|t_n - T|$ which we know we can make as small as we need to by choosing sufficiently large N. But its being multiplied by $|s_n|$ and we need to make sure the whole thing can be made small, so we should worry about what if $|s_n|$ is getting really big? But this isn't actually a worry - we know s_n is convergent, so its bounded, so there is some B where $|s_n| < B$ for all n. Now we can make $|t_n - T|$ as small as we like, (say, smaller than ϵ/B or $\epsilon/2B$ or whatever we need).

Since each of these terms can be made small as we need individually, choosing large enough n's we can make them both simultaneously small, so the whole difference $|s_nt_n - ST|$ is small (less than ϵ) which proves convergence. ☐

Exercise 7.4. Write the sketch of an argument above in the right order, as a formal proof.

Corollary 7.3. *If p is a positive integer then*

$$\lim \frac{1}{n^p} = 0$$

Hint: Induction on the power p

The next natural case to consider after sums and differences and products is quotients. We begin by considering the limit of a reciprocal:

Proposition 7.3 (Limit of a Reciprocal)**.** *Let s_n be a convergent nonzero sequence wtih a nonzero limit. Then the sequence $1/s_n$ of reciprocals is convergent, with*

$$\lim \frac{1}{s_n} = \frac{1}{\lim s_n}$$

Sketch. For any $\epsilon > 0$, want to show when n is very large, we can make

$$\left|\frac{1}{s_n} - \frac{1}{s}\right| < \epsilon$$

We can get a common denominator and rewrite this as

$$\left|\frac{1}{s_n} - \frac{1}{s}\right| = \frac{|s - s_n|}{|ss_n|}$$

Since s_n is not converging to zero, we should be able to bound it away from zero: that is, find some m such that $|s_n| > m$ for all $n \in \mathbb{N}$ (we'll have to prove we can actually do this). Given such an m we see the denominator $|ss_n| > m|s|$, and so

$$\left|\frac{1}{s_n} - \frac{1}{s}\right| < \frac{|s_n - s|}{m|s|}$$

We want this less than ϵ so all we need to do is choose N big enough that $|s_n - s|$ is less than $\epsilon m|s|$ and we're good. ☐

Exercise 7.5. Turn the sketch argument for $\lim \frac{1}{s_n} = \frac{1}{s_n}$ in Proposition 7.3 into a formal proof.

From here, its quick work to understand the limit of a general quotient.

Theorem 7.5 (Limit of a Quotient)**.** *Let s_n, t_n be convergent sequences, with $t_n \neq 0$ and $\lim t_n \neq 0$. Then the sequence s_n/t_n of quotients is convergent, with*

$$\lim \frac{s_n}{t_n} = \frac{\lim s_n}{\lim t_n}$$

Proof. Since t_n converges to a nonzero limit, by Proposition 7.3 we know that $1/t_n$ converges, with limit $1/\lim t_n$. Now, we can use Theorem 7.4 for the product $s_n \cdot \frac{1}{t_n}$:

$$\begin{aligned}\lim \frac{s_n}{t_n} &= \lim s_n \cdot \frac{1}{t_n} = (\lim s_n)\left(\lim \frac{1}{t_n}\right) \\ &= \lim s_n \frac{1}{\lim t_n} = \frac{\lim s_n}{\lim t_n}\end{aligned}$$

☐

Finally we look at square roots. We have already proven in Theorem 4.9 that nonnegative numbers have square roots, and so given a nonnegative sequence s_n we can consider the sequence $\sqrt{s_n}$ of its roots. Below we see that the limit concept respects roots just as it does the other field operations:

Theorem 7.6 (Root of Convergent Sequence)**.** *Let $s_n > 0$ be a convergent sequence, and $\sqrt{s_n}$ its sequence of square roots. Then $\sqrt{s_n}$ is convergent, with*

$$\lim \sqrt{s_n} = \sqrt{\lim s_n}$$

Sketch. Assume $s_n \to s$, and fix $\epsilon > 0$. We seek an N where $n > N$ implies $|\sqrt{s_n} - \sqrt{s}| < \epsilon$. This looks hard: because the fact we know is about $s_n - s$ and the fact we need is about $\sqrt{s_n} - \sqrt{s}$.
But what if we multiply and divide by $\sqrt{s_n} + \sqrt{s}$ so we can simplify using the difference of squares?

$$|\sqrt{s_n} - \sqrt{s}| \frac{\sqrt{s_n} + \sqrt{s}}{\sqrt{s_n} + \sqrt{s}} = \frac{|s_n - s|}{\sqrt{s_n} + \sqrt{s}}$$

This has the quantity $|s_n - s|$ that we know about in it! We know we can make this as small as we like by the assumption $s_n \to s$, so as long as the denominator does not go to zero, we can make this happen!

□

Formal. Let s_n be a positive sequence with $s_n \to s$ and assume $s \neq 0$ (we leave that case for the exercise below). Let $\epsilon > 0$, and choose N such that if $n > N$ we have $|s_n - s| < \epsilon\sqrt{s}$.

Now for any n, rationalizing the numerator we see

$$|\sqrt{s_n} - \sqrt{s}| = \frac{|s_n - s|}{\sqrt{s_n} + \sqrt{s}} < \frac{|s_n - s|}{\sqrt{s}}$$

Where the last inequality comes from the fact that $\sqrt{s_n} > 0$ by definition, so $\sqrt{s} + \sqrt{s_n} > \sqrt{s}$. When $n > N$ we can use the hypothesis that $s_n \to s$ to see

$$|\sqrt{s_n} - \sqrt{s}| < \frac{|s_n - s|}{\sqrt{s}} = \frac{\epsilon\sqrt{s}}{\sqrt{s}} = \epsilon$$

And so, $\sqrt{s_n}$ is convergent, with limit $\sqrt{s}$. □

Exercise 7.6. Prove that if $s_n \to 0$ is a sequence of nonnegative numbers, that the sequence of roots also converges to zero $\sqrt{s_n} \to 0$.

Hint: you don't need to rationalize the numerator or do fancy algebra like above

Together this suite of results provides an effective means of calculating limits from simpler pieces. They are often referred to together as *the limit theorems*

Theorem 7.7 (The Limit Theorems). *Let a_n and b_n be any two convergent sequences, and $k \in \mathbb{R}$ a constant. Then*

$$\lim ka_n = k \lim a_n$$

$$\lim(a_n \pm b_n) = (\lim a_n) \pm (\lim b_n)$$

$$\lim a_n b_n = (\lim a_n)(\lim b_n)$$

If $b_n \neq 0$ and $\lim b_n \neq 0$,

$$\lim \frac{a_n}{b_n} = \frac{\lim a_n}{\lim b_n}$$

And, if $a_n \geq 0$, then $\sqrt{a_n}$ is convergent, with

$$\lim \sqrt{a_n} = \sqrt{\lim a_n}$$

7.2.1. Example Computations

Example 7.4. Compute the limit of the following sequence s_n:

$$s_n = \frac{3n^3 + \frac{n^6-2}{n^2+5}}{n^3 - n^2 + 1}$$

Example 7.5. Compute the limit of the sequence s_n

$$s_n = \sqrt{\frac{1}{2^n} + \sqrt{\frac{n^2-1}{n^2-n+1}}}$$

7.3. Applications

7.3.1. Babylon and $\sqrt{2}$

We know that $\sqrt{2}$ exists as a real number (Theorem 4.7), and we know that the babylonian procedure produces excellent rational approximations to this value (Exercise 0.5), in the precise sense that the numerator squares to just one more than twice the square of the denominator.

Now we finally have enough tools to combine these facts, and prove that the babylonian procedure really does limit to $\sqrt{2}$.

Theorem 7.8. *Let $s_n = \frac{p_n}{q_n}$ be a sequence of rational numbers where both $p_n, q_n \to \infty$ and for each $p_n^2 = 2q_n^2 - 1$. Then $s_n \to \sqrt{2}$.*

Proof. We compute the limit of the sequence s_n^2. Using that $p_n^2 = 2q_n^2 + 1$ we can replace the numerator and do algebra to see

$$s_n^2 = \frac{p_n^2}{q_n^2} = \frac{2q_n^2 + 1}{q_n^2} = 2 + \frac{1}{q_n^2}.$$

Now, as by assumption $q_n \to \infty$ we have that $q_n^2 = q_n q_n$ also diverges to infinity (Exercise 7.11), and so its reciprocal converges to 0 (**?@prp-diverge-to-infty-equliv-converge-to-zero**). Thus, using the limit theorems for sums,

$$\lim \frac{p_n^2}{q_n^2} = \lim \left(2 - \frac{1}{q_n^2}\right) = 2 - \lim \frac{1}{q_n^2} = 2$$

That is, the limit of the squares approaches 2. Now we apply Theorem 7.6 to this sequence s_n^2, and conclude that

- $s_n = \sqrt{s_n^2}$ converges.
- $\lim s_n = \lim \sqrt{s_n^2} = \sqrt{\lim s_n^2} = \sqrt{2}$

□

This provides a *rigorous justification* of the babylonian's assumption that if you are patient, and compute more and more terms of this sequence, you will *always* get better and better approximations of the square root of 2.

Exercise 7.7. Build a sequence that converges to $\sqrt{n}$ by following the babylonian procedure, starting with a rectangle of area n.

7.3.2. Rational and Irrational Sequences

Combining the squeeze theorem and limit theorems with the density of the (ir)rationals allows us to prove the existence of certain sequences that will prove quite useful:

Theorem 7.9. *For every $x \in \mathbb{R}$ there exists a sequence r_n of rational numbers with $r_n \to x$.*

Proof. Let $x \in \mathbb{R}$ be arbirary, and consider the sequence $x + \frac{1}{n}$. Because the constant sequence $x, x, x \dots$ and the sequence $1/n$ are convergent, by the limit theorem for sums we know $x + \frac{1}{n}$ is convergent and

$$\lim \left(x + \frac{1}{n} \right) = x + \lim \frac{1}{n} = x$$

Now for each $n \in \mathbb{N}$, by the density of the rationals we can find a rational number r_n with $x < r_n < x + \frac{1}{n}$. This defines a sequence of rational numbers squeezed between x and $x + \frac{1}{n}$: thus, by the squeeze theorem we hav

$$x < r_n < x + \frac{1}{n} \implies \lim r_n = x$$

□

Through a similar argument using Exercise 4.5 we find the existence of a sequence of irrational numbers converging to any real number.

Exercise 7.8. For every $x \in \mathbb{R}$ there exists a sequence y_n of irrationals with $y_n \to x$.

7.4. Problems

7.4.1. ⋆ Infinity

Given the formal defintion of *divergence to infinity* as meaning *eventually gets larger than any fixed number*, we can formulate analogs of the limit theorems for such divergent sequences. We will not need any of these in the main text but it is good practice to attempt their proofs:

Exercise 7.9. If $s_n \to \infty$ and $k > 0$ then $ks_n \to \infty$.

Exercise 7.10. If t_n diverges to infinity, and s_n either converges, or also diverges to infinity, then $s_n + t_n \to \infty$.

Exercise 7.11. If t_n diverges to infinity, and s_n either converges, or also diverges to infinity, then $s_n t_n \to \infty$.

Note that there is *not* an analog of the division theorem: if $s_n \to \infty$ and $t_n \to \infty$, with only this knowledge we can learn nothing about the quotient s_n/t_n.

Exercise 7.12. Give examples of sequences $s_n, t_n \to \infty$ where

$$\lim \frac{s_n}{t_n} = 0$$

$$\lim \frac{s_n}{t_n} = 2$$

$$\lim \frac{s_n}{t_n} = \infty$$

These limit laws are the precise statement behind the "rules" often seen in a calculus course, where students may write $2+\infty = \infty$, $\infty+\infty = \infty$, or $\infty \cdot \infty = \infty$, but they may not write ∞/∞. (If you are looking at this last case and thinking *l'Hospital*, we'll get there in **?@thm-Lhospital**!)

8. Monotone Convergence

Highlights of this Chapter: We prove the monotone convergence theorem, which is our first theorem that tells us a sequence converges, without having to first know its limiting value. We show how to use this theorem to find the limit of various recursively defined sequences, including two important examples.

- We prove the infinite sequence of roots $\sqrt{1+\sqrt{1+\sqrt{1+\cdots}}}$ converges to the golden ratio.
- We prove the sequence $\left(1+\frac{1}{n}\right)^n$ converges to the number $e = 2.71828\ldots$
- We begin a treatment of irrational exponents, by looking at the limit of sequences with rational exponents.

8.1. Monotone Convergence

The motivation for inventing sequences is to work with infinite processes, where we have a precise description of each finite stage, but cannot directly grasp the "completed" state "at infinity". In the first section of this chapter we computed a few specific limits, and then in the second we showed *how to find new, more complicated limits* given that you know the value of some simpler ones via algebra.

But what we haven't done, since our original motivating discussion with the nested intervals theorem, is actually return to the part of the theory we are most interested in: rigorously assuring that certain sequences converge, *without* knowing the value of their limit ahead of time! The most useful theorem in this direction is the *monotone convergence theorem*, which deals with monotone sequences.

Definition 8.1 (Monotone Sequences)**.** A sequence s_n is monotone increasing (or more precisely, monotone *non-decreasing*) if

$$m \le n \implies s_m \le s_n$$

A sequence is monotone decreasing (non-increasing) if

$$m \le n \implies s_m \ge s_n$$

Note: constant sequences are monotone: both monotone increasing and monotone decreasing.

The original inspiration for a monotone sequence is the sequence of upper bounds or lower bounds from a collection of nested intervals: as the intervals get smaller, the lower bounds monotonically increase, and the upper bounds monotonically decrease. The *Monotone convergence theorem* guanatees that such sequences always converge. Its proof is below, but could actually be extracted directly from Theorem 3.3.

Theorem 8.1 (Monotone Convergence). *Let s_n be a monotone bounded sequence. Then s_n is a convergent sequence.*

Proof. Here we consider the case that s_n is monotone increasing, and leave the decreasing case as an exercise. Let $S = \{s_n \mid n \in \mathbb{N}\}$. Then S is nonempty, and is bounded above (by any upper bound for the sequence s_n, which we assumed is bounded). Thus by completeness, it has a supremum $s = \sup S$.

We claim that s_n is actually a convergent sequence, which limits to s_n. To prove this, choose $\epsilon > 0$, and note that as s is the *least* upper bound, $s - \epsilon$ is not an upper bound for S, so there must be some N where $s_N > s - \epsilon$. But s_n is monotone increasing, so if $n > N$ it follows that $s_n > s_N$. Recalling that for all n we know $s_n \leq s$ (since s is an upper bound), we have found some N where for all $n > N$ we know $s - \epsilon < s_n < s$. This further implies $|s_n - s| < \epsilon$, which is exactly the definition of convergence! Thus

$$s_n \to s$$

So it is a convergent sequence, as claimed. □

Though straightforward to prove, this theorem has tons of applications, as it assures us that many of the difficult to describe recursively defined sequences that show up in practice actually do converge, and thus we may rigorously reason about their limits. We will give several interesting ones below.

8.2. Application: Defining Irrational Powers

We have already defined rational powers of a number in terms of iterated multiplication/division, and the extraction of roots: but how does one define a real numbered power? We can use sequences to do this! To motivate this, let's consider the example of defining 2^π. We can write π as the limit of a sequence of rational numbers, for instance

$$3,\ 3.1,\ 3.14,\ 3.141,\ 3.1415\ \ldots$$

And since rational exponents make sense, from this we can produce a sequence of exponentials

$$2^3,\ 2^{\frac{31}{10}},\ 2^{\frac{314}{100}},\ 2^{\frac{3141}{1000}},\ 2^{\frac{31415}{10000}},\ldots$$

Then we may ask if this sequence has a limit: if it does, it's natural to try and define this value two to the power of pi. To make sure this makes sense, we need to check several potential worries:

- Does this sequence converge?
- Does the limit depend on the particular sequence chosen?

For example if you tried to define $3^{\sqrt{2}}$ using the babylonian sequence for $\sqrt{2}$, and your friend tried to use the sequence coming from the partial fraction, you'd better get the same number if this is a reasonable thing to define! Because we are in the section on monotone convergence, we will restrict ourselves at the moment to *monotone sequences* though we will see later we can dispense with this if desired.

Proposition 8.1. *If $r_n \to x$ is a monotone sequence of rational numbers converging to x, and $a > 0$ then the sequence a^{r_n} converges.*

Proof. Recall for a fixed positive base a, exponentiation by rational numbers is monotone increasing, so $r < s$ implies $a^r < a^s$.
Thus, given a monotone sequence r_n, the exponentiated sequence a^{r_n} remains monotone (for monotone increasing we see $r_n \leq r_{n+1} \implies a^{r_n} \leq a^{r_{n+1}}$ and the equalities are reversed if r_n is monotone decreasing).

Now that we know a^{r_n} is monotone, we only need to see its bounded to apply Monotone Convergence. Again we have two cases, and will deal here with the monotone increasing case. As $r_n \to x$ and x is a real number, there must be some natural number $N > x$. Thus, N is greater than r_n for all n, and so a^N is greater than a^{r_n}: our sequence is bounded above by a^N. Thus all the hypotheses of monotone convergence are satisfied, and $\lim a^{r_n}$ exists. □

Now that we know such sequences make sense, we wish to clear up any potential ambiguity, and show that if two different sequences both converge to x, the value we attempt to assign to a^x as a limit is the same for each. As a lemma in this direction, we look at sequences converging to zero.

Exercise 8.1. Let r_n be any sequence of rationals converging to zero. Then for any $a > 0$ we have

$$\lim a^{r_n} = 1$$

Corollary 8.1. *If r_n, s_n are two monotone sequences of rationals each converging to x, then*

$$\lim a^{r_n} = \lim a^{s_n}$$

for any $a > 0$.

Proof. Let $z_n = r_n - s_n$, so that $z_n \to 0$. Because r_n and s_n are monotone, we know $\lim a^{r_n}$ and $\lim a^{s_n}$ exist. And by the exercise above, we have $a^{z_n} \to 1$. Noting that $r_n = s_n + z_n$ and that the laws of exponents apply for rational exponents, we have

$$a^{r_n} = a^{s_n + z_n} = a^{s_n} a^{z_n}$$

But as all quantities in question converge we can use the limit theorems to compute:

$$\begin{aligned}
\lim a^{r_n} &= \lim a^{s_n + z_n} \\
&= \lim a^{s_n} a^{z_n} \\
&= (\lim a^{s_n})(\lim a^{z_n}) \\
&= \lim a^{s_n}
\end{aligned}$$

□

Thus, we can unambiguously define the value of a^x as the limit of *any monotone sequence a^{r_n}* without specifying the sequence itself.

Definition 8.2. ## Irrational Powers Let $a > 0$ and $x \in \mathbb{R}$. Then we define a^x as a limit

$$a^x = \lim a^{r_n}$$

For r_n any monotone sequence of rational numbers converging to x.

Perhaps upon reading this definition to yourself you wonder, is the restriction to monotone sequences important, or just an artifact of our currently limited toolset? Once we build more tools we will see the latter is the case; you will show on homework that arbitrary convergent sequences $r_n \to x$ can be used to unambiguously define a^x.

8.3. Applicaiton: Recursive Sequences

The monotone convergence theorem is particularly adept to working with *recursive sequences*, as one may aim to prove such a sequence is monotone and bounded by induction. This guarantees the limit exists, at which point we can rigorously give that limiting value a name, and use limit theorems to find its value.

8.3.1. The Golden Ratio

Consider the recursive sequence defined by $s_{n+1} = \sqrt{1+s_n}$ starting from $s_0 = 1$:

$$s_0 = 1$$
$$s_1 = \sqrt{1+\sqrt{1}}$$
$$s_2 = \sqrt{1+\sqrt{1+\sqrt{1}}}$$
$$\cdots$$

Because such sequences follow a regular pattern, we can use a shorthand notation with ellipsis for their terms. For example, in the original sequence above, writing the first couple steps of the pattern followed by an ellipsis

$$\sqrt{1+\sqrt{1+\sqrt{1+\ldots}}}$$

we take to mean the sequence of terms s_n where $s_{n+1} = \sqrt{1+s_n}$ itself. Thus, writing $\lim\sqrt{1+\sqrt{1+\cdots}}$ means the limit of this sequence, implicitly defined by this infinite expression.

Exercise 8.2. Here are some other infinite expressions defined by recursive sequences: can you give the recursion relation they satisfy?

$$\sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\cdots}}}$$

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

$$\cos(\cos(\cos(\cdots\cos(5)\cdots)))$$

In all of these sequences it is not clear *at all* how to find their limit value from scratch, or how we could possibly apply any of the limit theorems about field axioms and inequalities. But, recursive sequences are set up for using induction, and monotone convergence! We can build a sort of recipe for dealing with them:

Recursive Sequence Operation Manual:

- Prove its bounded, by induction.
- Prove its monotone, by induction.
- Use Monotone convergence to conclude its convergent.

- Use the recursive definition, and the limit theorems, to find an equation satisfied by the limit.
- Solve that equation, to find the limit.

A beautiful and interesting example of this operations manual is carried out below:

Proposition 8.2. *The sequence* $\sqrt{1+\sqrt{1+\cdots}}$ *converges to the golden ratio.s*

Proof. The infinite expression $\sqrt{1+\sqrt{1+\cdots}}$ defines the recursive sequence $s_{n+1} = \sqrt{1+s_n}$ with $s_1 = 1$.

Step 1: s_n is monotone increasing, by induction First we show that $s_2 > s_1$. Using the formula, $s_2 = \sqrt{1+\sqrt{1}} = \sqrt{2}$, which is larger than $s_1 = 1$. Next, we assume for induction $s_n > s_{n-1}$ and we use this to prove that $s_{n+1} > s_n$. Starting from our induction hypothesis, we add one to both sides yielding $1+s_n > 1+s_{n-1}$ and then we take the square root (which preserves the inequality, by Proposition 2.5) to get

$$\sqrt{1+s_n} > \sqrt{1+s_{n-1}}$$

But now, we simply note that the term on the left is the definition of s_{n+1} and the term on the right is the definition of s_n. Thus we have $s_{n+1} > s_n$ as claimed, and our induction proof works for all n.

Step 2: s_n is bounded, by induction It is hard to guess an upper bound for s_n without doing a little calculation, but plugging the first few terms into a calculator shows them to be less than 2, so we might try to prove $\forall n\, s_n < 2$. The base case is immediate as $s_1 = 1 < 2$, so assume for induction $s_n < 2$. Then $1+s_n < 3$ and so $\sqrt{1+s_n} < \sqrt{1+2} = \sqrt{3}$, and $\sqrt{3} < 2$ (as $3 < 2^2 = 4$) so our induction has worked, and the entire sequence is bounded above by 2.

Conclusion: s_n converges! We have proven the sequence s_n is both monotone increasing and bounded above by 2. Thus the monotone convergence theorem assures us that there exists some L with $s_n \to L$. It only remains to figure out what number this is!

Step 3: The Limit Theorems Because truncating the beginning of a sequence does not change its limit, we see that $\lim s_n = \lim s_{n+1} = L$. But applying the limit theorems to $s_{n+1} = \sqrt{1+s_n}$, we see that as $s_n \to L$, it follows that $1+s_n \to 1+L$ and thus that $\sqrt{1+s_n} \to \sqrt{1+L}$. This gives us an equation that L must satisfy!

$$\sqrt{1+L} = L$$

Simplifying this becomes $1+L = L^2$, which has solutions $(1 \pm \sqrt{5})/2$. This argument only tells us so far that one of these numbers must be our limit L: to figure out which we need to bring in more information. Noticing that only one of the two is positive, and all the terms of our sequence are positive singles it out:

$$\sqrt{1+\sqrt{1+\sqrt{1+\cdots}}} = \frac{1+\sqrt{5}}{2} \approx 1.618\ldots$$

This number is known as the *golden ratio*. □

Example 8.1. The final step of the proof above suggests a way one might find a recursive sequence to use as a calculating tool: if we started with the golden ratio

$$\phi = \frac{1+\sqrt{5}}{2}$$

we could observe that ϕ solves the quadratic equation $1 + L = L^2$, and hence $L = \sqrt{1+L}$. This sets up a recursive sequence, as we can plug this relation into itself over and over:

$$L = \sqrt{1+L} = \sqrt{1+\sqrt{1+L}} = \sqrt{1+\sqrt{1+\sqrt{1+\cdots}}}$$

Which immediately suggests the recursion $s_{n+1} = \sqrt{1+s_n}$ as a candidate for generating a sequence that would solve the original equation.

Exercise 8.3. Find a recursive sequence whose limit is the positive real root of $x^2 - 2x - 5$. Then prove that your proposed sequence actually converges to this value.

Exercise 8.4. What number is this?

$$\sqrt{1-2\sqrt{1-2\sqrt{1-2\sqrt{\cdots}}}}$$

8.3.2. $\sqrt{2}$

Recall the babylonian sequence converging to $\sqrt{2}$ was recursively defined, starting from the side length $x_1 = 2$ of a 2×1 rectangle and replacing it with the average of the two sides (a rectangle closer to a square). As a formula this is the process $x \mapsto (x + \frac{2}{x})/2$ from which we product a recursive sequence

$$x_{n+1} = \frac{x_n + \frac{2}{x_n}}{2}$$

Our goal here is to provide a proof of convergence using the strategy laid out above. First, we aim to show that x_n is monotone decreasing. To make the algebra simpler, we first give a little lemma simplfying the condition:

Exercise 8.5. Let x_n be a sequence of positive numbers satisfying the babylonian recurrence relation.
Show that $x_{n+1} < x_n$ if and only if $2 < x_n^2$.

Proposition 8.3. *Starting from $x_0 = 2$, the recursive procedure for x_n defines a monotone decreasing sequence.*

Proof. We proceed by induction. For the base case we compute $x_1 = \frac{2+\frac{2}{2}}{2} = \frac{3}{2}$ so $x_1 < x_0$ as required. For the inductive step we assume $x_n < x_{n-1}$ and aim to show $x_{n+1} < x_n$. By the exercise above, this is equivalent to assuming that $2 < x_{n-1}^2$ and using this to prove that $2 < x_n^2$.

Writing out the recursive definition of x_n we see

$$\begin{aligned}
x_n^2 &= \left(\frac{x_{n-1} + \frac{2}{x_{n-1}}}{2}\right)^2 \\
&= \frac{x_{n-1}^2 + 4 + \frac{4}{x_{n-1}^2}}{4} \\
&= \left(\frac{x_{n-1}}{2}\right)^2 + 1 + \left(\frac{2}{x_{n-1}}\right)^2
\end{aligned}$$

The first term is > 1 by the inductive hypothesis, and so the first two terms sum to greater than 2. Since the last term is a square its *positive* and can't possibly make things smaller, so the entire thing sums to something strictly larger than 2, as required. Thus x_n is monotone decreasing. □

The next step in our process is to prove the sequence is bounded.

Exercise 8.6. Prove that the sequence with $x_0 = 2$ satisfying the babylonian recurrence relation is bounded below by 1.

Now, since x_n is monotone and bounded it converges to some limiting value L. Since truncating the beginning of a sequence has no effect on the eventual limit, we know $\lim x_{n+1} = \lim x_n = L$, but expanding the first term's recursive definition, this implies that

$$\lim \frac{x_n + \frac{2}{x_n}}{2} = L$$

Since we know $x_n \to L$ and $L \neq 0$ (since its bounded below by 1), we can use the limit laws to compute

$$\lim \frac{x_n + \frac{2}{x_n}}{2} = \frac{\lim x_n + \frac{2}{\lim x_n}}{2} = \frac{L + \frac{2}{L}}{2}$$

Thus, whatever the limiting value L is it must satisfy the equation

$$\frac{L + \frac{2}{L}}{2} = L$$

Multiplying by $2L$ to clear denominators we see

$$L^2 + 2 = 2L^2 \implies 2 = L^2$$

We know L is positive, and the only positive solution to this equation is $L = \sqrt{2}$. So we have another proof of the convergence of the babylonian procedure!

8.4. ⋆ The Number *e*

In this section we aim to study, and prove the convergence of the following sequence of numbers

$$\left(1 + \frac{1}{n}\right)^n$$

We will later see that the limit of this sequence is the number *e* (indeed, many authors take this sequence itself as *the definition of e* as it is perhaps the first natural looking sequence limiting to this special value. We will instead define *e* in terms of *exponential functions* to come, and then later show its value coincides with this limit).

We begin by proving a_n is monotone as a prelude to applying monotone convergence.

Example 8.2. The sequence $a_n = \left(\frac{n+1}{n}\right)^n$ is monotone increasing.

Proof. To show a_n is increasing we will show that the ratio $\frac{a_n}{a_{n-1}}$ is greater than 1. Simplifying,

$$\frac{a_n}{a_{n-1}} = \frac{\left(\frac{n+1}{n}\right)^n}{\left(\frac{n}{n-1}\right)^{n-1}} = \left(\frac{n+1}{n}\right)^n \left(\frac{n-1}{n}\right)^{n-1}$$

Multiplying by $\frac{n-1}{n}$ and its inverse we can make the powers on each of these terms the same, and combine them:

$$= \left(\frac{n+1}{n}\right)^n \left(\frac{n-1}{n}\right)^n \frac{n}{n-1} = \left(\frac{n^2-1}{n^2}\right)^n \frac{n}{n-1}$$

Simplifying what is in parentheses, we notice that we are actually in a perfect situation to apply Bernoulli's Inequality (Exercise 2.6) to help us estimate this term. Recall this says that if r is any number such that $1+r$ is positive, $(1+r)^n \geq 1+nr$. When $n \geq 2$ we can apply this to $r = -\frac{1}{n^2}$, yielding

$$\frac{a_n}{a_{n-1}} = \left(1 - \frac{1}{n^2}\right)^n \frac{n}{n-1} \geq \left(1 - \frac{n}{n^2}\right) \frac{n}{n-1}$$
$$= \frac{n-1}{n} \frac{n}{n-1} = 1$$

Thus $\frac{a_n}{a_{n-1}} \geq 1$, so $a_n \geq a_{n-1}$ and the sequence is monotone increasing for all n, as claimed. □

Next we need to show that a_n is bounded above. Computing terms numerically, it *seems* that a_n is bounded above by 3, but of course no amount of computation can substitute for a proof. And after a bit of trying, it seems hard to prove *directly* that it actually is bounded above.

So instead, we will employ a bit of an ingenious trick. We will study a second sequence, which appears very similar to the first:

$$b_n = \left(\frac{n+1}{n}\right)^{n+1}$$

Indeed, this is just our sequence a_n multiplied by one extra factor of $\frac{n+1}{n}$! But this extra factor changes its behavior a bit: computing the first few terms, we see that it appears to be decreasing:

$$b_1 = (1+1)^2 = 4,\ b_2 = \left(1+\frac{1}{2}\right)^3 = \frac{27}{8} = 3.375,\ b_3 = \left(1+\frac{1}{3}\right)^4 \approx 3.1604$$

Indeed, a proof that its decreasing can be constructed following an identical strategy to a_n in Example 8.2.

Exercise 8.7. The sequence $b_n = \left(\frac{n+1}{n}\right)^{n+1}$ is monotone decreasing.

Now that we understand the behavior of b_n we can use it to prove that a_n is bounded above:

Corollary 8.2. *The sequence $a_n = \left(1 + \frac{1}{n}\right)^n$ is convergent*

Proof. Note that the sequence b_n and a_n are related by

$$b_n = \left(\frac{n+1}{n}\right)^{n+1} = a_n \left(\frac{n+1}{n}\right)$$

Since $\frac{n+1}{n} > 1$ we see that $b_n > a_n$ for all n. But b_n is decreasing, so $b_n \leq b_1 = 2^2 = 4$, and so a_n is bounded above by 4. □

Note that we can also readily see that b_n is itself convergent (though we did not actually need that fact for our analysis of a_n): we proved its monotone decreasing, and its a sequence of positive terms - so its trivially bounded below by zero!

We can also see that a_n and b_n have the same limit, using the limit theorems. Since $\frac{1}{n} \to 0$, we know that $1 + \frac{1}{n} \to 1$, and hence that

$$\begin{aligned}
\lim b_n &= \lim \left[a_n \left(\frac{n+1}{n}\right)\right] \\
&= (\lim a_n) \cdot \left(\lim \frac{n+1}{n}\right) \\
&= \lim a_n
\end{aligned}$$

As mentioned previously, we will later see that this limit is the number called e. But believing for a moment that we *should* be interested in this particular limit, having the two sequences a_n and b_n lying around actually proves quite practically useful for estimating its value.

Since $\lim a_n = e = \lim b_n$ and $a_n < b_n$ for all n, we see that the number e is contained in the interval $I_n = [a_n, b_n]$, and hence is the limit of the nested intervals:

Corollary 8.3.

$$\{e\} = \bigcap_{n \geq 1} \left[\left(1 + \frac{1}{n}\right)^n, \left(1 + \frac{1}{n}\right)^{n+1}\right]$$

Taking any finite n, this interval gives us both an upper and lower bound for e: for example

$$\begin{aligned}
n = 10 &\implies 2.59374 \leq e \leq 2.85311 \\
n = 100 &\implies 2.7048 \leq e \leq 2.73186 \\
n = 1,000 &\implies 2.71692 \leq e \leq 2.71964 \\
n = 1,000,000 &\implies 2.71826 \leq e \leq 2.71829
\end{aligned}$$

Thus, correct to four decimal places we know $e \approx 2.7182$

8.5. Problems

9. Subsequences

Highlights of this Chapter: We define the concept of subsequence, and investigate examples where subsequences behave much simpler than the overall sequence with the example of continued fractions. We then investigate the relationship between the convergence of subsequences and the convergence of a sequence as a whole. This leads to several nice theorems:

- A continued fraction description of the golden ratio and $\sqrt{2}$
- Theorem: a sequence converges if it is a union of subsequences converging to the same limit.
- Theorem: every bounded sequence contains a convergent subsequence.

Definition 9.1. A *subsequence* is a *subset* of a sequence which is itself a sequence. As sequences are infinite ordered lists of real numbers, an equivalent definition is that a subsequence is *any* infinite subset of a sequence.

We often denote an abstract subsequence like s_{n_k}, meaning that we have kept only the n_k terms of the original, and discarded the rest.

Example 9.1 (Example Subsequences)**.** In the sequence of all n-gons inscribed in a circle, the collection studied by archimedes (CITE EALRIER CHAP) by doubling is the subsequence

$$\begin{aligned} P_{3\cdot 2^k} &= (P_{3\cdot 2^1}, P_{3\cdot 2^2}, P_{3\cdot 2^3}, P_{3\cdot 2^4}, \ldots) \\ &= (P_6, P_{12}, P_{24}, P_{48}, \ldots) \end{aligned}$$

Archimedes began his estimation of π using a simple idea: create a sequence of nested intervals (upper and lower bounds) from inscribing and circumscribing n-gons. But then he realized calculations would be much simpler if he focused only on a *subsequence*, namely that generated by side-doubling. We too will often run into situations like Archimedes, where the overall behavior of a sequence is difficult to understand, but we can pull out subsequences that are much easier to work with.

9.1. Continued Fractions

In the previous section, we uncovered a beautiful formula for the golden ratio as the limit of an infinite process of square roots. However, practically speaking (if you were interested in calculating the value of the golden ratio, as the ancient mathematicians were) this series is *useless.* The golden ratio itself involves a square root, so if you are seeking a method of approximations its fair to assume that you cannot evaluate the square root function exactly. But what does our sequence of approximations look like? To calculate the n^{th} term, you need to take n square roots! The very terms of our convergent sequence are actually much *much* more algebraically complicated than their limiting value.

To be practical, we would like a sequence that (1) contains easy to compute terms, and (2) converges to the number we seek to understand. By **?@thm-rational-sequence**, we know for any real number there exists a sequence of rationals that converges to it, and so it's natural to seek a method of producing such a thing.

One method is the *continued fraction*, which is best illustrated by example. We know that the golden ratio L satisfies the equation $L^2 = L + 1$, and dividing by L this gives us an equation satisfied by L and $1/L$:

$$L = 1 + \frac{1}{L}$$

Just like we did above, we can use this self-referential equation to produce a series, by plugging it into itself over and over. After one such substitution we get

$$L = 1 + \frac{1}{1 + \frac{1}{L}}$$

And then after another such we get

$$L = 1 + \frac{1}{1 + \frac{1}{1+\frac{1}{L}}}$$

Continuing this way over and over, we push the L "off to infinity" on the right hand side, and are left with an infinite expression for L, as a limit of a sequence of fractions.

$$L = 1 + \cfrac{1}{1 + \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cdots}}}}}}}$$

Of course, this 'infinite manipulation' is not itself rigorous, but we can interpret this as a recursive sequence exactly as above. Setting $s_1 = 1$, we have the rule $s_{n+1} = 1 + \frac{1}{s_n}$, and we wish to understand $\lim s_n$.

Example 9.2 (Continued Fraction of the Golden Ratio)**.** The continued fration

$$1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cdots}}}}}}}}$$

defined by the recursive sequence $s_1 = 1$, $s_{n+1} = 1 + \frac{1}{s_n}$ limits to the golden ratio.

A continued fraction is a recursive sequence, so we can compute everything with the starting value and a single simple rule. To get a feel for the sequence at hand, let's compute the first few terms:

$$s_1 = 1, s_2 = 2, s_3 = \frac{3}{2}, s_4 = \frac{5}{3}, s_5 = \frac{8}{5}, s_6 = \frac{13}{8}, s_6 = \frac{21}{13}, \dots$$

What's one thing we notice about this sequence from its first few terms? Well - it looks like the fractions are all ratios of Fibonacci numbers! This won't actually be relevant but it's a good practice of induction with the sequence definition, so we might as well confirm it:

Example 9.3 (Fibonacci Numbers and the Golden Ratio)**.** Recall that the Fibonacci numbers are defined by the recurrence relation $F_1 = F_1 = 2$ and $F_{n+2} = F_{n+1} + F_n$. Show that the n^{th} convergent s_n of the continued fraction for the golden ratio is the ratio of the Fibonacci numbers F_{n+1}/F_n.

Proof. This is true for the first convergent which is 1, and $F_2/F_1 = 1/1 = 1$. Assume the n^{th} convergent is $s_n = F_{n+1}/F_n$, and consider the $n+1^{st}$: this is

$$s_{n+1} = 1 + \frac{1}{s_n} = 1 + \frac{1}{\frac{F_{n+1}}{F_n}}$$

$$= 1 + \frac{F_n}{F_{n+1}} = \frac{F_{n+1} + F_n}{F_{n+1}} = \frac{F_{n+2}}{F_{n+1}}$$

□

The more important thing we notice is that looking at the magnitude of the terms, it is neither increasing or decreasing, but it appears the sequence is zig-zagging up and down. Its straightforward to prove this is actually the case:

Example 9.4. If n is odd, then $s_n < s_{n+1}$. If n is even, $s_n > s_{n+1}$.

Proof. Again, we proceed by induction: we prove only the first case, and leave the second as an exercise. Note first $s_1 = 1$, $s_2 = 2$ and $s_3 = \frac{3}{2}$ so $s_1 < s_2$ and $s_2 > s_3$: the base case of each is true.

Now, assume that n is odd, and $s_n < s_{n+1}$. Computing from here

$$s_n < s_{n+1} \implies \frac{1}{s_n} > \frac{1}{s_{n+1}} \implies 1 + \frac{1}{s_n} > 1 + \frac{1}{s_{n+1}}$$

The last line of this computation is the definition of $s_{n+1} > s_{n+2}$,so we see the next one is decreasing as claimed. And applying the recurrence once more:

$$s_{n+1} > s_{n+2} \implies \frac{1}{s_{n+1}} < \frac{1}{s_{n+2}} \implies 1 + \frac{1}{s_{n+1}} < 1 + \frac{1}{s_{n+2}}$$

Where now the last line of the calculation is the definition of $s_{n+2} < s_{n+3}$, fininshing our induction step! □

While the overall sequence isn't monotone, it seems to be built of two different monotone sequences, interleaved with one another! In particular the *odd subsequence* $s_1, s_3, s_5, \ldots$ is monotone increasing, and the *even subsequence* $s_2, s_4, s_6, \ldots$ is monotone decreasing.

To study these subsequences separately, we first need to find a recurrence relation that gives us s_{n+2} in terms of s_n: applying this to either s_1 or s_2 will then produce the entire even or odd subsequence.

$$s_{n+2} = 1 + \frac{1}{s_{n+1}} = 1 + \frac{1}{1 + \frac{1}{s_n}}$$

Example 9.5. The subsequence $s_1, s_3, s_5, s_7, \ldots$ is monotone increasing.

Proof. We prove this by induction. Starting from $s_1 = 1$, we compute

$$s_3 = 1 + \frac{1}{1 + \frac{1}{1}} = 1 + \frac{1}{2} = \frac{3}{2}$$

So $s_1 < s_3$, completing the base case. Next, assume for induction that $s_{n+2} > s_n$. We wish to show that $s_{n+4} > s_{n+2}$. Calculating from our assumption:

$$\begin{aligned}
s_{n+2} > s_n &\implies \frac{1}{s_{n+2}} < \frac{1}{s_n} \\
&\implies 1 + \frac{1}{s_{n+2}} < 1 + \frac{1}{s_n} \\
&\implies \frac{1}{1+\frac{1}{s_{n+2}}} > \frac{1}{1+\frac{1}{s_n}} \\
&\implies 1 + \frac{1}{1+\frac{1}{s_{n+2}}} > 1 + \frac{1}{1+\frac{1}{s_n}} \\
&\implies s_{n+4} > s_{n+2}
\end{aligned}$$

This completes the induction step, so the subsequence of odd terms is monotone increasing as claimed! □

A nearly identical argument applies to the even subsequence:

Exercise 9.1. The subsequence $s_2, s_4, s_6, s_8, \dots$ is monotone decreasing.

Exercise 9.2. Let $f(x) = 1 + \frac{1}{x}$. Show that if $x < y$ then $f(x) > f(y)$; that is, f reverses the ordering of numbers. Use this to give a more streamlined proof that the even and odd subsequences are both monotone, but the overall sequence zigzags.

Now that we know each sequene is monotone, we are in a position similar to the previous chapter where we played two sequences off one another to learn about e. The same trick works to show they are bounded.

Example 9.6. The odd subsequence of s_n is bounded above, and the even subsequence is bounded below.

Proof. The even subseqece is monotone decreasing, but consists completely of positive terms. Thus, its bounded below by zero. Now we turn our attention to the odd subsequence: if n is odd, we know that s_n is bounded above by s_{n+1}, but s_{n+1} is a member of the monotone decreasing even subsequence, so $s_{n+1} < s_2 = 2$. Thus, for all odd n, s_n is bounded above by 2. □

Now we know by monotone convergence that both the even and odd subsequences converge! Next, we show they converge to the same value:

Example 9.7. Both the even and odd subsequences converge to the same value.

Proof. Let $e_n = s_{2n}$ be the even subsequence and $o_n = s_{2n-1}$ the odd subsequence, and write $\lim e_n = E$ and $\lim o_n = \Theta$. We wish to show $E = \Theta$.

Using the recurrence relation we see

$$o_{n+1} = 1 + \frac{1}{e_n} \qquad e_n = 1 + \frac{1}{o_n}$$

and so, using the limit laws and the convergence of e_n, o_n

$$\Theta = 1 + \frac{1}{E} \qquad E = 1 + \frac{1}{\Theta}$$

Therefore we see $\Theta - E = \frac{1}{E} - \frac{1}{\Theta}$, which after getting a common denominator implies

$$\Theta - E = \frac{\Theta - E}{\Theta E}$$

So whatever number $\Theta - E$ is, it has the property that it is unchanged when divided by the number ΘE. But the only number unchanged by multiplication and division is zero! Thus

$$\Theta - E = 0$$

□

Now we know that not only the even and odd subsequences converge but that they converge to the same limit! Its not too much more work to show that the entire sequence converges.

Example 9.8. The sequence s_n converges.

Proof. Call the common limit of the even and odd subsequences L. Let $\epsilon > 0$ Since $s_{2n-1} \to L$ we know there is an N_1 with $n > N_1$ implying $|s_{2n-1} - L| < \epsilon$. Similarly since $s_{2n} \to L$ we can find an N_2 where $n > N_2$ implies $|s_{2n} - L| < \epsilon$.

Set $N = \max\{N_1, N_2\}$. Then if $n > N$ we see both the even and odd subsequences are within ϵ of L by construction, and thus all terms of the sequence are within ϵ of L. But this is the definition of convergence! Thus s_n is convergent and $\lim s_n = L$. □

Finally! Starting with a zigzag sequence where monotone convergene did not apply, we broke it into two subsequences, each of which were monotone, and each of which we could prove converge. Then we showed these subsequences have the same limit and hence the overall sequence converges. We made it! Now its quick work to confirm the limit is what we expected from our construction: the golden ratio.

Example 9.9. The sequence s_n converges to the golden ratio.

Proof. Since throwing away the first term of the sequence does not change the limit, we see $\lim s_{n+1} = \lim s_n = L$. Using the recurrence relation and the limit laws, this implies

$$\lim s_{n+1} = \lim 1 + \frac{1}{s_n} = 1 + \frac{1}{L}$$

THus, the limit L satisfies the equation $L = 1 + 1/L$ or $L^2 = L + 1$. This has two solutions

$$\frac{1 \pm \sqrt{5}}{2}$$

Only one of which is positive. Thus this must be the limit

$$1 + \cfrac{1}{1 + \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cdots}}}}}}}} = \frac{1+\sqrt{5}}{2}$$

s □

We can apply this same process to discover another sequence of rational approximations to $\sqrt{2}$, by algebraic means (in contrast wtih the geometric approach of the babylonians). To start, we need to find a recursive formula that is satisfied by $\sqrt{2}$, and involves a reciprocal: something like

$$\sqrt{2} = \text{Rational stuff} + \frac{1}{\text{Rational stuff and } \sqrt{2}}$$

We can get such a formula through some trickery: first, using the difference of squares $a^2-b^2 = (a+b)(a-b)$ we see that $1 = 2-1 = (\sqrt{2}+1)(\sqrt{2}-1)$, which can be re-written

$$\sqrt{2} - 1 = \frac{1}{1+\sqrt{2}}$$

Now, substitute this into the obvious $\sqrt{2} = 1 + \sqrt{2} - 1$ to get

$$\sqrt{2} = 1 + \frac{1}{1+\sqrt{2}}$$

This is a *self-referential equation*, meaning $\sqrt{2}$ appears on both sides.

Example 9.10 (Continued Fraction of $\sqrt{2}$)**.** The continued fraction

$$1 + \cfrac{1}{2 + \cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cdots}}}}}$$

converges to the square root of 2.

9.2. Subsequences and Convergence

Hopefully this exploration into continued fractions has shown the usefulness of looking for easy-to-work-with subsequences, when theorems such as monotone convergence don't automatically apply. It is then our goal to try and piece this information back together: if we know the limits of various subsequences, what can we say about the entire sequence?

A direct generalization of the even/odd case from our example above shows that if s_n can separated into two subsequences which we can (separately) prove have the same limit, then the entire sequence converges to that

Proposition 9.1 (Union of Two Subsequences)**.** *If s_n is the union of two subsequences s_{n_k} and s_{n_ℓ} which both converge to the same limit, then s_n converges to that limit.*

Proof. Let L be the common limit of s_{n_k} and s_{n_ℓ}, and fix and $\epsilon > 0$. Then there is an N_1 such that for all $k > N_1$ we are assured that $|s_{n_k} - L| < \epsilon$ and there is an N_2 such that for all $\ell > N_2$ we know $|s_{n_\ell} - L| < \epsilon$. Choosing any particular $k \geq N_1$ and $\ell \geq N_2$ we set $N = \max n_k, n_\ell$.

Now choose arbitrary $n > N$. The term s_n is a a member of one of our two subsequences, but by design if $n = n_k$ then $k > N_1$ and if $n = n_\ell$ then $\ell > N_2$, so we know $|s_n - L| < \epsilon$, and our sequence converges as claimed. □

Theorem 9.1. *Let s_n be a sequence, and assume that s_n is the union of N subsequences, all of which converge to the same limit L. Then s_n is convergent, with limit L.*

Sketch. One can prove this directly, but choosing useful notation is tedious. The idea is as follows: for each of the N sequences, let $M_1, M_2, \ldots M_N$ be the threshold beyond which the subsequence is within ϵ of L for some fixed $\epsilon > 0$. Then set $M = \max\{M_1, \ldots, M_N\}$ and note that for all $n > M$ each of the subsequences is within ϵ of L. Because the entire sequence is just the union of these N subsequences, this means that every term of the sequence is within ϵ of L. But this is precisely the definition of $s_n \to L$. So we are done. □

However, a simple example shows its not enough to simply say a sequence is a union of *convergent subsequences*: we do have to know they all have the same limit!

Example 9.11. The sequence $s_n = (-1)^n$ diverges, but its even and odd subsequences form constant (thus convergent) subsequences:

$$\begin{aligned} s_{2n} &= (-1)^{2n} = 1, 1, 1, \ldots \\ s_{2n+1} &= (-1)^{2n-1} = -1, -1, -1, \ldots \end{aligned}$$

Or, more generally

Exercise 9.3. Prove that if a sequence s_n has two subsequences which converge to different values, then the overall sequence diverges. Rephrasing this positively gives a useful criterion: *If s_n is a convergent sequence, then all of its subsequences converge, and have the same limit.*

Remark 9.1. This can be turned into a useful technique to prove two sequences a_n, b_n have the same limit: interleave their terms $a_1, b_1, a_2, b_2, a_3, b_3, \cdots$ and try to prove the resulting sequence converges. If it does, then we know all subsequences have the same limit, and so both a_n and b_n converge to L.

9.3. Bolzano-Weierstrass

What about sequences that *don't* converge? The theorem above says that it cannot be true that all their subsequences converge, but Example 9.11 does show that a divergent sequence can still contain convergent subsequences. A natural question then is - do they always? Alas, a simple counterexample shows us that is too much to ask:

Example 9.12. The sequence $s_n = n^2$ diverges, and all subsequences of it diverge.

But the problem here is not serious, its simply that the original sequence is *unbounded* and cannot possibly contain anything that converges. The perhaps surprising fact that this is the *only constraint preventing the existence of a convergent subsequence* is known as the Bolzano Weierstrass theorem.

Theorem 9.2 (Bolzano-Weierstrass)**.** *Every bounded sequence has a convergent subsequence*

There are many ways to prove this, but a particularly elegant one uses (of course!) the monotone convergence theorem.

At first this sounds suspicious: we must confront head on the issue we ran into above, that not every sequence is monotone! However, the weaker property we actually need *is* true: while not every sequence is monotone, every sequence contains a monotone subsequence. There is a very clever argument for this, which needs one new definition.

Definition 9.2 (Peak of a Sequence)**.** Let s_n be a sequence and $N \in \mathbb{N}$. Then s_N is a *peak* if it is larger than all following terms of the sequence:

$$s_N \geq s_m \;\; \forall m > N$$

Theorem 9.3 (Monotone Subsequences)**.** *Every sequence contains a monotone subsequence*

Proof. Let s_n be an arbitrary sequence. Then there are two options: either s_n contains infinitely many peaks or it does not.

If s_n contains infinitely many peaks, we can build the subsequence of all peaks. This is monotone decreasing: if p_1 is the first peak, then its greater than or equal to all subsequent terms s_n, and so its greater than or equal to the second peak p_2. (But, nothing here is special about 1 and 2, this holds for the n^{th} and $n + 1^{st}$ peak without change).

Otherwise, if s_n contains only finitely many peaks, we will construct a monotone increasing subsequence as follows. Since there are finitely many peaks there must be a *last peak*, say this occurs at position N. Then s_{N+1} is not a peak, and we will take this as the first term of our new sequence (let's call it q_1). Because its not a peak, by definition there is some term farther down the sequence which is *larger* than s_{N+1} - say this happens at index N_2 and call it q_2. But q_2 is *also* not a peak (as there were only finitely many, and we are past all of them), so there's a term even farther down - say at index N_3 which is larger: call it q_3. Now we have $q_1 < q_2 < q_3$, and we can continue this procedure inductively to build a monotone increasing subsequence for all n. □

Now, given that every sequence has a monotone subsequence, we know that every *bounded* sequence has a *monotone and bounded subsequence.* Such things converge by MCT, so we know every sequence has a convergent subsequence! This is the *Bolzano Weierstrass Theorem.*

We will use this often in whats to come to produce examples of convergent subsequences where it might otherwise be difficult to do so. Here's a first example of such an argument

Proposition 9.2 (Analyzing All Convergent Subsequences)**.** *If s_n is a bounded sequence such that every* convergent subsequence *converges to the same value, then s_n converges.*

Proof. Assume that every convergent (proper) subsequence of s_n converges to L, but that s_n itself does not. Then fixing some $\epsilon > 0$ for each $n = 1, 2, 3, \ldots$ we can find an k such that $|s_{n_k} - L| > \epsilon$. This is an infinite sequence of terms all of which are farther than ϵ from L, and is bounded as s_n itself was bounded. Thus, by the Bolzano Weierstrass theorem there is a *subsequence* of this that converges. Its limit cannot be L because all terms in the sequence are more than ϵ away from L, so we've found a subsequence of the original sequence that converges to a different value, contradicting the original assumption. □

In addition to applications like the above, in time will come to appreciate it as one of the most elegant tools available to us. There will come many times (soon, when dealing with functions) where we can easily produce a sequence of points satisfying some property, but to make progress we need a *convergent* sequence of such points. The BW theorem assures us that we don't have to worry - we can always make one by just throwing out some terms, so long as the sequence we have is bounded.

9.4. Limsup and Liminf

When a sequence doesn't converge, it can have various subsequences that converge to different limits. One way to reign in the complexity of such things is via the concepts of *limit supremum* and *limit infimum.*

Definition 9.3 (Limsup and Liminf)**.** Let s_n be a bounded sequence, and for each $N \in \mathbb{N}$ define $u_N = \sup_{n\geq N}\{s_n\}$. Then we define the limit superior of s_n as

$$\limsup s_n := \lim u_N = \lim_{N\to\infty} \sup_{n\geq N}\{s_n\}$$

Similarly, with $\ell_N = \inf_{n\geq N}\{s_n\}$ we define the limit inferior of s_n to be

$$\liminf s_n := \lim \ell_N = \lim_{N\to\infty} \inf_{n\geq N}\{s_n\}$$

We will have occasional use for this definition during the course, mostly because of the following: that the limsup and liminf exist for all (bounded) sequences, not just convergent ones.

Proposition 9.3 (Existence of Limsup, Liminf)**.** *Let s_n be a bounded sequence. Prove that* $\limsup s_n$ *and* $\liminf s_n$ *both exist*

Proof. We verify for the limsup case, and leave the analgoous liminf as an exercise. For each N let $u_N = \sup_{n\geq N}\{s_n\}$ and notice that u_N is a monotone decreasing sequence as we are taking the supremum over smaller and smaller sets (Exercise 3.4). Its easy to construct a lower bound for u_N: since s_n is bounded we may take any lower bound L and note for any N this is a lower bound for the tail $\{s_n \mid n > N\}$. As the supremum is an upper bound and all lower bounds are $\leq$ all upper bounds we see $L \leq \sup_{n>N}\{s_n\} = u_N$. Thus u_N is a monotone decreasing sequence that is bounded below, which converges by the Monotone Convergence Theorem. □

One use of these quantities is to bound the possible values of subsequential limits:

Exercise 9.4 (Bounding Subsequential Limits)**.** Let s_n be any bounded sequence and s_{n_k} a convergent subsequence with limit L. Then

$$\liminf s_n \leq L \leq \limsup s_n$$

Proposition 9.4. *A sequence s_n converges if and only if*

$$\limsup s_n = \liminf s_n$$

Proof. For any N we have $\ell_N = \inf_{n\geq N} s_n \leq s_N \leq \sup_{n\geq N} s_n = u_N$. But $\lim \ell_N = \liminf s_n = \limsup s_n = \lim u_N$ by assumption, so the squeeze theorem (Theorem 7.1) implies s_N also converges, and has the same limit. $\square$

It is often useful to broaden our use of lim sup and lim inf to the extended reals $\mathbb{R} \cup \{\pm\infty\}$. Here its true that *every sequence* has a limsup and liminf: either the sequence is bounded and they exist as proven above, or the sequence is unbounded with $\limsup, \liminf = \pm\infty$. This slight generalization makes certain theorems easier to state, as one can use lim sup and lim inf without first checking they exist.

Exercise 9.5 (Subsequences & Limsup, Liminf)**.** Let s_n be a bounded sequence. Then there exists a subsequence which converges to the limsup and a subsequence which converges to the liminf.

9.5. Problems

Exercise 9.6. For any fixed n, prove that the following continued fraction exists, and find its vallue.

$$\cfrac{1}{n+\cfrac{1}{n+\cfrac{1}{n+\cfrac{1}{n+\cfrac{1}{n+\cfrac{1}{n+\cfrac{1}{n+\cdots}}}}}}}$$

Exercise 9.7 (Continued Fractions for Roots)**.** Let p be any prime number, find the continued fraction for $\sqrt{p}$.

Knowing such sequences is extremely useful for computation, in the age before computers: if n is a composite number we can find $\sqrt{n}$ by multiplying together the square roots of its prime factorization

Exercise 9.8. Find a rational approximation to $\sqrt{6}$ by calculating the first three terms in the continued fraction expansions for $\sqrt{2}$ and $\sqrt{3}$.

We could also find a continued fraction directly for cases like this, with a little more care:

Exercise 9.9. Find the continued fraction expansion for $\sqrt{pq}$ if p and q are primes. What happens to your procedure when $p = q$?

9.5.1. Limsup and Liminf

Exercise 9.10. Let a_n, b_n be bounded sequences. Prove that

$$\limsup(a_n + b_n) \leq \limsup a_n + \limsup b_n$$

$$\liminf a_n + \liminf b_n \leq \liminf(a_n + b_n)$$

Provide counterexamples to show that equality does not always hold.

Exercise 9.11. Let a_n be convergent and b_n be an arbitrary bounded sequence. Show that

$$\limsup(a_n + b_n) = \lim a_n + \limsup b_n$$
$$\liminf(a_n + b_n) = \lim a_n + \liminf b_n$$

Exercise 9.12. Let a_n be convergent and b_n be an arbitrary bounded sequence. Show that

$$\limsup(a_n b_n) = (\lim a_n)(\limsup b_n)$$
$$\liminf(a_n b_n) = (\lim a_n)(\liminf b_n)$$

9.5.2. ⋆ An Alternative Proof of Bolzano-Weierstrass

An alternative argument for the BW theorem proceeds via the nested interval property. Here's an outline of how this works

- If s_n is bounded then there is some a, b with $a \leq s_n \leq b$ for all n. Call this interval I_0, and inductively build a sequence of nested closed intervals as follows
- At each stage $I_k = [a_k, b_k]$, bisect the interval with the midpoint $m_k = \frac{a_k + b_k}{2}$. This divides I_k into two sub-intervals, and since I_k contains infinitely many points of the sequence, one of these two halves must still contain infinitely many points. Choose this as the interval I_{k+1}.
- Now, this sequence of nested intervals has nonempty intersection by the Nested Interval Property. So, let $L \in \mathbb{R}$ be a point in the intersection.

- Now, we just need to build a subsequence of s_n which converges to L. We build it inductively as follows: let the first term be s_1, and then for each k choose some point $s_{n_k} \in I_k$ that is distinct from all previously chosen points (we can do this because there are infinitely many points available in I_k and we have only used finitely many so far in our subsequence).
- This new sequence is trapped between a_k and b_k, which both converge to L. Thus it converges by the squeeze theorem!

Exercise 9.13. In this problem, you are to check the main steps of this proof to ensure it works. Namely, given the above situation prove that

- If $I_k = [a_k, b_k]$, the sequences a_k and b_k of endpoints converge. *Hint: Monotone Convergence*
- $\lim a_k = \lim b_k$, so the Squeeze theorem really does apply *Hint: use that at each stage we are bisecting the intervals: can you find a formula for the sequence $b_k - a_k$, and prove this converges to zero?

Exercise 9.14 (Simultaneous Bolzano Weierstrass)**.** Given two bounded sequences x_n, y_n there is a subsequence n_k of indices such that both x_{n_k} and y_{n_k} converge. Prove this, and then use induction to prove that for any finite number of bounded sequences, one can choose a subsequence of indices so they all converge.

10. Cauchy Sequences

10.1. Definition

One reasonably ambitious sounding goal in the study of sequences is to find a nice criterion to determine exactly when a sequence converges or not. We made partial progress towards this in the previous two chapters, and our goal in this chapter is to provide an alternative complete characterization, by a single simple property. But what could such a property be? One (good!) thought is the following

> When a sequence converges, terms eventually get close to some limit L. Thus the terms of the sequence eventually get close to one another.

One way to formalize this is as follows:

Definition 10.1 (Cauchy Sequence)**.** A sequence s_n is Cauchy if for all $\epsilon > 0$ there is a threshold past which any two terms of the sequence differ from one another by at most ϵ. As a logic sentence,

$$\forall \epsilon > 0 \ \exists N \ \forall m, n > N \ |s_n - s_m| < \epsilon$$

Example 10.1 (Cauchy Sequences: An Example)**.** The sequence $s_n = \frac{1}{n}$ is cauchy: we can see this because for any n, m

$$\left|\frac{1}{n} - \frac{1}{m}\right| < \left|\frac{1}{n} - 0\right| = \frac{1}{n}$$

And we already know that for any ϵ we can choose N with $n > N$ implying $1/n < \epsilon$.

Example 10.2 (Cauchy Sequences: A Nonexample)**.** The sequence $s_n = 1, 0, 1, 0, 1, 0 \ldots$ is not Cauchy, as the difference between any two consecutive terms is 1. Thus for $\epsilon = 1/2$ there is no N where past that N, every s_n is within $1/2$ of each other.

10.1.1. ⋆ Why This Definition?

There are several ways one could have attempted to mathematically make rigorous the statement *the terms of the sequence get close to one another.* But not all of them force a sequence to converge. For example, what if we had only looked at consecutive terms, and proposed instead

For all $\epsilon > 0$ there is an N where if $n > N$ then $|a_n - a_{n+1}| < \epsilon$

Unfortunately this doesn't work. Perhaps surprisingly, it *is possible* for consecutive terms of a sequence to all get within ϵ of one another, but for the overall sequence to diverge.

Example 10.3 ($|a_n - a_{n+1}|$ small but a_n diverges!)**.** Consider the sequence $a_n = \sqrt{n}$. Then for all $\epsilon > 0$ there is an N where $n > N$ implies $|\sqrt{n} - \sqrt{n+1}| < \epsilon$, but nonetheless a_n diverges (to infinity).

Proof. We can estimate the difference between consecutive terms with some algebra:

$$\begin{aligned}\sqrt{n+1} - \sqrt{n} &= \left(\sqrt{n+1} - \sqrt{n}\right)\frac{\sqrt{n+1}+\sqrt{n}}{\sqrt{n+1}+\sqrt{n}} \\ &= \frac{(n+1)-n}{\sqrt{n+1}+\sqrt{n}} \\ &= \frac{1}{\sqrt{n+1}+\sqrt{n}} \\ &< \frac{1}{\sqrt{n}}\end{aligned}$$

Thus for any $\epsilon > 0$ we can just take $N = \frac{1}{\epsilon^2}$ and see that if $n > N$ we have

$$|a_{n+1} - a_n| < \frac{1}{\sqrt{n}} < \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{\frac{1}{\epsilon^2}}} = \epsilon$$

Nowever, a_n is not converging to any finite number, as for any $M > 0$, if $n > M^2$ then $a_n = \sqrt{n} > M$, so $a_n \to \infty$ by Definition 6.4 □

Example 10.4 ($|a_n - a_{n+1}|$ small but a_n diverges, again!)**.** Perhaps the most famous example with this property is the *harmonic series*

$$a_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

Here it is clear that $a_n - a_{n+1} = \frac{1}{n+1}$ and we know this can be made smaller than any $\epsilon > 0$. However, as we will prove in future chapters, this sequence nonetheless diverges to infinity.

So, we need to ask for a *stronger condition.* What went wrong? Well, even though we forced a_n to be close to a_{n+1} for all n, the small differences between consecutive terms could still manage to add up to *big differences* between terms: even if a_n was within 0.01 of a_{n+1} for all n, its totally possible that $a_{n+100,000}$ could differ from a_n by $(0.01)(10,000) = 100$! So, to strengthen our definition we might try to impose that *all terms of the sequence eventually stay close together*:

10.2. ⋆ Properties

A good way to get used to a new definition is to *use it.* This definition looks very similar to the limit definition, which means we can often formulate analogous theorems and proofs to things we've seen before:

Note the proofs in this section are not *logically required* as the next section will render them superfluous: once we know Cauchy and convergent are equivalent, these all follow as immediate corollaries of the limit laws! Nonetheless it is instructive to see their direct proofs:

Proposition 10.1 (Cauchy Implies Bounded)**.** *If s_n is Cauchy then its bounded: there exists a B such that $|s_n| < B$ for all $n \in \mathbb{N}$.*

Very similar to proof for convergent seqs Proposition 7.2 in style, where we show after some N all the terms are bounded by some particular number, and then take the max of this and the (finitely many!) previous terms to get a bound on the entire sequence.

Proof. Set $\epsilon = 1$. Since a_n is Cauchy we know there is some N beyond which $|a_n - a_m| < 1$ for all $n, m > N$. In particular, this means every |a_n-a_{N+1}|<1$ so

$$a_{N+1} - 1 < a_n < a_{N+1} + 1$$

Thus for the (infinitely many terms!) after a_N, we can bound all of them above by $a_{N+1} + 1$ and below by $a_{N+1} - 1$. To extend these to bounds for the whole sequence, we just take the max or min with the (finitely many!) previous terms:

$$L = \min\{a_1, a_2, \dots, a_N, a_{N+1} - 1\}$$

$$U = \max\{a_1, a_2, \dots, a_N, a_{N+1} + 1\}$$

Now we have for all n, $L \leq a_n \leq U$ so $\{a_n\}$ is bounded. □

Proposition 10.2 (Sums of Cauchy Sequences)**.** *If a_n and b_n are Cauchy sequences, so is $a_n + b_n$.*

Proof. Let $\epsilon > 0$. Then choose N_a and N_b such that for all n, m greater than N_a, N_b respectively, we have $|a_n - a_m| < \epsilon/2$ and $|b_n - b_m| < \epsilon/2$. Set $N = \max\{N_a, N_b\}$ and let $n, m > N$. Then each of the above two inequalities hold, and so by the triangle inequality

$$\begin{aligned} |(a_n + b_n) - (a_m - b_m)| &= |(a_n - a_m) + (b_n - b_m)| \\ &\leq |a_n - a_m| + |b_n - b_m| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

Thus, $a_n + b_n$ is Cauchy as well. □

Exercise 10.1 (Constant Multiples of Cauchy Sequences)**.** Let a_n be Cauchy, and $k \in \mathbb{R}$ be constant. Then ka_n is Cauchy.

Proposition 10.3 (Products of Cauchy Sequences)**.** *Let a_n, b_n be Cauchy. Then $s_n = a_n b_n$ is a cauchy sequence.*

First, some scratch work: we are going to want to work with the condition $|s_n - s_m| = |a_n b_n - a_m b_m|$. But we only know things about the quantities $|a_n - a_m|$ and $|b_n - b_m|$. So, we need to do some algebra, involving *adding zero in a clever way* and *applying the triangle inequality*:

$$\begin{aligned} |a_n b_n - a_m b_m| &= |a_n b_n + (a_n b_m - a_n b_m) - a_m b_m| \\ &= |(a_n b_n - a_n b_m) + (a_n b_m - a_m b_m)| \\ &= |a_n(b_n - b_m) + b_m(a_n - a_m)| \\ &\leq |a_n(b_n - b_m)| + |b_m(a_n - a_m)| \\ &= |a_n||b_n - b_m| + |b_m||a_n - a_m| \end{aligned}$$

Because we know Cauchy sequences are bounded, we can get upper estimates for both $|a_n|$ and $|b_n|$. And then as we know that the sequences are Cauchy, we can make $|a_n - a_m|$ and $|b_n - b_m|$ as small as we need, so that this overall term is small. We carry this idea out precisely in the proof below.

Proof. Let a_n and b_n be Cauchy, and choose an $\epsilon > 0$. Then each are bounded, so we can choose some M_a with $|a_n| < M_a$ and M_b where $|b_n| < M_b$ for all n. To make notation easier, set $M = \max\{M_a, M_b\}$ so that we know both a_n and b_n are bounded by the same constant M.

Using that each is Cauchy, we can also get an N_a and N_b where if n, m are greater than these respectively, we know that

$$|a_n - a_m| < \frac{\epsilon}{2M} \qquad |b_n - b_m| < \frac{\epsilon}{2M}$$

Then set $N = \max\{N_a, N_b\}$, and choose arbitrary $n, m > N$. Since in this case both of the above hypotheses are satisfied, we know that

$$|a_n||b_n - b_m| \leq M\frac{\epsilon}{2M} = \frac{\epsilon}{2} \qquad |b_m||a_n - a_m| \leq M\frac{\epsilon}{2M} = \frac{\epsilon}{2}$$

Together, this means their sum is less than ϵ, and from our scratch work we see their sum is already an upper bound for the quantity we are actually interested in:

$$|a_nb_n - a_mb_m| \leq |a_n||b_n - b_m| + |b_m||a_n - a_m| \leq \epsilon$$

□

Exercise 10.2 (Reciprocals of Cauchy Sequences)**.** Let a_n be a Cauchy sequence with $a_n \neq 0$ for all n, which does not converge to zero. Then the sequence of reciprocals $s_n = \frac{1}{a_n}$ is Cauchy.

Just like for convergence, once we know the results products and reciprocals, quotients follow as an immediate corollary:

Corollary 10.1 (Quotients of Cauchy Sequences)**.** *If a_n and b_n are Cauchy with $b_n \neq 0$ and $\lim b_n \neq 0$ then the quotients a_n/b_n form a Cauchy sequence.*

Exercise 10.3. Show the hypothesis $b_n \not\to 0$ is necessary in Corollary 10.1 by giving an example of two Cauchy sequences a_n, b_n where $b_n \neq 0$ for all n, yet $\frac{a_n}{b_n}$ is *not* a Cauchy sequence.

10.3. Convergence

Now we move on to the main act, where we prove convergence is equivalent to Cauchy by showing an implication in both directions.

Exercise 10.4 (Convergent Implies Cauchy)**.** If s_n is a convergent sequence, then s_n is Cauchy. *Hint: The triangle inequality and $|a_n - a_m|$ for a sequence converging to L can tell you....what?*

More difficult, and more interesting, is the converse:

Proposition 10.4 (Cauchy Implies Convergent)**.** *If s_n is a Cauchy sequence, then s_n is convergent.*

Proof. Let s_n be a Cauchy sequence. Then it is bounded, by Proposition 10.1, so by the Bolzano Weierstrass theorem (**?@thm-thm-bolzano-weierstrass**) we can extract a subsequence s_{n_k} which converges to some real number L.

Now we have something to work with, and all we need to show is that the rest of the sequence also converges to L. So, let's fix an $\epsilon > 0$. Since $s_{n_k} \to L$ there exists an N_1 where if $n_k > N_1$ we know $|s_{n_k} - L| < \epsilon/2$. And, since s_n is Cauchy, we know there is an N_2 where for any $n, m > N_2$ we know $|s_n - s_m| < \epsilon/2$.

Let $N = \max\{N_1, N_2\}$, and choose any $n > N$. If s_n is in the subsequence, we are good because $n > N_1$ and we know for such elements of the subsequence $|s_n - L| < \epsilon/2 < \epsilon$. But if s_n is not in the subsequence, choose any m such that $m > N$ and s_m is in the subsequence, and apply the triangle inequality:

$$|s_n - L| = |s_n - s_m + s + m - L| \leq |s_n - s_m| + |s_m - L| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Where the first inequality is because of the Cauchy condition, and the second is the convergence of the subsequence. □

Together these imply the main theorem we advertised.

Theorem 10.1 (Cauchy $\iff$ Convergent)**.** *The conditions of being a Cauchy sequence and a convergent sequence are logically equivalent.*

10.4. Problems

Exercise 10.5. Is the sequence $s_n = 1 - \frac{(-1)^n}{n}$ cauchy nor not? Prove your claim.

Exercise 10.6. Let s_n be a periodic sequence (meaning after some period P we have $s_n = s_{P+n}$ for all n). Prove that if s_n is Cauchy then it is constant. *Hint: what's the contrapositive?*

Exercise 10.7. Prove directly from the definition of Cauchy: if s_n is Cauchy and s_{n_k} is a subsequence whose limit is L then $s_n \to L$.

11. Contraction Maps

The cauchy condition (that terms "bunch up") appears in many natural situations, which makes it a very useful equivalent to convergence. Here we investigate one such instance, when sequences are generated by iterating certain functions.

Definition 11.1 (Contractive Sequence)**.** A sequence s_n is *contractive* if there is some positive constant $k < 1$ such that for all n

$$|a_{n+1} - a_n| < k|a_n - a_{n-1}|$$

Definition 11.2 (Contraction Map)**.** A function f is a *contraction map* if there is some positive constant $k < 1$ such that for all x, y in the domain of f,

$$|f(x) - f(y)| < k|x - y|$$

Proposition 11.1. *If f is a contraction map, iterating f starting at any point of the domain produces a convergent sequence.*

Proof. We prove the *existence* of a fixed point of f by constructing a sequence that converges to it. Start by choosing any $x_0 \in \mathbb{R}$ and set $\delta = |f(x_0) - x_0|$. We can define a sequence x_n by iterating f: $x_{n+1} := f(x_n)$. Our goal is to show that x_n is Cauchy by bounding $|x_m - x_n|$ appropriately. As a starting point, note that as f is a contraction map we can choose a positive $k < 1$ where $|f(x) - f(y)| \leq k|x - y|$ for all x, y and compute

$$\begin{aligned}
|x_{n+1} - x_n| &= k|f(x_n) - f(x_{n-1})| \\
&\leq k^2|f(x_{n-1}) - f(x_{n-2})| \\
&\leq \cdots \\
&\leq k^n|f(x_0) - x_0| \\
&= k^n\delta
\end{aligned}$$

This is not the quantity we are really interested in however, we wish to bound $|x_m - x_n|$. Using the triangle inequality, we can replace this with $n-m$ bounds of the form above:

$$\begin{aligned}|x_m - x_n| &= |x_m - x_{m-1} + x_{m-1} - x_{m-2} + x_{m-2} - \cdots - x_{n+1} + x_{n+1} - x_n| \\ &\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \cdots + |x_{n+1} - x_n| \\ &\leq k^{m-1}\delta + k^{m-2}\delta + \cdots + k^n\delta\end{aligned}$$

Simplifying this we can factor out the δ to get a sum

$$\delta\left(k^{m-1} + k^{m-2} + \cdots + k^n\right) = \delta \sum_{j=n}^{m-1} k^j = \delta k^n \sum_{j=0}^{m-n-1} k^j$$

The final sum here we may recognize as a *geometric series* (Exercise 1.6), where

$$\sum_{j=0}^{N} k^j = \frac{1-k^{N+1}}{1-k} \leq \frac{1}{1-k}$$

Putting this all together, we have managed to bound $|x_m - x_n|$ by

$$|x_m - x_n| \leq \delta k^n \sum_{j=0}^{m-n-1} k^j \leq \delta k^n \frac{1}{1-k}$$

The numbers δ and $1/(1-k)$ are both constants, and as $|k| < 1$ we know by Example 6.8, $k^n \to 0$. Thus by the limit laws our bound $\delta k^n \frac{1}{1-k} \to 0$, which means for every ϵ there is some N where $n > N$ implies this is less than ϵ, and hence that $|x_m - x_n| < \epsilon$. This means the sequence of x's is Cauchy, and hence convergent. Thus there is some $x \in \mathbb{R}$ with $x_n \to x$. □

Thinking harder about the limit of this sequence proves a rather important theorem in analysis:

Theorem 11.1 (The Contraction Mapping Theorem)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a contraction map. Then there is a unique real number x such that $f(x) = x$.*

Proof. Starting from any x_0 in the domain, iterating f produces a sequence converging to some limit point x. Since $\lim x_n = \lim x_{n+1} = \lim f(x_n)$ we see this same point x is also the limit of the sequence $f(x_n)$, so it suffices to prove that $f(x_n) \to f(x)$. Then since f is a contraction map, there's a $k < 1$ where $|f(x_n) - f(x)| \leq k|x_n - x|$. This second sequence converges to 0 by assumption, so $f(x_n) \to f(x)$ as required (Exercise 6.17). Finally by uniqueness of limits (Theorem 6.1) since $f(x_n) \to x$ and $f(x_n) \to f(x)$ we conclude $x = f(x)$ is a fixed point.

Finally, we check uniqueness. Assume there are two fixed points x, y with $x = f(x)$ and $y = f(y)$. We apply the condition that f is a contraction map to $|x - y|$ to get

$$|x - y| \leq k|x - y|$$

Since k is strictly less than 1, the only solution to this is that $|x - y| = 0$, so $x = y$ and there is only one fixed point after all. □

The contraction mapping theorem makes quick work of some proofs that previously took us rather a lot of work.

Example 11.1 (Proving $a^n \to 0$ via contraction mapping)**.** If $|a| < 1$ the sequence a^n converges to zero.

Proof. This sequence is produced by iteration of the map $f(x) = ax$ starting from $x = 1$ at $n = 0$. When $|a| < 1$ we see immediately that f is a contraction map as $|f(x) - f(y)| = |ax - ay| = |a||x - y|$, which is less than $k|x - y|$ for any $k \in (a, 1)$. Its immediate to see the fixed point of $x \mapsto ax$ is $x = 0$, and the proof above assures this fixed point is the limit of any sequence of iterates of f. Thus $a^n \to 0$. □

Example 11.2 ($\sqrt{1 + \sqrt{1 + \sqrt{1 + \ldots}}}$ via Contraction Mapping)**.**

Proof. The recursive sequence $\sqrt{1 + \sqrt{1 + \sqrt{1 + \ldots}}}$ we previously studied with monotone convergence is generated by iterating the function $f(x) = \sqrt{1 + x}$. This map sends nonnegative numbers to nonnegative numbers, so we can consider it a function $f : [0, \infty) \to [0, \infty)$.

Furthermore, its straightforward to verify its a contraction map on this domain, by a clever factoring (difference of squares). Starting with $x, y \geq 0$ and noticing $|x - y| = |(1 + x) - (1 + y)|$, we factor as

$$|(1 + x) - (1 + y)| = \left|\left(\sqrt{1 + x} + \sqrt{1 + y}\right)\left(\sqrt{1 + x} - \sqrt{1 + y}\right)\right|$$

As $x, y \geq 0$ we see $\sqrt{1 + x} \geq 1$, $\sqrt{1 + y} \geq 1$ so the first factor $\sqrt{1 + x} + \sqrt{1 + y} \geq 2$. Thus

$$|x - y| = \left|\sqrt{1 + x} + \sqrt{1 + y}\right|\left|\sqrt{1 + x} - \sqrt{1 + y}\right| \geq 2\left|\sqrt{1 + x} - \sqrt{1 + y}\right|$$

Dividng by 2 directly shows f is a contraction map with constatn $k = \frac{1}{2}$:

$$\frac{1}{2}|x - y| \geq \left|\sqrt{1 + x} - \sqrt{1 + y}\right| = |f(x) - f(y)|$$

Thus, the contraction mapping theorem applies, and iterating f from any starting point produces a convergent sequence, whose limit is the unique fixed point of f. Solving for this fixed point we see

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

Which together with the contstraint $x \in [1, \infty)$ (which is the domain we proved f is a contraction on) yields $x = \frac{1+\sqrt{5}}{2}$, the golden ratio.

Note this proves something stronger than our original claim: we know not just the sequence starting with $x_0 = 1$ converges to the golden ratio, but *any starting point ≥ 1, for example,

$$\sqrt{17}, \sqrt{1+\sqrt{17}}, \sqrt{1+\sqrt{1+\sqrt{17}}}, \dots$$

□

Exercise 11.1 (Contraction Maps and $\sqrt{2}$)**.** The babylonian procedure for approximating $\sqrt{2}$ defined the recursive sequence

$$w_{n+1} = \frac{1}{2}\left(w_n + \frac{2}{w_n}\right)$$

which is generated by iterating the function

$$f(x) = \frac{1}{2}\left(x + \frac{2}{x}\right)$$

- Show that f maps $[1, \infty)$ into $[1, \infty)$.
- Show that on this domain, f is a contraction map.
- Show that if $f(x) = x$ then $x^2 = 2$.

Now apply the contraction mapping theorem to re-do a large collection of previous work; proving that (1) $\sqrt{2}$ exists in $\mathbb{R}$ and (2) that the babylonian sequence starting from *any possible initial rectangle of area 2* converges to $\sqrt{2}$. (Before we'd only shown this for the 2×1 rectangle, so $w_0 = 1$.)

We can combine the contracting mapping theorem with other techniques to continue re-proving known results in easier ways. Here we take another look at our argument for the basic limit $a^{1/n} \to 1$.

Exercise 11.2 (Proving $a^{1/n} \to 1$ by Contraction Mapping)**.** Fill in the following outline to prove $a^{1/n} \to 1$ for all $a > 0$:

The sequence $a^{1/n}$ is not generated by iterating a function, but it has many subsequences that are. For example the subsequence

$$a, a^{1/2}, a^{1/4}, a^{1/8}, \dots$$

Is generated by iterating $f(x) = \sqrt{x}$.

- Prove that this map is a contraction map on the domain $[1, \infty)$, which guarantees this subsequence converges for every $a \geq 1$. Find the limiting value.

- Now, for $a > 1$ prove the full sequence $a^{1/n}$ is monotone decreasing and bounded below by 1. This ensures it converges (by monotone convergence).

- Putting these two facts together, we see the entire sequence converges to 1, as if we have a convergent sequences, all subsequences have the same limit!

- Finally, use the limit laws to argue the same holds even when $a < 1$.

For some applications, its useful to have around a slight generalization of the contraction mapping theorem: what can we say about the case where f is *not* a contraction map, but some number of iterations f^N is? The contraction mapping theorem applies to f^N, showing this map has a fixed point. But that doesn't directly imply f does: after all, being fixed by f^N just means that a point *returns to itself* after N applications: it could be a *periodic point* of f. However, a little deeper thought shows this is not the case:

Theorem 11.2 (Root of a Contraction). *Let $f : X \to X$ be a map such that the N-fold composition f^N of f with itself is a contraction map. (That is, "f is the Nth root of a contraction"). Then the conclusion of the Contraction Mapping Theorem applies to f: namely f has a unique fixed point, and iterating f on any initial point $x_0 \in X$ produces a sequence which converges to this fixed point.*

Proof. Since f^N is a contraction map it has a unique fixed point $a \in X$. Thus one way of showing that something is equal to a is to show that its fixed by f^N. But a clever trick lets us see that f^N fixes $f(a)$: since we are just composing f with itself over and over,

$$f^N(f(a)) = f^{N+1}(a) = f(f^N(a)) = f(a)$$

where the last inequality follows as $f^N(a) = a$ by definition. Thus f^N fixes $f(a)$ so this must be the unique fixed point a itself! $f(a) = a$ as desired. Additionally, this is the *only* fixed point of f as any points fixed by f are fixed by all repeated compositions of f, so would be fixed points of f^N.

It remains only to see that starting from an arbitrary $x_0 \in X$, repeated iteration produces a convergent sequence with limit a. The trick is to break the sequence $f^n(a)$ down into N subsequences, one for each $r \in \{0, \ldots N-1\}$:

$$\begin{array}{cccccc}
x_0 & f^N(x_0) & f^{2N}(x_0) & \dots, & f^{nN}(x_0) & \dots \\
f(x_0) & f^{1+N}(x_0) & f^{1+2N}(x_0) & \dots, & f^{1+nN}(x_0) & \dots \\
f^2(x_0) & f^{2+N}(x_0) & f^{2+2N}(x_0) & \dots, & f^{2+nN}(x_0) & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
f^r(x_0) & f^{r+N}(x_0) & f^{r+2N}(x_0) & \dots, & f^{r+nN}(x_0) & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
f^{N-1}(x_0) & f^{(N-1)+N}(x_0) & f^{(N-1)+2N}(x_0) & \dots, & f^{(N-1)+nN}(x_0) & \dots
\end{array}$$

Each of these sequences is the iteration of f^N on some initial point $f^r(x_0)$, and so by the contraction mapping theorem converges to the unqiue fixed point a of f^N. Thus, we've written our sequence as the union of N sequences, each of which has the same limit! So the entire sequence converges to this limit, $f^n(x_0) \to a$ as desired. $\square$

Part III.

Series

- In Chapter 12 we define infinite series and inifnite products, and study some basic examples.
- In Chapter 13 we develop theorems (known as *convergence tests*) to help us determine when a series converges, even if we cannot find its value.
- In **?@sec-series-limits** we look at limits of infinite series, a special case of the *iterated limits* studied previously.
- In Chapter 15 we take a brief look at some advanced techniques for working with infinite series, including summation by parts and double summation
- In **?@sec-series-rearrangement** we explore the vast differences between *conditionally convergent series* and *absolutely convergent series.*

12. First Examples

Highlights of this Chapter: We define infinite series and infinite products, and look at some examples where one can compute the sum exactly: telescoping and geometric series. We study the sums of reciprocal powers of n, and show that the harmonic series $\sum \frac{1}{n}$ diverges, whereas $\sum \frac{1}{n^2}$ converges.

We've met several types of sequences so far where its possible to precisely describe their terms, which basically fall into one of two categories: those with *closed forms* like $a_n = \frac{\sqrt{1+3n^2}}{2n+1}$ where each term is given explicitly in terms of n, and *recursive sequences* where each term is given in terms of the previous ones. The simples recursive sequences are defined by just iterating a single function, which we've successfully attacked with monotone convergence and the Contraction Mapping Theorem. Perhaps the next simplest recursive sequences are from iterating a *process* like addition or multiplication, so we will study these now.

Definition 12.1 (Series)**.** A series s_n is a recursive sequence defined in terms of another sequence a_n by the recurrence relation $s_{n+1} = s_n + a_n$. Thus, the first terms of a series are

$$s_0 = a_0, \qquad s_1 + a_0 + a_1 \qquad s_2 = a_0 + a_1 + a_2 \ldots$$

We use *summation notation* to denote the terms of a series:

$$s_n =_0 +a_1 + \cdots + a_n = \sum_{k=0}^{n} a_k$$

Remark 12.1. It is important to carefully distinguish between the sequence a_n of *terms* being added up, and the sequence s_n of partial sums.

When a series converges, we often denote its limit using summation notation as well. The traditional 'calculus notation' sets n to infinity as the upper index; and another common notation is to list just the subset of integers over which we sum in as the lower bound: all of the following are acceptable

$$\lim s_n = \lim \sum_{k=0}^{n} a_k = \sum_{k=0}^{\infty} a_k = \sum_{k\geq 0} a_k$$

Remark 12.2. Because the sum of any finitely many terms of a series is a finite number, we can remove any finite collection without changing whether or not the series converges. In particular, when proving convergence we are free to ignore the first finitely many terms when convenient. Because of this, we often will just write $\sum a_n$ when discussing a series, without giving any lower summation bound.

There are many important infinite series in mathematics: one that we encountered earlier is the Basel series first summed by Euler.

$$\sum_{n\geq 1}\frac{1}{n^2}=\frac{\pi^2}{6}$$

When the sequences a_n consists of functions of x, we may define an infinite series *function* for each x at which it converges. These describe some of the most important functions in mathematics, such as the Riemann zeta function

$$\zeta(s)=\sum_{n\geq 1}\frac{1}{n^s}$$

One of our big accomplishments to come in this class is to prove that exponential functions can be computed via infinite series, and in particular, the *standard exponential* of base e has a very simple expression

$$\exp(x)=\sum_{n\geq 0}\frac{x^n}{n!}$$

The other infinite algebraic expression we can conjure up is infinite products:

Definition 12.2 (Infinite Products). An infinite product p_n is a recursive sequence defined in terms of another sequence a_n by the recurrence relation $p_{n+1}=p_na_n$. Thus, the first terms of a series are

$$s_0=a_0, \qquad s_1+a_0a_1 \qquad s_2=a_0a_1a_2\ldots$$

We use *product notation* to denote the terms of a series:_

$$s_n=_0 a_1\cdots a_n=\prod_{k=0}^{n}a_k$$

Again, like for series, when such a sequence converges there are multiple common ways to write its limit:

$$\lim p_n=\lim\prod_{k=0}^{n}p_k=\prod_{k=0}^{\infty}p_k=\prod_{k\geq 0}p_k$$

The first infinite product to occur in the mathematics literature is Viete's Product for π

$$\frac{2}{\pi}=\frac{\sqrt{2}}{2}\frac{\sqrt{2+\sqrt{2}}}{2}\cdots$$

This product is derived from Archimedes' side-doubling procedure for the areas of circumscribed n-gons; hence the collections of nested roots!

Another early and famous example being Wallis' infinite product for $2/\pi$, which instead is derived from Euler's infinite product for the sine function.

$$\frac{\pi}{2}=\prod_{n\geq 1}\frac{4n^2}{4n^2-1}$$
$$=\frac{2}{1}\frac{2}{3}\frac{4}{3}\frac{4}{5}\frac{6}{5}\frac{6}{7}\frac{8}{7}\frac{8}{9}\frac{10}{9}\frac{10}{11}\frac{12}{11}\frac{12}{13}\frac{14}{13}\frac{14}{15}\cdots$$

In 1976, the computer scientist N. Pippenger discovered a modification of Wallis' product which converges to e:

$$\frac{e}{2}=\left(\frac{2}{1}\right)^{\frac{1}{2}}\left(\frac{2}{3}\frac{4}{3}\right)^{\frac{1}{4}}\left(\frac{4}{5}\frac{6}{5}\frac{6}{7}\frac{8}{7}\right)^{\frac{1}{8}}\left(\frac{8}{9}\frac{10}{9}\frac{10}{11}\frac{12}{11}\frac{12}{13}\frac{14}{13}\frac{14}{15}\frac{16}{15}\right)^{\frac{1}{16}}\cdots$$

Pippenger wrote up his result as a paper...but due to the relatively ancient tradition of mathematics he was adding to - he decided to write it in Latin! The paper appears as *"Formula nova pro numero cujus logarithmus hyperbolicus unitas est".* in IBM Research Report RC 6217. I am still trying to track down a copy of this! So if any of you are better at the internet than me, I would be very grateful if you could locate it.

Alluded to above, one of the most famous functions described by an infinite product is the sine function, which Euler expanded in his proof of the Basel sum

$$\frac{\sin\pi x}{\pi x}=\prod_{n\geq 1}\left(1-\frac{x^2}{n^2}\right)$$

AAs well as our friend the Riemann zeta function from above, which can be written as a product over all the primes! (Alluding to its deep connection to number theory)

$$\zeta(s)=\prod_{p\text{ prime}}\frac{1}{1-p^{-s}}$$

Perhaps in a calculus class you remember seeing many formulas for the convergence of series (we will prove them here in short order), but did not see many infinite products. The reason for this is that it is enough to study one class of these recursive sequences, once we really understand exponential functions and logarithms: we can use these to convert between the two. Because of this we too will focus most of our theoretical attention on series, though interesting products of historical significance will make several appearances.

12.0.1. Elementary Properties

To finish this introduction, we give several properties of infinite series which follow directly from their definition as *limits of sequences of partial sums.*

Definition 12.3 (Cauchy Criterion). A series $s_n = \sum a_n$ satisfies the Cauchy criterion if for every $\epsilon > 0$ there is an N such that for any $n, m > N$ we have

$$\left|\sum_m^n a_k\right| < \epsilon$$

Exercise 12.1. Prove a series satisfies the Cauchy criterion if and only if its sequence of partial sums is a Cauchy sequence.

Proposition 12.1 (The Addition of Series). *If $\sum a_n$ and $\sum b_n$ both converge, then the series $\sum(a_n + b_n)$ converges and*

$$\sum_{n\geq 0}(a_n + b_n) = \sum_{n\geq 0} a_n + \sum_{n\geq 0} b_n$$

Proof. For each N, let $A_N = \sum_{n=0}^N a_n$ and $B_N = \sum_{n=0}^N b_n$. Then the value of the infinite sums are $\sum_{n\geq 0} a_n = \lim A_N$ and $\sum_{n\geq 0} b_n = \lim B_N$. For any finite N, we can use the commutativity and associativity of addition to see

$$\begin{aligned}
\sum_{n=0}^N (a_n + b_n) &= (a_0 + b_0) + (a_1 + b_1) + \cdots + (a_N + b_N) \\
&= (a_0 + a_1 + \cdots + a_N) + (b_0 + b_1 + \cdots + b_N) \\
&= \sum_{n=0}^N a_n + \sum_{n=0}^N b_n \\
&= A_N + B_N
\end{aligned}$$

Since A_N and B_N both converge by assumption, we can apply the limit law for sums to see

$$\lim(A_N + B_N) = \lim A_N + \lim B_N$$

Putting this all together we have proven what we want, $\sum_{n\geq 0}(a_n + b_n)$ exists, and equals the sum of $\sum_{n\geq 0} a_n$ and $\sum_{n\geq 0} b_n$. □

Exercise 12.2 (Constant Multiple of a Series)**.** Prove that if $\sum a_n$ is a convergent series and $k \in \mathbb{R}$ a constant, then the series $\sum ka_n$ is convergent, and

$$\sum_{n\geq 0} ka_n = k\sum_{n\geq 0} a_n$$

Remark 12.3. Multiplying series is more subtle, as the terms of $\left(\sum_{n=0}^N a_n\right)\left(\sum_{n=0}^N b_n\right)$ are not just the pairwise products a_nb_n: we need to *multiply it all out*. The resulting construction is called the *Cauchy Product*, and we will later show that under the right conditions if $\sum a_n = A$ and $\sum b_n = B$ then the Cauchy product converges to AB.

Exercise 12.3. Let $\prod_{k\geq 1} a_k = \alpha$ and $\prod_{k\geq 1} b_k = \beta$ be convergent infinite products. Prove that $\prod_{k=1} a_kb_k$ converges, with limit $\alpha\beta$.

12.1. Telescoping Series

There are rare cases when we can sum a series directly, but these will prove very useful as *basic series* much as our *basic sequences* underlied much of our earlier work. The simplest way to directly sum a series is to find an exact formula for its partial sums, and *telescoping series* are a particularly nice example, where algebra makes this almost trivial

Definition 12.4 (Telescoping Series)**.** A telescoping series is a series $\sum a_n$ where the terms themselves can be written as *differences* of consecutive terms of another sequence, for example if $a_n=t_{n-1}-t_n.

Telescoping series are the epitome of a math problem that looks difficult, but is secretly easy. Once you can express the terms as differences, everything but the first and last cancels out! For example:

$$\begin{aligned}
s_n &= \sum_{k=1}^n a_k \\
&= \sum_{k=1}^n (t_k - t_{k-1}) \\
&= (t_n - t_{n-1}) + (t_{n-1} - t_{n-2}) + \cdots + (t_2 - t_1) + (t_1 - t_0) \\
&= t_n + (t_{n-1} - t_{n-1}) + \cdots + (t_2 - t_2) + (t_1 - t_1) - t_0 \\
&= t_n - t_0
\end{aligned}$$

Thus, evaluating the sum is just as easy as evaluating the limit of t_n:

$$\lim s_n = \lim(t_n - t_0) = (\lim t_n) - t_0$$

Thus, once a series has been identified as telescoping, often proving its convergence is straightforward: you get a direct formula for the partial sums, and then all that remains is to calculate the limit of a *sequence*. Because there are many ways a sequence might telescope its easier to look at examples than focus on the general theory.

Example 12.1. The sum $\sum_{k\geq 1}\frac{1}{k}-\frac{1}{k+1}$ telescopes. Writing out a partial sum s_n, everything collapses so $s_n = 1-\frac{1}{n+1}$.

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

Now we no longer have a series to deal with, as we've found the partial sums! All that remains is the sequence $s_n = 1-\frac{1}{n+1}$. And this limit can be computed immediately from the limit laws:

$$s = \lim s_n = 1 - \lim\frac{1}{n+1} = 1$$

Of course, sometimes a bit of algebra needs to be done to reveal that a series is telescoping:

Example 12.2. Compute the sum of the series

$$\sum_{n\geq 1}\frac{1}{n(n+1)}$$

Performing a partial fractions decomposition to $\frac{1}{n(n+1)}$ we seek A, B with $\frac{A}{n}+\frac{B}{n+1} = \frac{1}{n(n+1)}$ which is satisfied by $A=1, B=-1$, so

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

Thus our series telescopes, with partial sums

$$\sum_{n=1}^{N}\frac{1}{n(n+1)} = \left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\cdots+\left(\frac{1}{N-1}-\frac{1}{N}\right) = 1-\frac{1}{N}$$

Taking the limit

$$\sum_{n\geq 1}\frac{1}{n(n+1)} = \lim_N \sum_{n=1}^{N}\frac{1}{n(n+1)} = \lim 1 - \frac{1}{N} = 1$$

Telescoping series don't need to cancel *consecutive* terms, but rather it can take a bit of time before the telescoping begins:

Example 12.3. Compute the sum of the series

$$\sum_{k\geq 1}\frac{1}{k^2+3k}$$

Doing partial fractions to the term $1/(k^2+3k)$ we find

$$\frac{1}{k^2+3k} = \frac{2}{k(k+3)} = \frac{1}{3}\left(\frac{1}{k} - \frac{1}{k+3}\right)$$

We'll ignore the factor of $1/3$ while doing some scratch work below but be careful to bring it back later. Adding up the first two terms we don't see any cancellations like we expect of a telescoping series

$$\left(1-\frac{1}{3}\right)+\left(\frac{1}{2}-\frac{1}{4}\right)$$

But, after more terms the cancellations begin: the sixth term is

$$\left(1-\frac{1}{3}\right)+\left(\frac{1}{2}-\frac{1}{4}\right)+\left(\frac{1}{3}-\frac{1}{5}\right)+\left(\frac{1}{4}-\frac{1}{6}\right)+\left(\frac{1}{5}-\frac{1}{7}\right)+\left(\frac{1}{6}-\frac{1}{8}\right)$$
$$= 1+\frac{1}{2}-\frac{1}{7}-\frac{1}{8}$$

Seeing the pattern here, you can prove by induction that the N^{th} term is

$$\frac{1}{3}\left(1+\frac{1}{2}-\frac{1}{N+1}-\frac{1}{N+2}\right)$$

So, taking the limit as the number of terms we add goes to infinity we can use the limit laws together with $1/N \to 0$ to conclude

$$\sum_{k\geq 1}\frac{1}{k^2+3k} = \lim \frac{1}{3}\left(1+\frac{1}{2}-\frac{1}{N+1}-\frac{1}{N+2}\right)$$

$$= \frac{1}{3}\left(1+\frac{1}{2}-0-0\right) = \frac{1}{3}\frac{3}{2} = \frac{1}{2}$$

Exercise 12.4. Show that the following series is telescoping, and then find its sum

$$\sum_{n\geq 1}\frac{1}{4n^2-1}$$

Hint: factor the denominator, and do a partial fractions decomposition!

A telescoping product is defined analogously

Definition 12.5 (Telescoping Product)**.** A telescoping product is a product $\prod a_n$ where the terms themselves can be written as *ratios* of consecutive terms of another sequence, for example $a_n = \frac{t_n}{t_{n-1}}$.

Exercise 12.5. Find the value of the following infinte product by showing its telescoping and computing an exact formula for its partial sums:

$$\prod_{n\geq 5}\left(\frac{1}{n^2}\right)$$

An example of historical importance is below:

Example 12.4 (Viete's Product for π)**.** Viete's infinite product $\frac{2}{\pi} = \frac{\sqrt{2}}{2}\frac{\sqrt{2+\sqrt{2}}}{2}\cdots$ which we derived back in the introductory historical chapter to this text from an infinite application of the half angle formula, can also be derived as a *telescoping product*, where each term represents the ratio of the area of a circumscribed polygon and and its side-doubled cousin.

- The first term, $\sqrt{2}/2$ is the ratio of the area of an square to a octagon.
- The second term, $\sqrt{2+\sqrt{2}}/2$ is the ratio of the area of a octagon to an 16-gon.
- The n^{th} term is the ratio of the area of a 2^{n+1}-gon to a 2^{n+2}-gon.

When multiplying these all together, the intermediaries cancel, and in the limit this gives the ratio of the area of a square to the area of a circle.

12.2. Geometric Series

Definition 12.6. A series $\sum a_n$ is *geometric* if all consecutive terms share a common ratio: that is, there is some $r \in \mathbb{R}$ with $a_n/a_{n-1} = r$ for all n.

In this case we can see inductively that the terms of the series are all of the form ar^n. Thus, often we factor out the a and consider just series like $\sum r^n$.

Exercise 12.6 (Geometric Partial Sums)**.** For any real r, the partial sum of the geometric series is:

$$1 + r + r^2 + \cdots + r^n = \sum_{k=0}^{n} r^n = \frac{1 - r^{n+1}}{1 - r}$$

Like telescoping series, now that we have explicitly computed the partial sums, we can find the exact value by just taking a limit.

Theorem 12.1. *If $|r| < 1$ then $\sum r^n$ converges, and*

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1 - r}$$

Conversely if $|r| > 1$, the geometric series $\sum r^n$ diverges.

Proof. We begin with the case $|r| < 1$. By the partial sum formula, we have

$$\sum_{n \geq 0} r^n = \lim \sum_{k=0}^{n} r^n = \lim \frac{1 - r^{n+1}}{1 - r}$$

Since $|r| < 1$, we know that $r^n \to 0$, and so $r^{n+1} = rr^n \to 0$ by the limit theorems (or by truncating the first term of the sequence). Again by the limit theorems, we may then calculate

$$\lim \frac{1 - r^{n+1} 1 - r}{=} \frac{1 - \lim r^{n+1}}{1 - r} = \frac{1 - 0}{1 - r} = \frac{1}{1 - r}$$

For $|r| > 1$, we directly see the sum is unbounded as each $r^k > 1^k = 1$ so

$$s_N = \sum_{k=0}^{N} r^k > 1 + 1 + \cdots + 1 = N + 1$$

As convergent sequences are bounded, this must diverge. □

Exercise 12.7. Show if $r < -1$ that $\sum r^k$ diverges.

Hint: look at the subsequence $s_0, s_2, s_4, s_6 \ldots$ of partial sums.

Remark 12.4. Its often useful to commit to memory the formula also for when the sum starts at 1:

$$\sum_{k \geq 1} r^k = \frac{r}{1 - r}$$

Example 12.5. What should the infinite decimal $0.99999\cdots$ mean?

Because this holds for *all values of* r between -1 and 1, this gives us our first taste of a function defined as an infinite series. For any $x \in (-1, 1)$ we may define the function

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

and the argument above shows that $f(x) = 1/(1-x)$. Thus, we have two expressions of the same function: one in terms of an infinite sum, and one in terms of familiar algebraic operations. This sort of thing will prove extremely useful in the future, where switching between these two viewpoints can often help us overcome difficult problems.

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

The theory of geometric series began with Archimedes' famous paper *The Quadrature of the Parabola*, and we can now make his final argument rigorous in a modern form. (We will not make rigorous the first steps of the argument, which deal mainly in geometry, but re review them briefly here)

Archimedes' big idea was to divide a parabolic up into triangles recursively by drawing the largest triangle which inscribes in the segment. This divides the parabolic segment into a triangle and two *smaller* parabolic segments, on which the process repeats.

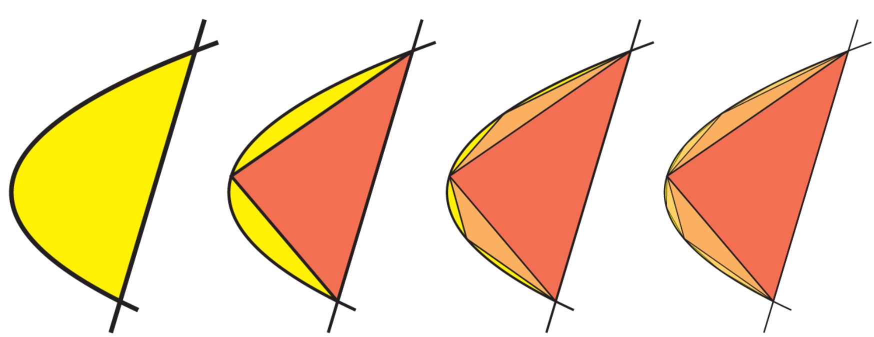


Figure 12.1.: Archimedes' infinite construction of the parabolic segment from triangles

Denote by T_n the sum of the areas of the triangles which appear at the $nth stage of this process (so T_0 is one triangle T_1 consists of two triangles, T_2 of four triangles, etc). Through clever use of the geometry of parabolas, archimedes shows that $T_{n+1} = \frac{1}{4}T_n$. And through further clever geometry, Archimedes argues that in the limit as $n \to \infty$, these triangles *completely fill the parabola*, so its area is the sum of their areas. That is

$$\text{Area} = \sum_{n\geq 0} T_n = \sum_{n\geq 0} \frac{T_0}{4^n} = T_0 \sum_{n\geq 0} \left(\frac{1}{4}\right)^n$$

Summing this geometric series yields the celebrated result:

Theorem 12.2. *The area of the segment bounded by a parabola and a chord is* $4/3^{rd}$*s the area of the largest inscribed triangle.*

12.3. Summing Reciprocals

Some of the most natural infinite series to consider are the sums of reciprocal natural numbers and their powers. The simplest of these is simply

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

and called the *Harmonic Series* (named after a distant connection to music). Other common examples are $\sum \frac{1}{n^2}$ or $\sum 1n^3$, etc. These arise everywhere throughout analysis, and find important applications in physics, number theory and beyond. In this final section of our introductory chapter we use what we've learned to calculate their values:

Theorem 12.3 (Divergence of the Harmonic Series)**.** *The harmonic series* $\sum_{n\geq 1} \frac{1}{n}$ *diverges.*

Proof. Let $s_N = \sum_{n=1}^{N} \frac{1}{N}$ denote the partial sums of the harmonic series, and note we have the following inequality relating s_{2N} with s_N:

$$\begin{aligned}
s_{2N} &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots + \frac{1}{2N-1} + \frac{1}{2N} \\
&= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6}\right) + \cdots + \left(\frac{1}{2N-1} + \frac{1}{2N}\right) \\
&> 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{6} + \frac{1}{6}\right) + \cdots + \left(\frac{1}{2N} + \frac{1}{2N}\right) \\
&= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N} \\
&= \frac{1}{2} + \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N}\right) \\
&= \frac{1}{2} + s_N
\end{aligned}$$

Now assume for the sake of contradiction that the harmonic series *does* converge, so $\lim s_N = L$ exists. Then all subsequences converge to the same limit, so restricting to the even subsequence, $\lim s_{2N} = L$ as well. But the inequality above ensures $s_{2N} > \frac{1}{2} + s_N$, and applying the limit theorems yields

$$\lim s_{2N} > \frac{1}{2} = \lim s_N \implies L > \frac{1}{2} + L$$

Subtracting L from both sides and multiplying by 2 gives $0 > 1$, a contradiction. □

Exercise 12.8. Give an alternate proof that the harmonic series $\sum \frac{1}{n}$ diverges, by comparing it with the partial sums of

$$1, 1/2, 1/4, 1/4, 1/8, 1/8, 1/8, 1/8, 1/16, \ldots$$

Hint: show that for each N the partial sum of the harmonic series is greater than the partial sums of this series. But then show the partial sums of this series are unbounded: for any integer k we can find a point where the sum surpasses k. This means the partial sums of the harmonic series are unbounded: but we know all convergent series are bounded! Thus the harmonic series cannot be convergent.

Theorem 12.4 (Convergence of the Reciprocal Squares)**.** *The series $\sum_{n\geq 1} \frac{1}{n^2}$ converges.*

Proof. Let s_N denote the partial sums of the series. Since $1/n^2 > 0$ for all n, we see the sequence s_N is monotone increasing for all N,

$$s_N = s_{N-1} + \frac{1}{N^2} > s_{N-1}$$

Thus to prove it converges we need only show its bounded above (and then apply the Monotone Convergence Theorem). As a first step, note that for every $n > 1$, we know $n-1 < n$ and so $\frac{1}{n^2} < \frac{1}{n(n-1)}$. Adding these up, we see

$$\sum_{n=2}^{N} \frac{1}{n^2} < \sum_{n=2}^{N} \frac{1}{n(n-1)}$$

This latter sum telescopes, so we can compute its partial sums directly:

$$\sum_{n=2}^{N} \frac{1}{n(n-1)} = \sum_{n=2}^{N} \left(\frac{1}{n-1} - \frac{1}{n}\right) = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{N-1} - \frac{1}{N}\right) = 1 - \frac{1}{N}$$

Thus, for all N we see $\sum_{n=2}^{N} \frac{1}{n^2} < 1 - \frac{1}{N^2} < 1$, so

$$\sum_{n=1}^{N} \frac{1}{n^2} = 1 + \sum_{n=2}^{N} \frac{1}{n^2} < 1 + 1 = 2$$

Together, our sequence of partial sums is monotone increasing and bounded above by 2, so its convergent. $\square$

While proving $\sum \frac{1}{n^2}$ is convergent was relatively straightforward, finding its value is what brought Lehonard Euler his first mathematical fame, when proved it equals exactly $\pi^2/6$.

Exercise 12.9. Prove that for $s \geq 2$ that $\sum \frac{1}{n^s}$ converges.

Hint: show its monotone; and shows its bounded by comparing partial sums with those of $\sum \frac{1}{n^2}$*. which we know converges.*

12.4. Problems

12.4.1. The Koch Fractal

The Koch Snowflake is a *fractal*, defined as the limit of an infinite process starting from a single equilateral triangle. To go from one level to the next, every line segment of the previous level is divided into thirds, and the middle third replaced with the other two sides of an equilateral triangle built on that side.

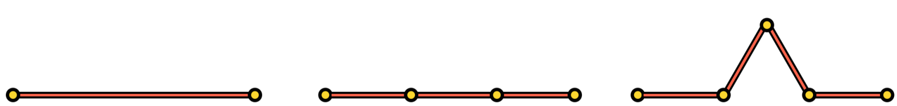


Figure 12.2.: The Koch subdivision rule: replace the middle third of every line segment with the other two sides of an equilateral triangle.

Doing this to *every line segment* quickly turns the triangle into a spiky snowflake like shape, hence the name. Denote by K_n the result of the n^{th} level of this procedure.

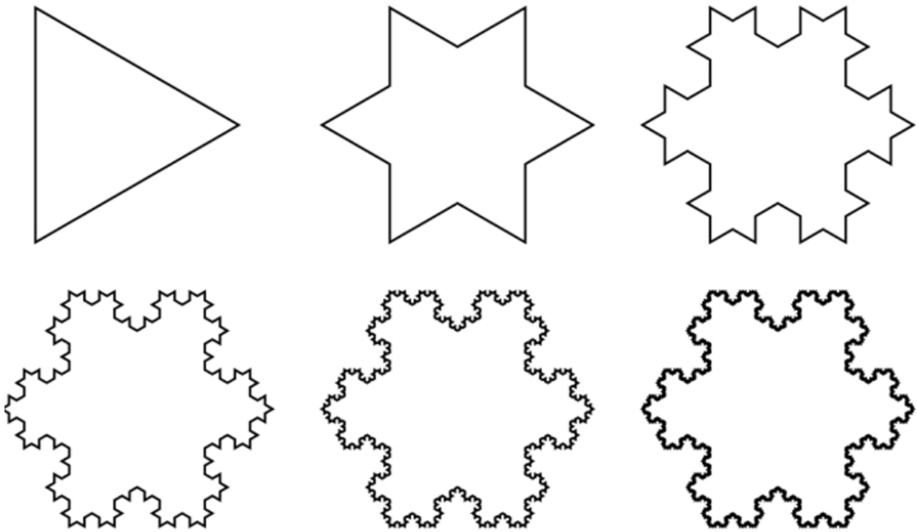


Figure 12.3.: The first six stages K_0, K_1, K_2, K_3, K_4 and K_5 of the Koch snowflake procedure. K_∞ is the fractal itself.

Say the initial triangle at level 0 has perimeter P, and area A. Then we can define the numbers P_n to be the perimeter of the n^{th} level, and A_n to be the area of the n^{th} level..

Exercise 12.10 (The Koch Snowflake Length)**.** What are the perimeters P_1, P_2 and P_3 of the first iterations? From this conjecture (and then prove by induction) a formula for the perimeter P_n and prove that P_n diverges. Thus, the limit cannot be assigned a length!

Next we turn to the area: recall that the area of an equilateral triangle can be given in terms of its side length as $A = \frac{\sqrt{3}}{2}s^2$

Exercise 12.11 (The Koch Snowflake Area)**.** What are the areas A_1, A_2 and A_3 in terms of the original area A? Find an infinite series that represents the area of the n^{th} stage A_n, and prove that your formula is correct by induction.

Now, use what we know about geometric series to prove that this converges: in the limit, the Koch snowflake has a finite area even though its perimeter diverges!

13. Convergence Tests

Highlights of this Chapter: Finding the value of a series explicitly is difficult, so we develop some theory to determine convergence without explicitly finding the limit. In particular, we provide a simple criterion to determine if a series diverges, formalize the technique of *comparison*, and develop the root and ratio tests as means of comparing with geometric series. We also introduce the notion of *absolute convergence*, and prove the convergence of *alternating series.*

In this section, we build up some technology to prove the convergence (and divergence) of series, without explicitly being able to compute the limit of partial sums. Such results will prove incredibly useful, as in the future we will encounter many theorems of the form *if* $\sum a_n$ *converges, then...* and we will need to a method of proving convergence to continue.

13.1. Divergent and Alternating Series

We begin with some low-hanging fruit: easy-to-check conditions on the terms of a series which either guarantee its convergence or divergence.

Corollary 13.1 (Divergence Test). *If a series* $\sum a_n$ *converges, then* $\lim a_n = 0$. *Equivalently, if* $a_n \not\to 0$ *then* $\sum a_n$ *diverges.*

Proof. Assume the sequence $s_N = \sum_{n=0}^{N} a_n$ of partial sums converges to some limit L. Then s_{N-1} also converges to L, as truncating the first term of a sequence doesn't affect convergence. Thus, we may use the limit law for subtraction to conclude the sequence $s_N - s_{N-1}$ converges, with limit zero. But

$$s_N - s_{N-1} = \sum_{n=0}^{N} a_n - \sum_{n=0}^{N-1} a_n = a_N$$

That is, a_N converges and $\lim a_N = 0$. □

Here's an alternative proof using the Cauchy criterion:

Proof. Let's apply the cauchy condition to the single value m. This says for all $\epsilon > 0$ there is some N where for $m > N$ we have

$$\left|\sum_{k=m}^{m} a_k\right| = |a_m| < \epsilon$$

But making $|a_m| < \epsilon$ for all $m > N$ is exactly the definition of $a_m \to 0$. □

Remark 13.1. The converse of the divergence test is **false** as the harmonic series $\sum \frac{1}{n}$ has terms going to zero, but we've already seen the overall sum diverges.

Happily, there are some minimal extra conditions one can add to *the terms go to zero* that do ensure convergence! The most famous such set of conditions is for the series to be *alternating* with terms converging monotonically to zero.

Theorem 13.1 (Alternating Series Test). *Let $a_n \to 0$ be a monotonically decreasing sequence. Then $\sum(-1)^n a_n$ converges.*

Proof. We will show the sequence $s_n = \sum_{k=0}^{n}(-1)^k a_k$ of partial sums converges by showing it is Cauchy. Looking at a concrete example

$$s_6 - s_{12} = (-1)^6 a_6 + (-1)^7 a_7 + (-1)^8 a_8 + (-1)^9 a_9 + (-1)^{10} a_{10} + (-1)^{11} a_{11} + (-1)^{12} a_{12}$$

$$= a_6 - a_7 + a_8 - a_9 + a_{10} - a_{11} + a_{12}$$

We can group the terms in two different ways to bound the difference $|s_6 - s_{12}|$: first,

$$= (a_6 - a_7) + (a_8 - a_9) + (a_{10} - a_{11}) + a_{12}$$

Here, each parentheses encloses a nonnegative term (since a_k is monotone decreasing), so the sum is nonnegative: $s_6 - s_{12} \geq 0$. But sliding our groupings down by one,

$$= a_6 + (-a_7 + a_8) + (-a_9 + a_{10}) + (-a_{11} + a_{12})$$

now each pair of parentheses includes a *non-positive* term, which we are taking away from a_6. Thus $s_6 - s_{12} \leq a_6$. Together these bounds imply $|s_6 - s_{12}| < a_6$. Indeed, one can confirm this holds generally, so for $m < n$, $|s_n - s_m|$ is bounded above by a_m. Using that $a_k \to 0$, for any $\epsilon > 0$ there's an N beyond which $0 < a_m < \epsilon$, and thus, for any $n > m > N$ we have

$$|s_n - s_m| < a_m < \epsilon$$

So, the sequence of partial sums is Cauchy, and thus convergent, as required. □

Exercise 13.1. Prove the general inequality used in the proof above: if $s_n = \sum_{k=0}^{n}(-1)^k a_k$ is the partial sums of a monotone decreasing sequence $a_n \to 0$, then

$$|s_n - s_m| \leq a_m \qquad \text{for all } m < n$$

Hint: think about different cases, when m, n are even and odd

Exercise 13.2. Give an alternative proof of the alternating series test, using the Nested Interval Theorem. Here's a potential outline:

Let $a_n \to 0$ be monotone decreasing and $s_n = \sum_{k=0}^{n}(-1)^k a_k$ be its partial sums.

- Show the *even subsequence* $s_0, s_2, s_4, \ldots$ is monotone decreasing
- Show the *odd subsequence* $s_1, s_3, s_5, \ldots$ is monotone increasing
- Show the intervals $[s_{2n-1}, s_{2n}]$ are nested, and their lengths are going to zero.
- Show there is a unique point in their intersection, and argue this is the limit of the partial sums s_n.

The monotonicity hypothesis of the Alternating Series test cannot be dropped, as the following example shows.

Example 13.1 (Monotonicity is Required)**.** Consider the infinite series

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

This series is alternating and its terms converge to zero, but it is *not* monotone. To see it diverges, look at the sum of two consecutive terms:

$$\frac{1}{\sqrt{n}-1} - \frac{1}{\sqrt{n}+1} = \frac{(\sqrt{n}+1)-(\sqrt{n}-1)}{n-1} = \frac{2}{n-1}$$

Thus, if we add up the first $2N$ terms of the series we get

$$\left(\frac{1}{\sqrt{2}-1} - \frac{1}{\sqrt{2}+1}\right) + \left(\frac{1}{\sqrt{3}-1} - \frac{1}{\sqrt{3}+1}\right) + \cdots + \left(\frac{1}{\sqrt{N}-1} - \frac{1}{\sqrt{N}+1}\right)$$

$$= \left(\frac{2}{1}\right) + \left(\frac{2}{2}\right) + \cdots + \left(\frac{2}{N-1}\right)$$

$$= 2\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-1}\right)$$

This is twice the sum of the first $n-1$ terms of the harmonic series, which we know diverges! Thus our series contains a subsequence of partial sums that diverges, and must diverge as well.

13.2. Absolute vs Conditional Convergence

Below we will develop several theorems that apply exclusively to *series of positive terms.* That may seem at first to be a significant obstacle, as many series involve both addition and subtraction! So, we take some time here to assuage such worries, and provide a means of probing a general series using information about its nonnegative counterpart.

Definition 13.1 (Absolute Convergence)**.** A series $\sum a_n$ converges absolutely if the associated series of absolute values $\sum |a_n|$ is convergent.

Of course, such a definition is only *useful* if facts about the nonnegative series imply facts about the original. Happily, that is the case.

Theorem 13.2 (Absolute Convergence Implies Convergence)**.** *Every absolutely convergent series is a convergent series.*

Proof. Let $\sum a_n$ be absolutely convergent. Then $\sum |a_n|$ converges, and its partial sums satisfy the Cauchy criterion. This means for any ϵ we can find an N where

$$|a_n| + |a_{n+1}| + \cdots + |a_m| < \epsilon$$

But, by the triangle inequality we know that

$$|a_n + a_{n+1} + \cdots + a_n| \leq |a_n| + |a_{n+1}| + \cdots + |a_m|$$

Thus, our original series $\sum a_k$ satisfies the Cauchy Criterion, as

$$\left|\sum_{k=m}^{n} a_k\right| < \epsilon$$

And, since Cauchy is equivalent to convergence, this implies $\sum a_k$ is a convergent series. □

Definition 13.2. A series converges *conditionally* if it converges, but is not absolutely convergent.

Such series caused much trouble in the foundations of analysis, as they can exhibit rather strange behavior. We met one such series in the introduction, the alternating sum of $1/n$ which seemed to converge to different values depending on the order we added its terms. Here we begin an investigation into such phenomena.

13.3. Comparison

One of the very most useful convergence tests for a series is comparison. This lets us show that a series we care about (that may be hard to compute with) converges or diverges by comparing it to a simpler series - much like the squeeze theorem did for us with sequences. This theorem gives less information than the squeeze theorem (it doesn't give us the exact value of the series we are interested in) but it is also easier to use (it only requires a bound, not an upper and lower bound with the same limit).

Theorem 13.3 (Comparison For Series). *Let $\sum a_n$ and $\sum b_n$ be two series of nonnegative terms, with $0 \leq a_n \leq b_n$.*

- *If $\sum b_n$ converges, then $\sum a_n$ converges.*
- *If $\sum a_n$ diverges, then $\sum b_n$ diverges.*

The proof is just a rehashing of our old friend, Monotone Convergence.

Proof. We prove the first of the two claims, and leave the second as an exercise. If $x_n \geq 0$ then the series $s_n = \sum_{k=0}^{n} x_k$ is monotone increasing (as by definition $s_n = s_{n-1} + x_n$ and $x_n \geq 0$ we see $s_n \geq s_{n-1}$ for all n).

Thus. $\sum a_n$ and $\sum b_n$ are monotone sequences. If $\sum b_n$ converges, we know by the Monotone Convergence Theorem that it its limit β is the supremum of the partial sums, so for all n

$$\sum_{k=0}^{n} b_k \leq \beta$$

But, since $a_k \leq b_k$ for all k, we see the same is true of the partial sums

$$\sum_{k=0}^{n} a_k \leq \sum_{k=0}^{n} b_k$$

Stringing these inequalities together, we see that $\sum a_k$ is bounded above by β. Since it is monotone (as the sum of nonnegative terms) as well, Monotone convergence assures us that it converges, as claimed. □

Exercise 13.3. Let $\sum a_n$ and $\sum b_n$ be two series of nonnegative terms, with $0 \leq a_n \leq b_n$. Prove that if $\sum a_n$ diverges, then $\sum b_n$ diverges.

A very effective means of proving the convergence or divergence of certain series is to compare them with geometric series, who we understand completely. Such comparisons will only work if the terms of our series are shrinking fast enough (quicker than a geometric progression, so at least exponentially), and coarse methods like this are bound to prove unhelpful for various particular examples. Nonetheless the ease of use of such comparisons is unparalleled, making them an essential element of our toolkit.

13.3.1. The Root Test

Geometric series are particularly simple to work with, as the n^{th} term is just a constant raised to the n^{th} power. Said another way, the n^{th} root of the n^{th} term is *constant*! This suggests a concrete way to compare a series with a geometric series: comparing the roots of its terms with a constant. This is commonly known as the *root test.*

Proposition 13.1 (The Root Test)**.** *Let $\sum a_n$ be a series, and assume that $\lim \sqrt[n]{|a_n|}$ converges to some limit L. Then*

- *$\sum a_n$ converges if $L < 1$*
- *$\sum a_n$ diverges if $L > 1$*

Proof. We work in cases, starting with $L < 1$. Here, for any r strictly between L and 1 we can eventually bound our sequence $\sqrt[n]{|a_n|}$ above by r: (set $\epsilon = r - L$, then there's an N beyond which $|\sqrt[n]{|a_n|} - L| < \epsilon$, so $\sqrt[n]{|a_n|} < L + \epsilon = r$).

Taking the n^{th} power of both sides, we see that for all $n \geq N$ we have $|a_n| < r^n$, and so we can compare the series $\sum_{n \geq N} |a_n|$ with the geometric series $\sum_{n \geq N} r^n$, which converges as $|r| = r < 1$. Thus $\sum_{n \geq N} |a_n|$ converges, and as the first finitely many terms of a series do not affect convergence, $\sum_{n \geq 0} |a_n|$ converges as well. But now we are done: this is exactly the statement that $\sum a_n$ converges absolutely, and absolute convergence implies convergence.

If $L > 1$ we can use the same sort of reasoning for any $1 < r < L$ to eventually bound our sequence $\sqrt[n]{|a_n|}$ to be *greater than* r, so that $|a_n| > r^n$ for all $n \geq N$. This sets up a comparison with a divergent geometric series which causes problems. But even more directly, if $|a_n| > r^n$ and $r^n > 1$, then $|a_n| > 1$ for all sufficiently large n, and its impossible that $a_n \to 0$. Thus $\sum a_n$ diverges by the *Divergence Test.* $\square$

Example 13.2 (Using the Root Test)**.** Consider the series

$$\sum_{n=1}^{\infty} \left(\frac{2n}{3n+1} \right)^n .$$

Applying the root test means we must take the n^{th} root of the absolute value of $a_n = \left(\frac{2n}{3n+1} \right)^n$. Since for all $n \in \mathbb{N}$ these terms are positive, this is

$$\sqrt[n]{|a_n|} = \frac{2n}{3n+1}.$$

Using the limit laws, we can check this limit exists:

$$\lim \frac{2n}{3n+1} = \frac{2}{3}$$

Since $\frac{2}{3} < 1$, the root test guarantees that the series $\sum_{n=1}^{\infty}\left(\frac{2n}{3n+1}\right)^n$ converges absolutely, and thus converges.

Remark 13.2. When $\lim \sqrt[n]{|a_n|} = 1$ we gain no information, as there are both convergent series (for example $\sum \frac{1}{n^2}$) and divergent series (for example, $\sum \frac{1}{n}$) with this property. Recalling that $n^{1/n} \to 1$, we see

$$\lim \sqrt[n]{\frac{1}{n}} = \lim \frac{1}{\sqrt[n]{n}} = \frac{1}{\lim n^{1/n}} = 1$$

$$\lim \sqrt[n]{\frac{1}{n^2}} = \lim \frac{1}{\sqrt[n]{n^2}} = \frac{1}{\lim n^{2/n}} = \frac{1}{\left(\lim n^{1/n}\right)^2} = 1$$

The root test is very powerful when it applies, but one of its hypotheses is that the limit $\lim \sqrt[n]{|a_n|}$ must exist. This is a rather big ask: and remembering the limsup one might wonder if we could instead prove an analog of the root test which looks at $\limsup \sqrt[n]{|a_n|}$ instead, as we know *this quantity always exists* so long as the sequence is bounded. Indeed we can, and after refreshing our memories of the definition of limsup, we see the proof barely changes either!

Exercise 13.4 (The Root Test with Limsup)**.** Prove that if $\sum a_n$ is a series and $L = \limsup \sqrt[n]{|a_n|}$, the series converges if $L < 1$ and diverges for $L > 1$. (Here we write $L = \infty$ as a shorthand to mean the sequence $\sqrt[n]{|a_n|}$ is unbounded.)

This more general version will be crucial to our understanding of general power series in the next chapter. It is still not *exact*, because it does not give any information when the limsup is exactly 1; but this is the only case of ambiguity

Corollary 13.2. *If* $\sum a_n x^n$ *is a convergent series, then* $\limsup \sqrt[n]{|a_n|} \leq 1$.

Proof. We proceed by process of elimination. If the $\limsup$ were not ≤ 1 then it would either be (1) greater than 1, or (2) not exist. In the first case, the root test directly proves the series diverges. In the second case, $\limsup$ exists for every bounded sequence, so this means the sequence of n^{th} roots is unbounded. Thus the original sequence of terms is also unbounded, so the series diverges (by the divergence test, as the terms aren't going to zero). $\square$

Below is an example in practice where the need for $\limsup$, over the more familiar $\lim$ may arise.

Example 13.3 (Using the Limsup Root Test)**.** Consider the series $\sum_{n=1}^{\infty}\left(\frac{1+(-1)^n}{3}\right)^n$. Computing the n^{th} roots of the absolute values required for the root test,

$$\sqrt[n]{\left|\left(\frac{1+(-1)^n}{3}\right)^n\right|} = \left|\frac{1+(-1)^n}{3}\right| = \begin{cases}\frac{2}{3}, & \text{if } n \text{ is even,}\\ 0, & \text{if } n \text{ is odd.}\end{cases}$$

The limit of this sequence does not exist, because it oscillates between two values. Rigorously, the even subsequence converges to $2/3$ and the odd subsequence converges to 0 (both are constant sequences), but for a convergent sequence all subsequences converge to the same limit. Thus, the original form of the root test does not apply, as the limit does not even exist.

However, the limsup does exist. Let r_n be the sequence of roots, so $r_n = 2/3$ if n is even and 0 if n is odd. Then for all N

$$\sup_{n\geq N} r_n = \sup\{\tfrac{2}{3}, 0, \tfrac{2}{3}, 0, \tfrac{2}{3}, \ldots\} = \frac{2}{3}$$

And thus

$$\limsup r_n = \lim_N \sup_{n\geq N}\{r_n\} = \frac{2}{3}$$

as its the limit of a constant sequence.

Because $2/3 < 1$, the limsup version of the root test applies without issue, proving that our series converges absolutely, and hence converges (even though our original test was not strong enough to say anything at all).

However, while good for building theory, using the root test *in practice* is rather annoying - nobody wants to be computing limits of n^{th} roots of arbitrary things! So its beneficial to look for another, perhaps simpler method of comparing with a geometric series, as we do below.

13.3.2. The Ratio Test

Consecutive terms of a geometric series $\sum r^n$ have the common ratio r. Thus a natural means of comparing with a geomeric series is to investigate the common ratio a_{n+1}/a_n of a series' terms:

Proposition 13.2 (The Ratio Test)**.** *Let $\sum a_n$ be a series and assume that the sequence of ratios $\left|\frac{a_{n+1}}{a_n}\right|$ converges, with limit L. Then $\sum a_n$ converges if $L < 1$ and $\sum a_n$ diverges when $L > 1$.*

Proof. We first consider the case $L < 1$. For any r with $L < r < 1$, since $\lim |a_{n+1}/a_n| < r$, our sequence is eventually less than r (setting $\epsilon = r - L$ there is an N beyond which $||a_{n+1}/a_n - L| < \epsilon$, or $|a_{n+1}/a_n| < L + \epsilon = r$). But this means that our sequence is eventually shrinking by a factor of r with each consecutive term:

$$\left|\frac{a_{n+1}}{a_n}\right| < r \implies |a_{n+1}| < r|a_n| \qquad \text{for } n \geq N$$

Thus, beyond N our series is bounded by a geometric series:

$$|a_{N+1}| < r|a_N| \qquad a_{N+1} < r|a_{N+1}| < r^2|a_N| \qquad |a_{N+3}| < r|a_{N+2}| < r^3|a_N|$$

continuing inductively yields immediately that for all k, $|a_{N+k}| < |a_N|r^k$. Thus, by comparison our series $\sum_{n\geq N} |a_n|$ converges:

$$\sum_{n\geq N} |a_n| = \sum_{k\geq 0} |a_{N+k}| < \sum_{k\geq 0} |a_N|r^k = |a_N| \sum_{k\geq 0} r^k$$

□

Exercise 13.5. Prove the divergence of $\sum a_n$ in the case $\lim |a_{n+1}/a_n| > 1$.

Example 13.4. Consider the series

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

The n^{th} term is defined by $a_n = \frac{3^n}{n!}$, so applying the ratio test requires us to compute $\left|\frac{a_{n+1}}{a_n}\right|$. Since all terms involved are positive, we can drop the absolute values and compute

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{3^{n+1}/(n+1)!}{3^n/n!} = \frac{3^{n+1} \cdot n!}{3^n \cdot (n+1)!}.$$

Simplifying,

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{3^{n+1}}{3^n} \cdot \frac{n!}{(n+1)!} = \frac{3 \cdot 3^n}{3^n} \cdot \frac{1}{n+1} = \frac{3}{n+1}.$$

Now, taking the limit using our limit laws and basic limits

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

Since $0 < 1$, the ratio test guarantees that the series $\sum_{n=1}^{\infty} \frac{3^n}{n!}$ converges absolutely, and thus converges.

Remark 13.3. When $\lim\left|\frac{a_{n+1}}{a_n}\right| = 1$ we gain no information, as there are both convergent series (for example $\sum \frac{1}{n^2}$) and divergent series (for example, $\sum \frac{1}{n}$) with this property.

$$\lim \frac{\frac{1}{n+1}}{\frac{1}{n}} = \lim \frac{n}{n+1} = \lim \frac{1}{1+\frac{1}{n}} = \frac{1}{1+\lim \frac{1}{n}} = 1$$

$$\lim \frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}} = \lim \left(\frac{n}{(n+1)}\right)^2 = \left(\lim \frac{n}{(n+1)}\right)^2 = 1^2 = 1$$

Like the root test, one might seek a version of the root test which doesn't require the limit $\lim |a_{n+1}/a_n|$ to converge. Again, we succeed by replacing lim with lim sup and barely modifying the proof.

Exercise 13.6 (The Ratio Test with Limsup)**.** Prove that if $\sum a_n$ is a series and $L = \limsup \left|\frac{a_{n+1}}{a_n}\right|$, the series converges if $L < 1$ and diverges for $L > 1$.

Example 13.5. Example using ratio test + limsup

The ratio test, while easy to apply, has some obvious failure modes even in this more general version. What if some of the terms being added up are *zero*, so that consecutive ratios are undefined? (For example, if every other term of the series is zero, then the consecutive ratios alternate between being zero and undefined, completely independently of the values of the nonzero terms). One might be tempted to try and fix this problem by re-indexing; removing all terms that are zero before applying the ratio test. While this would remove some problems, the fact still remains that comparing consecutive ratios just isn't that fine-grained of a tool to work with, and we can't take this as a *one size fits all* tool.

14. Power Series

Highlights of this Chapter: we introduce the definition of a power series, and testing for convergence via ratios.

We have developed some pretty powerful tools to prove the convergence of infinite series: for example, with little work we can apply the ratio test to immediately see $\sum \frac{n}{2^n}$ converges.

Indeed the same technique shows that the more general series $\sum \frac{n\alpha^n}{2^n}$ converges whenever $|\alpha| < 2$. This invites a bit of a change in perspective: we might think of the series above as a *function* that takes in one value of α (a parameter), and outputs the limit of the series. From this perspective, our calculation above actually helped us find the *domain* of the function - all the values of α that make sense to plug in.

Defining functions using sequences and series proves to be an incredibly powerful tool in analysis, and so we take a break from our more theoretical development to introduce the simplest and most useful case: power series. Polynomials or *finite* sums of multiples of powers of x, are some of the simplest functions we know. So its natural to wonder about their *infinite* counterparts: *power series*, arising as the limit of a sequence of polynomials of increasing degree

Definition 14.1 (Power Series)**.** A power series is a function defined as the limit of a sequence of polynomials

$$f(x) = \sum_{n\geq 0} a_n x^n$$

for a sequence a_n of real numbers. For each x, this defines an infinite series; the domain of a power series is the subset $D \subset \mathbb{R}$ of x values where the series converges.

Of course, polynomials themselves are a special case of power series, with $a_n = 0$ after some finite N. Perhaps the second simplest power series is the one with $a_n = 1$ for all n:

$$f(x) = 1 + x + x^2 + x^3 + x^4 + \cdots + x^n + \cdots$$

This is none other than the *geometric series in x*! So, it converges whenever the common ratio x satisfies $|x| < 1$: its domain is the interval $(-1, 1)$.

Power series are an extremely versatile tool to reach beyond the arithmetic of polynomials, while staying close to the fundamental operations of addition/subtraction and multiplication/division. One of our main uses of them will be to provide efficient means of computing important functions (exponentials, logs, trigonometric functions, etc).

Definition 14.2 (Power Series Representation)**.** Given a function $f(x)$, a power series representation of f is a series $s(x) = \sum a_n x^n$ such that $s(x) = f(x)$ whenever $s(x)$ converges.

Example 14.1. The function $f(x) = \frac{1}{1-x}$ has a power series representation on the interval $(-1, 1)$, where

$$\sum_{n\geq 0} x^n = \frac{1}{1-x}$$

This might not *seem* like an exciting discovery, as $1/(1-x)$ is a function that is easy for us to compute using the field operations, and now we've build a *more complicated looking expression* - an infinite series - to compute the same value! The ability to represent a function as a power series will be much more useful when looking at functions that are difficult to understand, like the exponential. Our best procedure for computing an irrational power right now is the definition: the limit of a sequence of rational powers. But such a limit is ridiculously hard to compute in practice. If we could instead represent the exponential as a power series, we could replace this limit with a series made out of just addition and multiplication! We will do exactly this, in a future chapter.

14.1. Convergence

Here we will study the most general theory of power series $\sum a_n x^n$ for arbitrary sequences a_n. The first thing to understand is their *domain*: for which values of x do the series converge? One point stands out immediately: for $x = 0$ the terms $a_n x^n$ are all equal to zero, so the partial sums of $\sum a_n x^n$ are $0 + 0 + 0 \cdots + 0 = 0$, and so the series converges to 0. Thus, $x = 0$ is in the interval of convergence of every power series.

Proposition 14.1. *If a power series $\sum a_n x^n$ converges at some $u > 0$, then it converges at all $x \in (-u, u)$.*

Proof. If $\sum a_n u^n$ converges, then by the root test $\limsup \sqrt[n]{|a_n u^n|} \leq 1$ (by process of elimination: if this sequence were unbounded or has limsup > 1 then we know it diverges). Now $x \in (-u, u)$ means $|x| < |u|$. Noting that $x = (x/u)u$, we may rewrite the limit we want as follows:

$$\sqrt[n]{|a_n x^n|} = |x|\sqrt{n}|a_n| = \frac{|x|}{|u|}|u|\sqrt[n]{|a_n|} = \frac{|x|}{|u|}\sqrt[n]{|a_n u^n|}$$

Because $|x|/|u|$ is a positive constant, we can pull it out of the limsup and compute

$$\limsup \frac{|x|}{|u|}\sqrt[n]{|a_n u^n|} = \frac{|x|}{|u|}\limsup \sqrt[n]{|a_n u^n|}$$

This is strictly less than 1 as $|x|/|u|$ is strictly less than 1 (since $|x| < |u|$) and the limsup term is less than or equal to 1 by our assumption of convergence. Thus, with limit less than 1, this series converges absolutely by the root test, so $\sum a_n x^n$ converges as claimed. □

This motivates the following definition:

Definition 14.3 (Radius of Convergence)**.** The radius of convergence of a power series is the largest value of $r > 0$ such that the series converges on $(-r, r)$.

With a more careful use of our convergence tests, we can compute an exact expression for the radius of convergence in terms of the coefficients. This will be of incredible theoretical utility in the chapters to come, and is often known as the *Cauchy Hadamard Theorem.*

Theorem 14.1 (Finding the Radius of Convergence: Cauchy-Hadamard)**.** *Let $\sum a_n x^n$ be a power series, and $\alpha = \limsup \sqrt[n]{|a_n|}$. Then the radius of convergence is $r = 1/\alpha$ (where $\alpha = 0$ means convergence on all of $\mathbb{R}$), and if $\sqrt[n]{|a_n|}$ is unbounded (so its* lim sup *is undefined, or infinite) then $\sum a_n x^n$ converges only at $x = 0$.*

Proof. First we deal with the generic case, where α is a nonzero real number. Since $\limsup \sqrt[n]{|a_n x^n|} = |x| \limsup \sqrt[n]{|a_n|} = |x|\alpha$, we know this series converges whenever $|x|\alpha < 1$, so $|x| < 1/\alpha$. Since the root test further ensures the series diverges whenever $|x|\alpha > 1$, $r = 1/\alpha$ is the *largest* number such that the series converges for all $|x| < r$, making it the radius of convergence.

The same reasoning works when $\alpha = 0$, as now for any x we see $\limsup \sqrt[n]{|a_n x^n|} = |x| \limsup \sqrt[n]{|a_n|} = |x|\alpha = 0$, which is less than 1: thus $\sum a_n x^n$ converges for all $x \in \mathbb{R}$.

Finally, if $\sqrt[n]{|a_n|}$ is an unbounded sequence, then for any nonzero $x \neq 0$ the sequence $|x|\sqrt[n]{|a_n|} = \sqrt[n]{|a_n x^n|}$ is also unbounded. Taking n^{th} powers shows $|a_n x^n|$ is unbounded, and in particular does not converge to zero. Thus $a_n x^n$ does not converge to zero, so the series diverges by the divergence test. Since all series converge for $x = 0$, this makes $\{0\}$ the entire domain, so the radius of convergence is zero. □

Corollary 14.1 (Absolute Convergence of Power Series)**.** *Let $f(x) = \sum a_n x^n$ be a power series with radius of convergence r, and let $u \in (-r, r)$. Then f converges absolutely at u.*

The root test further implies that for a power series with radius of convergence r, at any x with $|x| > |r|$ the series must diverge. Thus, the domain of a power series has a very limited form: it must be an interval centered at zero. Because of this we often call the domain the *interval of converence.*

Definition 14.4 (Interval of Convergence)**.** The interval of convergence of a power series $\sum a_n x^n$ is another word for its domain, the set of all x for which the series converges.

Exercises below show there are no further restrictions: any possible interval centered at zero is the interval of convergence for some power series: so domains can take the form $\{0\}, (-r, r), [-r, r], [-r, r), (-r, r]$, or $(-\infty, \infty)$; all of $\mathbb{R}$.

While the root test is of great theoretical utility to proving the above theorem, in practice n^{th} roots are rather unweildly to work with, and we might wish to instead apply hte *ratio test.* One can make a version of the ratio test specifically adapted to power series as follows:

Exercise 14.1. Let $\sum a_n x^n$ be a power series, and assume the sequence of ratios $\frac{a_{n+1}}{a_n}$ converges to some $\alpha \in \mathbb{R}$. Prove that the radius of convergence is $r = 1/\alpha$ when $\alpha \neq 0$, and converges on all of $\mathbb{R}$ when $\alpha = 0$.

14.1.1. Skipping Terms

The root and ratio tests as proven above apply to series $\sum a_n x^n$, where the n^{th} term is an n^{th} power of x. But there are very many natural series not of this form: for example, a series with only even powers of x,

$$1 + \frac{x^2}{4} + \frac{x^4}{16} + \frac{x^6}{64} + \cdots = \sum_{n \geq 0} \frac{x^{2n}}{4^n}$$

Or only odd powers, or only powers that are multiples of three, etc. It might be tempting to directly apply the root test (for example) to the coefficients of such a series and conclude

$$\sqrt[n]{1/4^n} = 1/4$$

So the series converges for $|x| < 4$. However this would be wrong! A more careful application of the root test to the entire series shows

$$\sqrt[n]{x^{2n}/4^n} = |x^2|/4$$

So we actually need $x^2/4 < 1$, so $x^2 < 4$ so $|x| < 2$. Being careful and applying the root test / ratio test *to the entire series* each time (instead of the shortcut where its applied to the coefficients) will avoid trouble in such cases.

14.2. Problems

14.2.1. Example Power Series

Power series provide us a means of describing functions via explicit formulas that we have not been able to thus far, by allowing a limiting process in their definition. For instance, we will soon see that the power series below is an *exponential function.*

Exercise 14.2. Show the power series $\sum \frac{x^n}{n!}$ converges for all $x \in \mathbb{R}$.

When a power series converges on a finite interval, its behavior at each endpoint may require a different argument than the ratio test (as that will give 1, and tell you nothing)

Example 14.2. Show the power series $\sum \frac{x^n}{n}$ has domain $[-1, 1)$. This shows that its possible for power series to have a domain which is closed on the left side and open on the right side of the interval of convergence.

To finish showing *all possible interval types exist*, create an example of a power series which converges on an interval of the form $(-a, a]$ for some $a > 0$ (and prove your claim).

Exercise 14.3. Show the power series $\sum \frac{x^n}{n^2}$ has domain $[-1, 1]$.

When the radius of convergence is 0, the power series converges at a single point:

Exercise 14.4. Show the power series $\sum n!x^n$ diverges for all $x \neq 0$.

Exercise 14.5. Series $\sum 2^n x^n$ converges on $[-1/2, 1/2)$. *Hint: substitution* $y = 2x$

Example 14.3. Where does $\sum 2^n x^{3n}$ converge? *Trickier! Need to worry about the exponents not being just n*

15. Advanced Techniques

The theory of infinite series is both deep and rich: there is much more we could say in this short chapter. With an eye towards calculus however we must march onwards, and so in this final chapter we collect some odds and ends about series that will prove useful in that quest. In particular, we prove the *dominated convergence theorem* allowing one to work simultaneously with limits and infinite sums, as well as *Abel's Lemma* - an application of *summation by parts.*

15.1. Switching Limits and Sums

It's rather common in practice to end up with an *infinite sequence of infinite series.* For example, if $f(x) = \sum_k a_k x^k$ is a power series, then one might be interested in evaluating this function along a sequence x_n of inputs values within its interval of convergence. This would produce the sequence of values $f(x_n) = \sum_k a_k (x_n)^k$, and if $x_n \to x$, its natural to wonder if $\sum_k a_k (x_n)^k \to \sum_k a_k x^k$. Unpacking this, we are asking if

$$\lim_n \sum_k a_k (x_n)^k = \sum_k a_k (\lim_n x_n)^k$$

By the limit laws this is the same as asking whether $\lim_n \sum_k a_k (x_n)^k = \sum_k \lim_n a_k (x_n)^k$, and so more abstractly we are asking if we can switch the order of a limit and a sum.

QUESTION: If $a_{n,k}$ is a sequence depending on two variables n and k, when is $\lim_n \sum_{k\geq 0} a_{n,k} = \sum_{k\geq 0} \lim_n a_{n,k}$?

Unfortunately this is subtle: its sometimes true and sometimes false:

Example 15.1 (When you can switch a limit and a sum)**.** Consider the geometric series $\sum_{k\geq 0} x^k$, and let a_n be a convergent sequence in $(-1, 1)$, with limit $a \in (-1, 1)$. For each fixed n the series $\sum_{k\geq 0} (a_n)^k$ converges, and has limit $\frac{1}{1-a_n}$ by the formula for a geometric series. Thus, taking the limit as $n \to \infty$ of these sums yields

$$\lim_n \sum_{k\geq 0} (a_n)^k = \lim_n \frac{1}{1-a_n} = \frac{1}{1-\lim a_n} = \frac{1}{1-a}$$

by the limit theorems. But this is precisely the value of $\sum_{k\geq 0} a^k = \frac{1}{1-a}$. And again, looking at each term a^k individually, by the limit theorems

$$a^k = (\lim_n a_n)^k = \lim_n (a_n)^k$$

Putting these together, we see $\sum_{k\geq 0} \lim_n (a_n)^k = \frac{1}{1-a}$ and so

$$\lim_n \sum_{k\geq 0} (a_n)^k = \frac{1}{1-a} = \sum_{k\geq 0} \lim_n (a_n)^k$$

Example 15.2 (When you can't switch a limit and a sum)**.** Written without summation notation, consider the following

$$\begin{aligned} 1 &= \frac{1}{2} + \frac{1}{2} \\ &= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} \\ &= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} \end{aligned}$$

Each row sums to 1, but the limit of each *term* $1/2^n \to 0$. So, if we took the limit of the terms first, we would get $1 = 0 + 0 + 0 + \cdots + 0 = 0$. Nonsense! Writing this precisely in summation notation, we define

$$a_{n,k} = \begin{cases} 1/2^n & 0 \leq k < 2^n \\ 0 & \text{else} \end{cases}$$

Then each of the rows above is the sum $1 = \sum_{k\geq 0} a_{n,k}$ for $n = 2, 3, 4$. Since this is constant it is true that the limit is 1, but it is *not true* that the limit of the sums is the sum of the limits, which is zero.

$$1 = \lim_n \sum_{k\geq 0} a_{n,k} \neq \sum_{k\geq 0} \lim a_{n,k} = \sum_{k\geq 0} 0 = 0$$

So, its hopefully clear that to be able to use series in realistic contexts, we are in desperate need of a theorem which tells us when we can interchange limits and summations. The precise theorem giving these conditions is sometimes called *Tannery's Theorem*, but we shall refer to it by its more descriptive name, *Dominated Convergence*.

Theorem 15.1 (Dominated Convergence for Series)**.** *Let $a_{n,k}$ be a double sequence such that $\lim_n a_{n,k}$ exists for each k, and $\sum_{k\geq 0} a_{n,k}$ converges for each n. Then if*

- *There is an M_k with $|a_{n,k}| \leq M_k$ for all n.*
- *$\sum M_k$ is convergent.*

It follows that $\sum_k \lim_n a_{n,k}$ *is convergent, and*

$$\lim_n \sum_k a_k(n) = \sum_k \lim_n a_{n,k}$$

Proof. For simplicity of notation define $a_{\infty,k} = \lim_n a_{n,k}$. First, we show that $\sum_k a_{\infty,k}$ converges. Since for all n, $|a_{n,k}| \leq M_k$ we know this remains true in the limit, so $\lim_n |a_{n,k}| = |a_{\infty,k}| < M_k$. Thus, by comparison we see $\sum_k |a_{\infty,k}|$ converges, and hence so does $\sum_k a_{\infty,k}$.

Now, the main event. Let $\epsilon > 0$. To show that $\lim_n \sum_k a_{n,k} = \sum_k a_{\infty,k}$ we will show that there there is some N beyond which these two sums always differ by less than ϵ.

Since $\sum_k M_k$ converges, by the Cauchy criterion there is some L where

$$\sum_{k\geq L} M_k < \frac{\epsilon}{3}$$

For arbitrary n, we compute

$$\begin{aligned}
\left|\sum_{k\geq 0} a_{n,k} - \sum_{k\geq 0} a_{\infty,k}\right| &= \left|\sum_{k<L}(a_{n,k} - a_k) + \sum_{k\geq L} a_{n,k} + \sum_{k\geq L} a_{\infty,k}\right| \\
&\leq \left|\sum_{k<L}(a_{n,k} - a_{\infty,k})\right| + \left|\sum_{k\geq L} a_{n,k}\right| + \left|\sum_{k\geq L} a_{\infty,k}\right| \\
&\leq \sum_{k<L} |a_{n,k} - a_{\infty,k}| + \sum_{k<L} |a_{n,k}| + \sum_{k\geq L} |a_{\infty,k}| \\
&\leq \sum_{k<L} |a_{n,k} - a_{\infty,k}| + 2\sum_{k>L} M_k \\
&< \sum_{k<L} |a_{n,k} - a_{\infty,k}| + \frac{2\epsilon}{3}
\end{aligned}$$

That is, for an arbitrary n we can bound the difference essentially in terms of the first L terms: the rest are uniformly less than $2\epsilon/3$. But for each of these L terms, we know that $a_{n,k} \to a_{\infty,k}$ so we can find an N making that difference as small as we like. Let's choose N_k such that $|a_{n,k} - a_{\infty,k}| < \epsilon/3L$ for each $k < L$ and then take

$$N = \max\{N_0, N_1, \ldots N_{L-1}\}$$

Now, for any $n > N$ we are guaranteed that |a_{n,k}-a_{∞,k}|<⊠/3L$ and thus that

$$\sum_{k<L} |a_{n,k} - a_{\infty,k}| < L\frac{\epsilon}{3L} = \frac{\epsilon}{3}$$

Combining with the above, we now have for all $n > N$,

$$\left|\sum_{k\geq 0} a_{n,k} - \sum_{k\geq 0} a_{\infty,k}\right| < \epsilon$$

as required. □

There is a natural version of this theorem for products as well (though we will not need it in this course, I will state it here anyway)

Theorem 15.2 (⋆ Dominated Convergence for Products)**.** *For each k let $a_{n,k}$ be a function of n, and assume the following:*

- *For each k, $a_{n,k} \to a_{\infty,k}$ is convergent.*
- *For each n, $\prod_{k\geq 0} a_{n,k}$ is convergent.*
- *There is an M_k with $|a_{n,k}| \leq M_k$ for all n.*
- *$\sum M_k$ is convergent.*

Then $\prod_{k\geq 0} \lim_n (1 + a_{n,k})$ is convergent, and

$$\lim_n \prod_{k\geq 0} (1 + a_{n,k}) = \prod_{k\geq 0} (1 + a_{\infty,k})$$

15.2. ⋆ Double Sums

A useful application of dominated convergence is to switching the order of a double sum. Given a double sequence, one may want to define an *double sum*

$$\sum_{m,n\geq 0} a_{m,n}$$

But, how should one do this? Because we have two indices, there are two possible orders we could attempt to compute this sum:

$$\sum_{n\geq 0}\sum_{m\geq 0} a_{m,n} \quad \text{or} \quad \sum_{m\geq 0}\sum_{n\geq 0} a_{m,n}$$

Definition 15.1 (Double Sum)**.** Given a double sequence $a_{m,n}$ its double sum $\sum_{m,n\geq 0} a_{m,n}$ is defined if both orders of iterated summation converge, and are equal. In this case, the value of the double sum is defined to be their common value:

$$\sum_{m,n\geq 0} a_{m,n} := \sum_{n\geq 0}\sum_{m\geq 0} a_{m,n} = \sum_{m\geq 0}\sum_{n\geq 0} a_{m,n}$$

We should be worried from previous experience that in general these two things need not be equal, so the double sum may not exist! Indeed, we can make this worry precise, by seeing that to relate one to the other is really an *exchange of order of limits*:

$$\sum_{m\geq 0} = \lim_M \sum_{0\leq m\leq M} \qquad \sum_{n\geq 0} = \lim_N \sum_{0\leq n\leq N}$$

And so, expanding the above with these definitions (and using the limit laws to pull a limit out of a finite sum) we see

$$\sum_{n\geq 0}\sum_{m\geq 0} a_{m,n} = \lim_N \sum_{0\leq n\leq N}\left(\lim_M \sum_{0\leq m\leq M} a_{m,n}\right)$$

$$= \lim_N \lim_M \left(\sum_{0\leq n\leq N}\sum_{0\leq m\leq M} a_{m,n}\right) = \lim_N \lim_M \sum_{\substack{0\leq m\leq M\\ 0\leq n\leq N}} a_{m,n}$$

Where in the final line we have put both indices under a single sum to indicate that it is a finite sum, and the order does not matter. Doing the same with the other order yields the exact same finite sum, but with the order of limits reversed:

$$\sum_{m\geq 0}\sum_{n\geq 0} a_{m,n} = \lim_M \lim_N \sum_{\substack{0\leq m\leq M\\ 0\leq n\leq N}} a_{m,n}$$

Because this is an exchange-of-limits-problem, we can hope to provide conditions under which it is allowed using Tannery's theorem.

Theorem 15.3. *Let $a_{m,n}$ be a double sequence, and assume that either*

$$\sum_{m\geq 0}\sum_{n\geq 0} |a_{m,n}| \qquad \textit{or} \qquad \sum_{n\geq 0}\sum_{m\geq 0} |a_{m,n}|$$

converges. Then the double sum $\sum_{m,n\geq 0} a_{m,n}$, meaning both iterated sums exist and are equal:

$$\sum_{m\geq 0}\sum_{n\geq 0} a_{m,n} = \sum_{n\geq 0}\sum_{m\geq 0} a_{m,n}$$

Exercise 15.1 (Cauchy's Double Summation Formula)**.** Use Dominated Convergence to prove the double summation formula (Theorem 15.3).

Hint: without loss of generality, assume that $\sum_{m\geq 0}\sum_{n\geq 0} |a_{m,n}|$ converges. Set $M_m = \sum_{n\geq 0} |a_{m,n}|$ and show the various hypotheses of Dominated convergence apply

15.3. Summation by Parts

Perhaps you remember from calculus 2 the formula for *integration by parts*, which states

$$\int_a^b u dv = uv\Big|_a^b - \int_a^b v du$$

We will of course prove this later, after we have defined the integral! But there is also a discrete version of this, for sums, which will prove useful beforehand. In fact this is a fact about *finite sums* so we could have proven it way back in the very first chapter of this book on the field axioms (like we did for the finite geometric series). But we were busy enough back then and did not, so instead the duty falls to us now in this odds-and-ends chapter of advanced techniques.

Theorem 15.4. *For sequences $\{a_n\}$ and $\{b_n\}$, we have*

$$\sum_{k=m}^{n}(a_{k+1}-a_k)b_{k+1}+\sum_{k=m}^{n}a_k(b_{k+1}-b_k)=a_{n+1}b_{n+1}-a_mb_m.$$

The good news is the proof is remarkably simple, now that we have the concept of a *telescoping series*

Proof. Combining the two terms on the left (which are finite sums, so this is no trouble), we obtain

$$\sum_{k=m}^{n}(b_{k+1}a_{k+1}-b_{k+1}a_k+a_kb_{k+1}-a_kb_k)=\sum_{k=m}^{n}(b_{k+1}a_{k+1}-a_kb_k).$$

This is a telescoping sum, and it simplifies to $a_{n+1}b_{n+1}-a_mb_m$ after all the cancellations. □

Summation by parts is often used in a slightly different form known as *Abel's Lemma*, named after the Norwegian Mathematician Niels Abel.

Theorem 15.5 (Abel's Lemma). *Let $\{a_n\}$ and $\{b_n\}$ be sequences and let $s_n = \sum_{k\leq n} a_k$ denote the nth partial sum of the series corresponding to the sequence $\{a_n\}$. Then for every $m<n$ we have*

$$\sum_{k=m+1}^{n}a_kb_k=s_nb_n-s_mb_m-\sum_{k=m}^{n-1}s_k(b_{k+1}-b_k).$$

Proof. We apply the summation by parts formula to the sequences $\{s_n\}$ and $\{b_n\}$, to obtain

$$\sum_{k=m}^{n-1}(s_{k+1}-s_k)b_{k+1}+\sum_{k=m}^{n-1}s_k(b_{k+1}-b_k)=s_nb_n-s_mb_m.$$

Since s_k is the sequence of partial sums of a_k, observe that $s_{k+1}-s_k=a_{k+1}$, and so

$$\sum_{k=m}^{n-1}a_{k+1}b_{k+1}+\sum_{k=m}^{n-1}s_k(b_{k+1}-b_k)=s_nb_n-s_mb_m.$$

Replacing k with $k-1$ in the first sum and bringing the second sum to the right, we get our result. □

This allowed Abel to produce another powerful test for convergence of a series:

Theorem 15.6 (Abel's Test). *If the series $\sum_{k=1}^{\infty}x_k$ converges, and if (y_k) is a monotone decreasing nonnegative sequence, then the series $\sum_{k=1}^{\infty}x_ky_k$ converges.*

Exercise 15.2. Prove Abel's test.

Hint: Use Abel's Lemma to observe that $\sum_{k=1}^{n}x_ky_k=s_ny_{n+1}+\sum_{k=1}^{n}s_k(y_k-y_{k+1})$, where $s_n=x_1+x_2+\cdots+x_n$ is the partial sums of x_n. Then use the comparison test to argue that

$$\sum_{k=1}^{\infty}s_k(y_k-y_{k+1})$$

converges absolutely, and show how this leads directly to a proof of Abel's Test.

Our main application of this result will be to understanding the continuity of power series at their endpoints, in a future chapter. But summation by parts makes quick work of many other calculations that might have otherwise been performed through a lengthy induction.
Here we take a look at one example: the summing of integer powers.

Example 15.3 (Summing Integers). For any $n\in\mathbb{N}$, $\sum_{k=1}^{n}k=\frac{n(n+1)}{2}$.

Let $a_k=k$ and $b_k=k-1$. Then each of the differences $a_{k+1}-a_k$ and $b_{k+1}-b_k$ equals 1, so by summation by parts, we have

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

This equality can be simplified to

$$2\sum_{k=1}^{n} k = n(n+1).$$

Dividing by two gives the claim

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

Example 15.4 (Summing Squares)**.** For any $n \in \mathbb{N}$, $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

To see this, let $a_k = k^2$ and $b_k = k - 1$. In this case,

$$a_{k+1} - a_k = (k+1)^2 - k^2 = 2k + 1,$$

and

$$b_{k+1} - b_k = 1.$$

So by the summation by parts formula, we have

$$\sum_{k=1}^{n}(2k+1)k + \sum_{k=1}^{n}(k^2)(1) = (n+1)^2 n.$$

Simplifying a bit, we get

$$3\sum_{k=1}^{n} k^2 + \sum_{k=1}^{n} k = (n+1)^3 n.$$

Since $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ from the previous example, after some algebra we end up with the desired result

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

Exercise 15.3 (Summing Cubes)**.** Prove that the following formula holds for the sum of cubes

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

Hint: follow the suggested steps below:

- Let $a_k = k^2$ and $b_k = (k-1)^2$, and apply summation by parts.
- Simplify the left side with algebra, and the sum of squares
- Divide both sides by 4, and recognize the right as the square of what we got when summing $\sum_{k=1}^{n} k$.

15.4. Problems

Exercise 15.4. Let $E(x) = \sum_{k\geq 0} x^k/k!$ Prove that if $x_n \to x$ is a convergent sequence, that

$$\lim_n E(x_n) = E(x)$$

using dominated convergence.

Exercise 15.5. Another nice application involving the series $\sum_{k\geq 0} \frac{x^k}{k!}$ is proving

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Hint: Setting $f_k(n) = \binom{n}{k}\frac{x^k}{n^k}$, it can be shown that $f_k = \lim_{n\to\infty} f_k(n) = \frac{x^k}{k!}$ and Dominated Convergence can be applied with the upper bound $M_k = \frac{|x|^k}{k!}$.

Exercise 15.6. Use Dominated Convergence to prove that

$$\frac{1}{2} = \lim_n \left[\frac{1+2^n}{2^n \cdot 3 + 4} + \frac{1+2^n}{2^n \cdot 3^2 + 4^2} + \frac{1+2^n}{2^n \cdot 3^3 + 4^3} + \cdots\right]$$

- Write in summation notation, and give a formula for the terms $a_k(n)$
- Show that $\lim_n a_k(n) = \frac{1}{3^k}$
- Show that for all n, $|a_k(n)| \leq \frac{2}{3^k}$

Use these facts to show that the hypotheses of dominated convergence hold true, and then use the theorem to help you take the limit.

Exercise 15.7 (Applying the Double Sum)**.** Since switching the order of limits involves commuting terms that are arbitrarily far apart, techniques like double summation allow one to prove many identities that are rather difficult to show directly. We will make a crucial use of this soon, in understanding exponential functions. But here is a first example:

For any $k \in \mathbb{N}$, prove the following equality of infinite sums:

$$\frac{z^{1+k}}{1-z} + \frac{(z^2)^{1+k}}{1-z^2} + \frac{(z^3)^{1+k}}{1-z^3} + \cdots = \frac{z^{1+k}}{1-z^{1+k}} + \frac{z^{2+k}}{z^{2+k}} + \frac{z^{3+k}}{1-z^{3+k}} + \cdots$$

Hint: first write each side as a summation:

$$\sum_{n\geq 1} \frac{z^{n(k+1)}}{1-z^n} = \sum_{m\geq 1} \frac{z^{m+k}}{1-z^{m+k}}$$

*Then setting $a_{m,n} = z^{n(m+k)}$, show that Cauchy summation applies to the double sum $\sum_{m,n} \geq 0 a_{m,n}$ and compute the sum in each order, arriving that the claimed equality.

Part IV.

Continuity

- In **?@sec-function-continuity** we give the definition of *continuity* and
- In **?@sec-function-properties** we look at basic properties of continuous functions and their arithmetic.
- In Chapter 18 we prove some foundational theorems about continuous functions, including the extreme value theorem and intermediate value theorem.
- In Chapter 17 we introduce the theory of power series
- In Chapter 20 we give rigorous definitions of the familiar *exponential*, *logarithmic* and *trigonometric* functions

16. Definition & Properties

Highlights of this Chapter: we formalize the concept of continuity, one of the foundational definitions in the analysis of functions. We provide an equivalent definition built out of sequences, and use it to prove 'continuity analogs' of the limit theorems. We also give the related definitions of *function limits*.

What does continuity mean? In pre-calculus classes, we often first hear something like "you can draw the graph without picking up your pencil". This is a good guide to start with for a formal definition: its clearly capturing some property that is easy to check by visual inspection! But it's not precise: terms like "you" and "pencil", as well as modal phrases like "can draw" are nowhere to be found in the axioms of ordered fields! How can we say the same thing, using words we have access to?

16.1. Epsilons and Deltas

First, a function is an input-output machine, so we should rephrase things in terms of inputs and outputs. When a graph makes a jump (where you'd have to pick up your pencil), the output changes a lot even when the input barely does. Thus, not having to pick up your pencil means you change the input by a little bit, the output changes by a little bit.

This is totally something we can make precise! A good start is by giving names to things: we want to say for any change in the input smaller than some δ, we know the output cant change that much: maybe its maximum is some other small change ϵ:

Definition 16.1 (Continuity with $\epsilon - \delta$)**.** A function f is continuous at a point a in its domain if for every $\epsilon > 0$ there is some threshold δ where if x is within δ of a, then $f(x)$ is within ϵ of $f(a)$. As a logic sentence:

$$\forall \epsilon > 0\, \exists \delta > 0\, \forall x\, |x-a| < \delta \implies |f(x) - f(a)| < \epsilon$$

A function is continuous on a set $X \subset \mathbb{R}$ if it is continuous at a for each $a \in \mathbb{R}$. A function is *continuous* if it is continuous on its domain.

16.1.1. Working with this Definition

This definition looks a lot like the sequence definition, at least in terms of the order of the quantifiers. And so we can work with it the same way: playing the "ϵ-δ game" instead of the ϵ-N game.

Example 16.1. Any constant function $f(x) = c$ is continuous at every real number a.

To prove this, we choose arbitrary $\epsilon > 0$, and observe that for any $x \in \mathbb{R}$, $f(x)-f(a) = c - c = 0$, which is less than ϵ. Thus, for any $\delta > 0$ (and if we want to be specific, choose say $\delta = 1$), if $|x - a| < \delta$ then $|f(x) - f(a)| < \epsilon$.

Example 16.2. The function $y = cx$ is continuous at every real number a.

Here's the scratch work: note that if $c = 0$ then $f(x) = 0$ is constant, and we are done by the previous example. So, we may assume $c \neq 0$. Given an aribtrary $a \in \mathbb{R}$, choose $\epsilon > 0$, and note that $|f(x) - f(a)| = |kx - ka| = |k||x - a|$. If $|x - a| < \delta$ this means $|f(x) - f(a)| < |k|\delta$, so we may choose $\delta = \epsilon/|k|$.

Remark: our value of δ is allowed to depend on ϵ, as well as properties of our function (like the k here)

While the $\epsilon - \delta$ definition is nice in that it *looks* like the sequence definition, we still end up having to play the ϵ game with every argument. Indeed, while some functions are well-suited these, for other relatively simple looking arguments, picking the right δ actually turns out to be a bit of work!

Example 16.3. The function $f(x) = x^2$ is continuous.

Scratch Work: Given $a \in \mathbb{R}$; we will prove f is continuous at a (here we do the case $a > 0$; it is only a small modification for $a < 0$: can you complete it?) Start by choosing arbitrary $\epsilon > 0$. We seek a δ such that when $|x - a| < \delta$, we can ensure $|f(x) - f(a)| = |x^2 - a^2| < \epsilon$. Using difference of squares,

$$|x^2 - a^2| = |x + a||x - a| < |x + a|\delta$$

for our future value of δ. To make further progress, let's decide to always choose a value of δ which is < 1 (if you originally had a larger δ, of course taking a smaller value will *also* work, so its no trouble to choose a maximal size). Then $|x - a| < \delta$ means x is always within 1 of a, so x can never be bigger than $a + 1$. Thus, $x + a$ can never be bigger then $(a + 1) + a$, or $2a + 1$ so we know

$$|x + a| < (2a + 1)\delta$$

For this to be less than ϵ, we can solve for δ, and set $\delta = \epsilon/(2a+1)$. Writing a rigorous proof by essentially starting with this claimed value for δ and "working backwards" confirms this works.

Like any definition, its good after seeing a few examples to also turn and look at non-examples:

Example 16.4. The step function

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

is discontinuous at 0, but is continuous at all other real numbers.

At 0, we prove discontinuity by fixing $\epsilon = 1/2$, and showing for any $\delta > 0$ there are points within δ of 0 whose values under f differ from $f(0)$ by more than ϵ. Indeed - we can just take $x = \delta/2$: this is positive, and $|x - 0| < \delta$, but $f(x) = 1$ whereas $f(0) = 0$, so $|f(x) - f(0)| = 1 > \epsilon$. However, for any nonzero $a \in \mathbb{R}$, h is continuous at a. Fixing an arbitrary $\epsilon > 0$, we can take $\delta = |a|$, and note that x being within δ of a implies x has the same sign as a (either positive or negative). Thus $f(x) = f(a)$, so $|f(x) - f(a)| = 0$ which is certainly less than ϵ.

Thus, a function with a jump in it is discontinuous right at the jump, as we expect. This shows its possible for a function to be *discontinuous* at a single point, but things can get much stranger!

Example 16.5. The characteristic function of the rational numbers is discontinuous *everywhere.*

$$b(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

Setting $\epsilon = 1/2$, note that proving *discontinuity at a* means showing that for *any$ $\delta > 0$ we can find an x within δ of a where $f(x)$ differs from $f(a)$ by more than $1/2$. The proof breaks into two cases depending on the (ir)rationality of a. First, for irrational a, by the density of rationals we may for any $\delta > 0$ find a rational number x with $a - \delta < x < a + \delta$, so $|x - a| < \delta$. But, $f(a) = 0$ since a is irrational and $f(x) = 1$ since its rational, thus $|f(x) - f(a)| = 1 > \epsilon$. The case of a rational is similar, but now we use the density of the irrationals to find an appropriate x.

We saw above a function that is discontinuous at a single point, and then one that is discontinuous everywhere. What's harder to imagine, is a function that is *continuous at a single point.* Try thinking about what this might mean!

Exercise 16.1. Show that the following function is continuous at 0 and discontinuous everywhere else:

$$g(x) = \begin{cases} x & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

There are even stranger functions out there: for instance, the Thomae function

$$\tau(x) = \begin{cases} \frac{1}{q} & x \in \mathbb{Q} \text{ and } \frac{p}{q} \text{ is lowest terms.} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

is continuous at the irrational numbers, and discontinuous at every rational.

As an example of proving something using continuity, we prove the useful fact that when an *continuous* function is nonzero at some point, it actually *stays* nonzero for a little bit on each side.

Proposition 16.1 (Nonzero on a Neighborhood)**.** *If f is continuous, $f(a) \neq 0$ then there is a small open interval about a where f is nonzero.*

Proof. Let $f(a) = c$ with $c > 0$, and set $\epsilon = |f(c)|/2$. By continuity, there is some δ such that if $|x - c| < \delta$ we know $|f(x) - f(c)| < \epsilon$. Unpacking this, for all $x \in (c - \delta, c + \delta)$ we know

$$-\epsilon = \frac{-|f(c)|}{2} < f(x) - f(c) < \frac{|f(c)|}{2} = \epsilon$$

And thus

$$f(c) - \frac{|f(c)|}{2} < f(x) < f(c) + \frac{|f(c)|}{2}$$

If $f(c)$ is positive, then the lower bound here is $f(c)/2$ which is still positive, so $f(x)$ is always positive in the interval. And, if $f(c)$ is negative, the upper bound here is $f(c)/2$ which is still negative: thus $f(x)$ is always negative in the interval. □

16.2. Continuity and Sequences

We spent a lot of time working with sequences so far, so it would be nice if we could leverage some of that knowledge as more than just analogy. And indeed we can! We give the formal result below, but pause to develop some intuition:

Theorem 16.1 (Sequences and Continuity)**.** *Let f be a real function, and a a point of its domain. Then f is continuous at a if and only if for every sequence a_n in the domain wtih $a_n \to a$, we have $f(a_n) \to f(a)$.*

This theorem is an *equivalence of definitions* or an *if-and-only-if* result, so the proof requires two parts: first we show that continuity implies sequence continuity, and then we show the converse.

Continuity Implies Sequence Continuity. Let f be continuous at a, and x_n an arbitrary sequence converging to a. We wish to show the sequence $f(x_n)$ converges to $f(a)$. Choosing an $\epsilon > 0$, we use the assumed continuity to get a $\delta > 0$ where $|x - a| < \delta$ implies that $|f(x) - f(a)| < \epsilon$.

But since $x_n \to a$, we know there must be some N such that for $n > N$ we have $|x_n - a| < \delta$: thus for this same N we have $|f(x_n) - f(a)| < \epsilon$.

Putting this all together, this is just the definition of convergence for the sequence $f(x_n)$ to $f(a)$: starting with $\epsilon > 0$ we got an N which for $n > N$ we can guarantee $|f(x_n) - f(a)| < \epsilon$. So we are done. □

Sequence Continuity Implies Continuity. Here we prove the contrapositive: that if f is *not continuous* at a then it is also *not sequence continuous* there.

If f is not continuous at a then there is some ϵ where for every $\delta > 0$ we can find points within δ of a where $f(x)$ is more than ϵ away from $f(a)$. From this we need to somehow produce a *sequence*, so we will take a sequence of such δ's and for each pick some such bad point x.

For example, if we let $\delta = 1/n$ then call x_n the point with $|x_n - a| < 1/n$ but $|f(x_n) - f(a)| > \epsilon$. Doing this for all n produces a sequence where

$$a - \frac{1}{n} < x_n < a + \frac{1}{n}$$

And so by the squeeze theorem we see that x_n converges, and its limit is a. But we also know (by our choices of x_n) that for *every element of this sequence* |f(x_n)-f(a)|>⊠$, so there's no way that $f(x_n)$ converges to $f(a)$.

Thus, we've shown by example that our function is not sequence continuous at a, as required. □

When working with this definition of continuity, its important to remember that we need to check $f(\lim x) = \lim f(x_n)$ for *all sequences* $x_n \to a$. If it fails for any individual sequence, that is enough to show the function is not continuous at that point. Thus when proving continuity we will always start with *let x_n be an arbitrary sequence converging to a*, and make use of convergence theorems to help us (since we cannot know the particular sequence), whereas for proving discontinuity all we need to do is produce a specific example sequence that fails.

With this definition, we can bring all of our theory on limits and put it to work. We see many of these benefits below; here we pause merely to re-do a single example for illustrative purposes:

Example 16.6. The function $f(x) = x^2$ is continuous on all of $\mathbb{R}$.

Let $a \in \mathbb{R}$ be arbitrary, and choose an arbitrary sequence a_n of real numbers with $a_n \to a$ (we know *at least one such sequence exists* since we have proven every real number is the limit of a sequence of rationals). By the limit theorem for products, since $a_n \to a$ we know $a_n \cdot a_n \to a \cdot a$. And as $f(x) = x^2$, we can rewrite this as $f(a_n) \to f(a)$. Since $a_n \to a$ was arbitrary, this holds for all such sequences, and so f is continuous at a. But since $a \in \mathbb{R}$ was arbitrary, f is continuous on the entire real line.

16.3. Building Continuous Functions

Because we have an equivalent characterization of continuity in terms of sequence convergence, and we have many theorems about this, we can use our characterization to rephrase these as results about *continuity.*

Proposition 16.2 (Continuity of Constant Multiples). *If f is continuous at $a \in \mathbb{R}$ and $k \in \mathbb{R}$ is a constant, then the function $kf : x \mapsto kf(x)$ is continuous at a.*

Using $\epsilon - \delta$. First note if $k = 0$ we are done as $kf(x) = 0$ is a constant function. Otherwise, let $\epsilon > 0$: since f is continuous at a there exists a $\delta > 0$ such that $|x-a| < \delta$ implies $|f(x) - f(a)| < \epsilon/|k|$. But this implies

$$|kf(x) - kf(a)| = |k||f(x) - f(a)| < k\frac{\epsilon}{|k|} = \epsilon$$

So $|x - a| < \delta$ implies $|kf(x) - kf(a)| < \epsilon$, and $kf(x)$ is continuous at a. □

Using Sequences. Let $a \in \mathbb{R}$ be arbitrary, and x_n a sequence converging to a. Then by the limit theorem for multiples, $kx_n \to ka$. Rephrasing this in terms of the function $f(x) = kx$, this just says that $\lim f(x_n) = f(\lim x_n)$ so f is continuous at a. □

Theorem 16.2 (Continuity and the Field Operations). *Let f, g be functions which are continuous at a point a. Then the functions $f(x) + g(x)$, $f(x) - g(x)$ and $f(x)g(x)$ are all continuous at a. Furthermore if $g(a) \neq 0$ then $f(x)/g(x)$ is also continuous at a.*

Proof. We prove the case for sums, and leave the rest as an exercise. Let f, g be any two continuous functions and let $a \in \mathbb{R}$ be a point in their domains. Let x_n be any sequence converging to a. Since f is continuous we know that $\lim f(x_n) = f(\lim x_n) = f(a)$ and similarly by the continuity of g, $\lim g(x_n) = f(\lim x_n) = g(a)$. Thus by the limit theorem for sums, the sequence $f(x_n) + g(x_n)$ is convergent, with

$$\lim (f(x_n) + g(x_n)) = \lim f(x_n) + \lim g(x_n) = f(a) + g(a)$$

So, $f + g$ is continuous at a. Since a was arbitrary, we see that $f + g$ is continuous at every point of its domain. The same argument applies for subtraction, multiplication, and division using the respective limit theorems for sequences. □

Exercise 16.2. Prove the remaining "continuity theorems".

Exercise 16.3 (Continuity of Polynomials)**.** Prove that every polynomial is a continuous function on the entire real line. *Hint: prove x^n is continuous for each n by induction. Then prove the result for polynomials by induction on their degree!*

Exercise 16.4 (Continuity of Rational Functions)**.** A rational function is a quotient of polynomials $r(x) = p(x)/q(x)$. Prove that every rational function is continuous, on every point of its domain.

One of the most important operations for functions is that of *composition*: if $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ then the function $g \circ f : \mathbb{R} \to \mathbb{R}$ is defined as $g \circ f(x) := g(f(x))$. More generally, so long as the domain of g is a subset of the range of f, the composition $g \circ f$ is well defined.

Theorem 16.3 (Continuity of Compositions)**.** *Let f, g be functions such that f is continuous at a, and g is continuous at $f(a)$. Then the composition $g \circ f(x) := g(f(x))$ is continuous at a.*

Proof. Let x_n be an arbitrary sequence converging to $a \in \mathbb{R}$: we wish to show that $\lim g(f(x_n)) = g(f(\lim x_n)) = g(f(a))$. Since f is continuous at $x = a$ we see immediately that $f(x_n)$ is a convergent sequence with $f(x_n) \to f(a)$. And now, since g is assumed to be continuous at $x = f(a)$ and $f(x_n)$ is a sequence converging to this point, we know $g(f(x_n)) = g(f(a))$ as required. □

Another limit theorem we had was the limit theorem for the square root: which translates directly to a continuity theorem as well!

Theorem 16.4 (Continuity of Roots)**.** *The function $R(x) = \sqrt{x}$ is continuous on $[0, \infty)$.*

Proof. We actually already proved this, as a limit theorem about the square root! CITE states that if $x \geq 0$ and x_n is any sequence of nonnegative numbers converging to x, then $\lim \sqrt{x_n} = \sqrt{\lim x_n}$. Thus $\sqrt{x}$ is continuous at x, and as x is an arbitrary nonnegative value, its continuous on its domain. □

Corollary 16.1 (Continuity of Absolute Value)**.** *The absolute value satisfies $|x| = \sqrt{x^2}$ for all real x. This is a composition of two continuous functions, and thus is continuous.*

The same is true for n^{th} roots, though we do not stop to prove it here, you may wish to for practice! This is a special case of a more general result on the continuity of inverse functions (as the square root is the inverse of x^2)

Theorem 16.5 (Continuity of Inverse Functions)**.** *Let $f : A \to B$ be a continuous invertible function for $A, B \subset \mathbb{R}$ bounded subsets. Then f^{-1} is continuous.*

By Contradiction. Assume for the sake of contradiction that $f : A \to B$ is continuous and invertible with $A, B \subset \mathbb{R}$ but f^{-1} is not continuous. Then there would be some sequence convergent sequence $b_n \to b$ where $f^{-1}(b_n) \not\to f^{-1}(b)$.

This sequence $f^{-1}(b_n)$ lies in A which is bounded, so it contains a convergent subsequence due to Bolzano Weierstrass. Its further possible to select such a subsequence $f^{-1}(b_{n_k})$ that converges to some value $a \neq f^{-1}(y)$ (if this were impossible, *all* convergent subsequences would converge to $f^{-1}(y)$, and so our sequence would have converged to this value!)

Now we use the fact that f is continuous. Since $f^{-1}(y_{n_k}) \to a$, we see $f\left(f^{-1}(y_{n_k})\right) \to f(a)$, and since f^{-1} is the inverse of f, this just means that $y_{n_k} \to f(a)$. But since f is invertible, its $1-1$, so the fact that $a \neq f^{-1}(y)$ means that $f(a) \neq y$. That is, we have found a subsequence y_{n_k} of y_n, which does not converge to y.

But this implies that the sequence y_n itself does not converge to y (else all subsequence would converge to y!) and this is a contradiction, as we assumed $y_n \to y$ at the very start. Thus, f^{-1} is actually continuous, as desired. $\square$

We can apply this to functions we care about like n^{th} powers, to prove the continuity of n^{th} roots.

Corollary 16.2. *The function $\sqrt[n]{x}$ is continuous on the positive reals.*

Proof. Let $x > 0$, we want to show $\sqrt[n]{}$ is continuous at x. So, we need to choose bounded sets A, B for our domains, to make sure things work. Taking $B = [0, x+1]$ will do, as it contains x, and then $A = [0, \sqrt[n]{x+1}]$, so x^n is an invertible function from A to B. Its continuous, so its inverse is continuous, meaning $\sqrt[n]{}$ is continuous on the interval $B = [0, x+1]$ which contains x. Thus its continuous at x, and as x was an arbitrary positive real number, $\square$

Exercise 16.5. Our argument above showed $\sqrt[n]{x}$ is continuous at all positive inputs. Show its continuous at zero.

Note to students! If you can think of a better proof of this, (especially one that doesn't have this awkward boundedness/bolzano weierstrass stuff) let me know. If its slick enough, I'll replace the proof in the textbook, and thank you in a footnote of the future editions!

The combination of these theorems allows us to prove many complicated functions are continuous, that would otherwise have been quite difficult directly from the definition!

Example 16.7. The following function is continuous on the entire real line.

$$\frac{(x+|x^2-1|)^4}{1+\sqrt[3]{1+|x-1|^7}}$$

To prove it, we work from *inside out*, like we do for using the limit laws. Starting with the numerator, we see x^2-1 is continuous as its a polynomial, so $|x^2+1|$ is continuous as the composition of two continuous functions, and $x+|x^2+1|$ is continuous as its the sum of two continuous functions. For the denominator, we similarly start with the continuity of $x-1$, compose with $||$ to get the continuity of $|x-1|$, then compose with x^7 to get $|x+1|^7$ is continuous, add the constant function 1 (which is continuous) take the cube root (composing with the continuous function $\sqrt[3]{x}$) and finally add the continuous function 1 once more. So, the numerator and denominator are continuous. Finally, the denominator is strictly positive for all x (hence nonzero), so the quotient is continuous.

Exercise 16.6 (Continuity of Max and Min)**.** Prove that for any $a, b \in \mathbb{R}$, we have

$$\max\{a,b\} = \frac{a+b+|a-b|}{2}$$

Intuitively, notice $(a+b)/2$ is the midway point between a and b, and $|a-b|/2$ is half the distance between them. So $(a+b)/2+|a-b|/2$ is the midway point plus half the distance, so its the larger of the two. But give a rigorous argument, perhaps by cases.

Use this to conclude that if f, g are two continuous functions then $M(x) = \max\{f(x), g(x)\}$ is also continuous. Propose and prove a similar formula for the minimum, and show that $\min\{f(x), g(x)\}$ is continuous in x.

16.4. Function Limits

A related but slightly different concept is the *limit* of a function. We include this here as the definition and techniques ties very closely to those for continuity; we will have use for this material when we introduce the *derivative*, and in other cases where we need to understand the behavior of a function *near* a point, without actually being able to compute the function's value *at* that point (perhaps, that point is outside the functions' domain).

Definition 16.2 (Limits of Functions)**.** Let $f : D \to \mathbb{R}$ and a be a limit point of D. Then we write $\lim_{x\to a} f(x) = L$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that if $x \in D$ and $|x-a| < \delta$ then $|f(x) - L| < \epsilon$.

One can alternatively phrase this in terms of sequences:

Exercise 16.7. Prove the following definition is equivalent to $\lim_{x\to a} f(x) = L$: Given any sequence $\{x_n\}$ in D with $x_n \neq a$ for all n, $x_n \to a$ implies that $f(x_n) \to L$.

Example 16.8.

$$\lim_{x\to 2} \frac{x^2-4}{x-2}$$

Let x_n be any sequence converging to 2, for which $x_n \neq 2$ for all n. Then since $x_n \neq 2$ the denominator of $(x^2-4)/(x-2)$ is never zero, and we can simplify with algebra:

$$\frac{x_n^2-4}{x_n-2} = \frac{(x_n+2)(x_n-2)}{x_n-2} = x_n + 2$$

Thus, for all n we have

$$\lim \frac{x_n^2-4}{x_n-2} = \lim x_n + 2 = \lim(x_n) + 2 = 4$$

Since x_n was arbitrary, this holds for all sequences and

$$\lim_{x\to 2} \frac{x^2-4}{x-2} = 4$$

We will be most interested in taking the limit of functions in cases where things are *not actually define* at a like the example above: the most important example being the derivative, defined as the limit $f'(a) = \lim_{x\to a}(f(x) - f(a))/(x-a)$. However a good sanity check with a new definition is to see it performs as expected in known situations

Theorem 16.6 (Limits of Continuous Functions). *If f is continuous at a, then $\lim_{x\to a} f(x) = f(a)$.*

Proof. Let x_n be a sequence converging to a, but not equal to a at any term. Since f is continuous at a, we know the sequence $f(x_n)$ converges to $f(a)$. Thus by the sequence definition of function limits $\lim_{x\to a} f(x) = f(a)$. □

As an exercise, re-prove this result using the $\epsilon - \delta$ definition.

16.4.1. ⧫ One-Sided Limits

The definition of function limit requires understanding *all sequences* limiting to a but not equal to a. In applications, its often important to consider more restricted limits, looking only at what happens when we approach a from *above* or from *below*.

Definition 16.3 (Left- and Right-Sided Limits)**.** Let f be a function

Similarly to above, these definitions have sequence counterparts (prove this, as an exercise):

Definition 16.4. Let f be a function. Then $\lim_{x\to a^+} f(x) = L$ if for every sequence $x_n \to a$ with $x_n > a$ we have $f(x_n) \to L$. Similarly $\lim_{x\to a^-} f(x) = L$ if for every $x_n \to a$ with $x_n < a$ we have $f(x_n) \to L$.

Exercise 16.8 (Limit Exists when Both Sides Agree)**.** Let f be a function defined on an interval containing a (but perhaps not at a). Then $\lim_{x\to a} f(x)$ exists if and only if both $\lim_{x\to a^+} f$ and $\lim_{x\to a^-} f$ both exist, and in this case is equal to their common value.

Exercise 16.9 (One Sided Limits of Monotone Functions)**.** Let f be a bounded monotone function on the interval (a, b). Then both of the one sided limits exist

$$\lim_{x\to a^+} f(x) \qquad \lim_{x\to b^-} f(x)$$

Hint: show they are the inf and sup of $\{f(x) \mid x \in (a, b)\}$

This proves useful in many cases where we know only that our function is monotone, but cannot compute its values. For us, the most important application is Proposition 24.1 where we show exponential functions are differentiable, when we have only assumed they are continuous.

16.5. Problems

Exercise 16.10. Let $f(x)$ be a continuous function, and assume that $f(x)^2$ is a constant function. Prove that $f(x)$ is constant. To show continuity is an essential assumption, give an example of an $f(x)$ where $f(x)^2$ is constant, but f is not.

Exercise 16.11. Recall that a function f is a contraction map if there exists a $k \in (0, 1)$ with $|f(x) - f(y)| < k|x - y|$ for all x, y. Prove that contraction maps are continuous.

Exercise 16.12. If f is continuous at a point a, then $|f|$ is continuous there, by using the reverse triangle inequality.

Exercise 16.13. The function

$$\operatorname{sgn}(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0 \end{cases}$$

is discontinuous at $x = 0$, but continuous at every other real number.

Exercise 16.14 (Removable and Jump Discontinuities)**.**

$$f(x) = \begin{cases} 0 & x < 0 \\ 17 & x = 0 \\ x & x > 0 \end{cases}$$

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

Next consider

$$g(x) = \begin{cases} 0 & x < 0 \\ 17 & x = 0 \\ x^2 + 1 & x > 0 \end{cases}$$

Show that $\lim_{x \to 0} g(x)$ does not exist.

Exercise 16.15 (The Pasting Lemma)**.** Let f, g be two continuous functions and $a \in \mathbb{R}$ is a point such that $f(a) = g(a)$. Prove that the piecewise function below is continuous at a.

$$h(x) = \begin{cases} f(x) & x \leq a \\ g(x) & x > a \end{cases}$$

17. Power Series

Chapter highlights: we prove that a power series is continuous on its entire domain: this is a combination of two theorems, that (1) its continuous on the interior and (2) that its continuous at boundary points, when defined there. Proving the continuity of series provides an opportunity to use the material we learned in *Series: Advanced Techniques.* In particular, we will use Dominated Convergence to prove the continuity of series within the interval of convergence, and Summation By Parts to prove continuity at any boundary points.

17.1. Continuity In the Interior

We have proven previously that a power series $\sum_{n\geq 0} a_n x^n$ either converges only at $\{0\}$, converges on the entire real line, or has a finite *radius of convergence*, where it converges on an interval of the form $(-r,r)$, $[r,r]$, $(-r,r]$ or $[-r,r)$. Plotting the partial sums of such a series shows that outside the radius things quickly blow up to infinity, but within the radius of convergence the result appears to be continuous. We confirm this below.

Theorem 17.1 (Continuity within Radius of Convergence)**.** *Let $f(x) = \sum_k a_k x^k$ be a power series with radius of convergence r. Then if $|x| < r$, f is continuous at x.*

Proof. Take $x > 0$ (leaving the trivial modifications for $x < 0$ as an exercise), and let x_n be an arbitrary sequence in $(-r,r)$ converging to x. We aim to show that $f(x_n) \to f(x)$.

As $x < r$ choose some y with $x < y < r$ (perhaps, $y = (x+r)/2$). Since $x_n \to x$ there is some N past which x_n is always less than y (take $\epsilon = y - x$ and apply the definition of $x_n \to x$). As truncating the terms of the sequence before this does not change its limit, we may without loss of generality assume that $x_n < y$ for all n. Thus, we may define $M_k = a_k y^k$, and we are in a situation to verify the hypotheses of Dominated Convergence:

- Since $x_n \to x$, we have $a_k x_n^k \to a_k x^k$ by the limit theorems.
- For each n, $f(x_n) = \sum_k a_k x_n^k$ is convergent as x_n is within the radius of convergence.
- $M_k = a_k y^k$ bounds $a_k x_n^k$ for all n, as $0 < x_n < y$.

- $\sum_k M_k$ converges as this is just $f(y)$ and y is within the radius of convergence.

Applying the theorem, we see

$$\lim_n f(x_n) = \lim_n \sum_k a_k x_n^k = \sum_k \lim a_k x_n^k = \sum_k a_k x^k = f(x)$$

Thus for arbitrary $x_n \to x$ we have $f(x_n) \to f(x)$, so f is continuous at x. $\square$

Remark 17.1. If the power series converges on all of $\mathbb{R}$, the same proof above holds, taking a sequence $x_n \to x$, but its a little easier: we don't have to be careful choosing our upper bound y. Any upper bound for the convergent sequence $\{x_n\}$ will do.

We pause to remark this result is something rather special to power series, and is not true in general: its quite possible to write down a sequence of continuous functions which converges to a *discontinuous* function. So the fact a sequence of (continuous) *partial sums* of a power series converges to a continuous limit is indeed a big deal! This is one of many things that makes power series particularly nice.

Exercise 17.1. Let $f(x) = \frac{x}{\sqrt{x^2+1}}$, and define the sequence of functions $f_n(x) = f(nx)$. As $n \to \infty$ prove that

- If $x > 0$ then $f_n(x) \to 1$
- If $x = 0$ then $f_n(x) \to 0$
- If $x < 0$ then $f_n(x) \to -1$

Thus, while f_n is continuous for each n (its a composition of continuous functions), the limit is discontinuous.

There is a lot of theoretical work in real analysis to determine more general conditions under which a sequence of continuous functions converges to a continuous limit. In this semester long course we won't have need for such results beyond the power series case above, but in the eventual extension of this book, we will develop the notion of *uniform convergence* for this purpose.

17.2. ⧫ Continuity at the Boundary

While we now completely understand a power series on the *interior* of its radius of convergence, there's a little more work to do to complete the picture.

Theorem 17.2 (Continuity at the Boundary: Abel's Theorem). *Let $f(x) = \sum_{k\geq 0} a_k x^k$ be a power series with radius of convergence r, which converges at an endpoint $\pm r$. Then f is continuous there.*

The full proof of this theorem is rather more technical than the previous result, and before proving it through a sequence of steps, we pause to appreciate why. First, note that *some* cases of this theorem really are easy: for instance, if the endpoint converges absolutely, you can carry out the exact same proof using Dominated Convergence as above.

Exercise 17.2. Let $\sum_{k\geq 0} a_k x^k$ be a power series with radius of convergence 1, and suppose $\sum_{k\geq 0} a_k$ converges absolutely. Then

$$\lim_{x\to 1^-}\sum_{k\geq 0} a_k x^k = \sum_{k\geq 0} a_k$$

The difficulty is then is what happens when $\sum_{k\geq 0} a_k$ converges, but *does not do so absolutely*. This is a real case, that actually shows up in important situations rather often: for example the power series

$$\sum_{n\geq 0}\frac{(-1)^n}{n+1}x^{n+1} \qquad \sum_{n\geq 0}\frac{(-1)^n}{2n+1}x^{2n+1}$$

Both converge at $x=1$, but their absolute values *diverge*. However, the continuity of these power series at their endpoints will be absolutely essential to us later on in the course, when deriving the amazing identities

$$\log(2) = 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots \qquad \frac{\pi}{4} = 1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots$$

We also pause to quickly dash any hopes there might be a general sort of proof: (perhaps one hopes that if a sequence of functions converges on $(-1,1)$ to a continuous limit, and also converges at a boundary point then it is automatically continuous at the boundary).

Example 17.1. Consider the sequence $f_n(x) = x^n$. As $n\to\infty$ this converges on the interval $(-1,1]$ and diverges everywhere else. Furthermore for $|x|<1$ it converges to the *zero function* which is constant - hence continuous. However at $x=1$ the sequence $f_n(1) = 1^n = 1$ is constantly equal to 1, so the limit is also 1. Thus the limit is defined at a boundary point, but not continuous there.

Finally, we proceed with the proof of the theorem. To simplify notation, we prove it for series with radius of convergence 1, and leave the simple rescaling to series converging for $|x|<r$ as an exercise. That is, we aim to prove

Proposition 17.1 (Abel's Theorem for $r=1$)**.** *Let $\sum_{k\geq} 0a_k x^k$ be a power series which converges for $|x|<1$ and assume $\sum_{k\geq 0} a_k$ also converges. Then*

$$\lim_{x\to 1^-}\sum_{k\geq 0} a_k x^k = \sum_{k\geq 0} a_k x^k$$

Throughout the proof, it is useful to introduce the notation $A_N = \sum_{k=0}^{N} a_k$ and $A = \lim A_N = \sum_{k\geq 0} a_k$ for the partial sums of of the coefficients and their limit. We begin with a lemma providing a means of rewriting the series within the radius of convergence.

Lemma 17.1. *Let $\sum a_k x^k$ be a power series which converges for $|x| < 1$, and $A_N = \sum_{k\leq N} a_k$ be the partial sums of the coefficients. Then for any $x \in (-1, 1)$*

$$\sum_{k\geq 0} a_k x^k = (1-x)\sum_{k\geq 0} A_k x^k$$

Proof. Recall the formulation of summation by parts given in Abel's lemma:

$$\sum_{k=1}^{N} x_k y_k = X_N y_N - X_0 y_0 - \sum_{k=0}^{N-1} X_k(y_{k+1} - y_k)$$

for X_N the partial sums of the sequence $\{x_k\}$. We can interpret a power series $\sum a_k x^k$ as summing a *product of two sequences*: the sequence a_k of coefficients and the sequence x^k of monomials. Looking at a partial sum of our power series and summing by parts

$$\begin{aligned}
\sum_{k=0}^{N} a_k x^k &= a_0 + \sum_{k=1}^{N} a_k x^k \\
&= a_0 + \left(A_N x^N - A_0 x^0 - \sum_{k=1}^{N-1} A_k(x^{k+1} - x^k)\right)
\end{aligned}$$

Using that $A_0 = a_0$ and $x^0 = 1$, this simplifies

$$\begin{aligned}
&= A_N x^N - \sum_{k=1}^{N-1} A_k x^k (x-1) \\
&= A_N x^N + (1-x)\sum_{k=1}^{N-1} A_k x^k
\end{aligned}$$

As we assumed $A = \sum_{k\geq 0} a_k$ converges, $A_N \to A$ as $N \to \infty$. And as $|x| < 1$, we know $x^N \to 0$ by our basic limits. Thus by the limit law for products, $A_N x^N \to 0$, so for the full power series

$$\begin{aligned}
\sum_{k\geq 0} &= \lim_N \sum_{k=0}^{N} a_k x^k \\
&= \lim_N \left(A_N x^N + (1-x) \sum_{k=1}^{N-1} A_k x^k \right) \\
&= 0 + (1-x) \lim_N \sum_{k=0}^{N-1} A_k x^k \\
&= (1-x) \sum_{k\geq 0} A_k x^k
\end{aligned}$$

□

Next, we use this new form of the series to convert our problem to something simpler

Lemma 17.2. *Let $f(x) = \sum a_n x^n$ be a power series with radius of convergence* 1*, which also converges at $x = 1$. Then f is continuous at* 1 *if and only if*

$$(1-x)\sum_{n\geq 0}(A_n - A)x^n \to 0$$

as $x \to 1^-$ for $A_N = \sum_{n\leq N} a_n$ and $A = \sum_{n\geq 0} a_n$.

Proof. The overall goal is to show

$$\lim_{x\to 1^-} \sum_{k\geq 0} a_k x^k = \sum_{k\geq 0} \lim_{x\to 1^-} a_k x^k = \sum_{k\geq 0} a_k = A$$

Subtracting A from both sides, we must show $\lim_{x\to 1^-} \sum_{k\geq 0} a_k x^k - A = 0$. Switching out the series for its alternative expression derived in the previous lemma, our problem is equivalent to showing

$$\lim_{x\to 1^-} \left((1-x) \sum_{k\geq 0} A_k x^k \right) - A = 0$$

To turn this into something useful, we do a sneaky trick. Recall that for $|x| < 1$ we know $\sum_{k\geq 0} x^k = \frac{1}{1-x}$ (the geometric series). Clearing the denominator, this means $(1-x)\sum_{k\geq 0} x^k = 1$, so we "multiply by 1"

$$
\begin{aligned}
(1-x)\sum_{k\geq 0} A_k x^k - A &= (1-x)\sum_{k\geq 0} A_k x^k - A\cdot 1 \\
&= (1-x)\sum_{k\geq 0} A_k x^k - A(1-x)\sum_{k\geq 0} x^k \\
&= (1-x)\left(\sum_{k\geq 0} A_k x^k - \sum_{k\geq 0} A x^k\right)
\end{aligned}
$$

Because both of the sums involved are convergent, we can add them term-by-term without changing the value (Proposition 12.1), combining the sums.

$$
\begin{aligned}
&= (1-x)\sum_{k\geq 0}(A_k x^k - A x^k) \\
&= (1-x)\sum_{k\geq 0}(A_k - A) x^k
\end{aligned}
$$

This is just a rewriting of our original series minus the proposed limit. So proving this converges to zero is logically equivalent to our desired result $\square$

Finally, we prove Abel's theorem by showing this does indeed limit to zero, as $x \to 1^-$.

Proof. We work directly with the limit definition: for arbitrary $\epsilon > 0$ we must provide a $\delta > 0$ such that if $1-\delta < x < 1$, our sum is less than ϵ in absolute value. By the triangle inequality and limit inequalities,

$$
\left|(1-x)\sum_{k\geq 0}(A_k - A)x^k\right| \leq |1-x|\sum_{k\geq 0}|A_k - A||x|^k = (1-x)\sum_{k\geq 0}|A_k - A|x^k
$$

Where the final equality holds as we are concerned with $\lim_{x\to 1^-}$ so we can without loss of generality assume $x > 0$. So, our goal is to show the right hand side is smaller than ϵ, when x is sufficiently close to 1. Since $A_k \to A$ by definition, there is an N such that for all $k \geq N$, $|A_k - A| < \epsilon$. So, we break our sum into two terms, to estimate separately:

$$
(1-x)\sum_{k\geq 0}|A_k - A|x^k = (1-x)\sum_{0\leq k<N}|A_k - A|x^k + (1-x)\sum_{k\geq N}|A_k - A|x^k
$$

In the second of these sums, we know that $|A_k - A| < \epsilon$, so

$$\sum_{k\geq N} |A_k - A|x^k \leq \sum_{k\geq N} \epsilon x^k = \epsilon \sum_{k\geq N} x^k \leq \epsilon \sum_{k\geq 0} x^k = \epsilon \frac{1}{1-x}$$

Substituting back into the full second term,

$$(1-x)\sum_{k\geq N} |A_k - A|x^k \leq (1-x)\frac{\epsilon}{1-x} = \epsilon$$

Thus the second term can be made as small as we like (independently of the value of x!), and we need only think about the first term. But this is actually easy: its a *finite sum*! For any positive $x < 1$ we know $x^k < 1^k = 1$, so

$$\sum_{0\leq k<N} |A_k - A|x^k < \sum_{0\leq k<N} |A_k - A|$$

Call the value of this finite sum on the right L. Then to make the term $(1-x)\sum_{0\leq k<N} |A_k - A|x^k$ less than ϵ it suffices to make $1-x$ less than ϵ/L. So, set $\delta = \epsilon/L$: then for any $x < 1$ with $|x-1| < \delta$, we know

$$(1-x)\sum_{0\leq k<N} |A_k - A|x^k < (1-x)\sum_{0\leq k<N} |A_k - A| = (1-x)L < \frac{\epsilon}{L}L = \epsilon$$

Putting it all back together, we see that for this δ, $|x-1| < \delta$ implies our sum is less than 2ϵ. So, we need to go back and replace some epsilons with $\epsilon/2$'s to complete the proof [Apologies: due to the length of this argument, I prioritized readability, and reduced clutter by not writing the correct $\epsilon/2$'s everywhere]. □

As a last step, we do the substitutions to return from $r = 1$ to general radius of convergence.

Exercise 17.3 (The General Case)**.**

- Let $f(x) = \sum_{n\geq 0} a_n x^n$ be a power series which converges on $(-r, r)$ and also at r. Prove that it is continuous at r. *Hint: consider $f(rx)$: show this has radius of convergence 1 and converges at 1. Then apply the previous theorem, and re-arrange to yield the result.
- Let $f(x) = \sum_{n\geq 0} a_n x^n$ be a power series which converges on $(-r, r)$ and also at $-r$. Prove that it is continuous at $-r$. *Hint: consider $f(-x)$.

17.3. ⧫ Uniqueness

Using this continuity result, we prove a theorem which is very helpful for not getting lost in the world of power series. Its natural to wonder if two distinct power series could converge to the same function (as limits of their partial sums). Of course if any of their coefficients differed no *finite* partial sums could be equal (as the finite sums are polynomials, and polynomials are fully determined by their coefficients). But this doesn't rule out any coincidence in the limit. After all, this happens with numbers all the time! *Finite* decimals are determined by their digits, but infinite decimals are not! $1.00000\ldots = 0.999999\ldots$, and $0.5000 = 0.4999\ldots$, etc.

But this cannot happen for power series: if a function $f(x)$ can be written as a power series in two ways $\sum a_k x^k$ and $\sum b_k x^k$, then these two are *precisely equal*: $a_k = b_k$ for all k. But we can get by with much less information than this: if two power series agree on a single *sequence converging to zero* thats already enough information to completely determine them! This is the content of the theorem below.

Theorem 17.3. *Let $f(x) = \sum_{k\geq 0} a_k x^k$ and $g(x) = \sum_{k\geq 0} b_k x^k$ be two power series with positive radii of convergence (not a priori assumed to be the same). Then if for some sequence c_n of nonzero terms converging to zero we have $f(c_n) = g(c_n)$, it follows that f and g are* the same power series*: $a_k = b_k$ for all k.*

Proof. We proceed step by step, starting with the constant terms.
Since $f(x)$ is continuous we can compute $f(0)$ as $\lim f(c_n)$. Similarly, the continuity of g lets us write $g(0) = \lim g(c_n)$. But by our assumption that $f(c_n) = g(c_n)$, this implies $f(0) = g(0)$. Plugging in zero to our power series shows

$$f(0) = a_0 + a_1 0 + a_2 0^2 + \cdots = a_0 \qquad g(0) = b_0 + b_1 0 + b_2 0^2 + \cdots = b_0$$

Thus $a_0 = b_0$, so the first two terms of our power series are equal.

This argument was remarkably efficient, so let's try to repeat it. Subtracting the constant terms from f and g yields functions divisible by x, and dividing by that x gives

$$f_1(x) = \frac{f(x) - a_0}{x} = a_1 + a_2 x + a_3 x^2 + \cdots$$
$$g_1(x) = \frac{g(x) - b_0}{x} = b_1 + b_2 x + b_3 x^2 + \cdots$$

And, as f_1 and g_1 are again power series with the *same* coefficients, they have the same radius of convergence (exercise!) and so are continuous at 0. This means we can re-run our trick:

$$\lim f_1(c_n) = f_1(\lim c_n) = f_1(0) = a_1$$
$$\lim g_1(c_n) = g_1(\lim c_n) = g_1(0) = b_1$$

But now think a bit about f_1 and g_1: because $a_0 = b_0$, we have actually done exactly the same operation to each function: subtracted the same number and then divided by x. Thus, the fact that $f(c_n) = g(c_n)$ immediately implies that $f_1(c_n) = g_1(c_n)$. So, the two sequences we are taking the limit of are the same, meaning their limits are the same: $a_1 = b_1$.

You can imagine how we continue from here: induction! □

Corollary 17.1. *Let $f(x) = \sum_{k\geq 0} a_k x^k$ and $g(x) = \sum_{k\geq 0} b_k x^k$ be two power series, which are equal on some neighborhood of zero. Then they are identical, and $a_k = b_k$ for all k.*

Proof. Let $\epsilon > 0$ be some small number where $f(x) = g(x)$ for all $|x| < \epsilon$. Take the sequence $1/n$ and truncate the first finitely many terms until $1/n < \epsilon$, producing a nonzero sequence converging to zero fully contained in $(-\epsilon, \epsilon)$. Then by our assumption f and g agree on this sequence, so they are equal by the Identity Theorem. □

18. Important Theorems

Highlights of this Chapter: we prove two foundational results about continuous functions whose proofs have several steps in common:

- Continuous Functions are determined by their values on dense sets.
- The Extreme Value Theorem: a continuous function achieves a max and min on any closed interval.
- The Intermediate Value Theorem: a continuous function must take every value between $f(a)$ and $f(b)$ on the interval $[a,b]$.

Just like we have seen various 'proof styles' for sequences (recurrent themes in proofs, like 'an $\epsilon/2$ argument') one of the biggest takeaways of this section is a proof technique for working with continuous functions. It has three steps, summarized below:

- Use whatever information you have to start, to construct a sequence of points.
- Use Bolzano Weierstrass to find a convergent subsequence.
- Apply f to that sequence and use continuity to know the result is also convergent.

This is to vague on its own to be useful, but in reading the proofs of the boundedness theorem, the extreme value theorem, and the intermediate value theorem below, look out for these three recurrent steps.

18.1. Dense Sets

Functions determined by values on dense set

Lemma 18.1. *If f is a continuous function such that $f(r)=0$ for every rational number r, then $f=0$ is the zero function.*

Proof. Let f be such a function, and $a\in\mathbb{R}$ any real number. Then there is a sequence r_n of rational numbers converging to a. Given that f is zero on all rationals, we see that $f(r_n)=0$ for all n. Thus $f(r_n)$ is the *constant* zero sequence, and so its limit is zero:

$$\lim f(r_n)=\lim 0=0$$

But, since f is assumed to be continuous, we know that we can move the limit inside of f:

$$0 = \lim f(r_n) = f(\lim r_n) = f(a)$$

Thus $f(a) = 0$, and since a was arbitrary, we see f is the constant function equal to zero at all real numbers. □

Proposition 18.1 (Equal on Rationals $\implies$ Equal)**.** *Let f, g be continuous functions such that for all $r \in \mathbb{Q}$ they are equal: $f(r) = g(r)$. Then in fact, $f = g$: for all $x \in \mathbb{R}$, $f(x) = g(x)$*

Proof. Since f and g are continuous, the function $h = f - g$ is continuous using the theorems for field operations. And, since $f(x) = g(x)$ for all rational x, we see $h(x) = 0$ on the rationals. Thus, by **?@prp-zero-on-rationals**, h itself must be the zero function on all of $\mathbb{R}$. Thus for every x, $h(x) = f(x) - g(x) = 0$, or rearranging,

$$\forall x,\ f(x) = g(x)$$

□

This has a the pretty significant consequence that if we have a function and we know it is continuous, then being able to calculate its values at the rational numbers is good enough to completely determine the function on the real line. In particular, this can be used to prove various *uniqueness results*: you can show a certain function is uniquely defined if you can prove that its definition implies (1) continuity and (2) determines the rational points (or more generally, the values on a dense set).

Theorem 18.1 (Equal on a Dense Set $\implies$ Equal)**.** *Continuous functions are determined by their values on a dense subset of their domains: if $f, g : X \to \mathbb{R}$ and $D \subset X$ is dense with $f = g$ on D, then $f(x) = g(x)$ for all $x \in X$.*

Exercise 18.1. Prove this (following the ideas for the special case of rationals)

We will use this property in understanding exponential functions (where their value at rational numbers are determined by powers and roots) and trigonometric functions (whose values on certain dyadic multiples of π are determined by the half-angle identities.)

There are many useful theorems of this type, that check a property of a function on a dense set and use it to conclude the same property holds generally. We give two more examples below, that prove useful in upcoming work

Proposition 18.2. *If f is continuous and monotone on a dense set, then it is monotone on its entire domain.*

Proof. Assume for contradiction that f is monotone increasing on a dense set D in the domain of f, but that it is *not* monotone increasing on the entire domain. This means that there exists a pair $x < y$ in the domain where $f(x) > f(y)$, call the difference $f(y) - f(x) = D$ and set $\epsilon = D/3$. Then by continuity of f there is a δ_x about x such that $|x - a| < \delta_x$ implies $|f(x) - f(a)| < \epsilon$, and similarly for a δ_y about y.

We are going to use these δ neighborhoods to choose points in $d_x, d_y \in D$ near x and y, so we need to be careful: we wish to ensure $d_x < d_y$ just as $x < y$, so we want our δ neighborhoods to not overlap. And since we got the values δ_x and δ_y from continuity we don't have any control over their size, so they might be rather large! But this is no serious problem, we can easily shrink them if needed: if $\delta = |y - x|$ we can set δ_x to be the minimum of its original value and $\delta/2$, and same for δ_y.

Now, by the density of D in the domain, there is a $d_x \in D$ within δ_x of x and a $d_y \in D$ within δ_y of y. Together with the above this implies that $f(d_x)$ is *at least* $f(x) - \epsilon$ and $f(d_y)$ is *at most* $f(y) + \epsilon$. But the distance between $f(x)$ and $f(y)$ was $D = 3\epsilon$, so $f(d_x) - f(d_y) \geq \epsilon > 0$ and hence $f(d_x) > f(d_y)$. But this contradicts the fact that f is increasing on D as $d_x < d_y$. □

Exercise 18.2. Modify the above proof to show that if f is continuous and *strictly increasing* or *strictly decreasing* on a dense set, then it is strictly increasing/decreasing everywhere on its domain.

Exercise 18.3. If f is continuous and convex on a dense set, then it is convex on its entire domain.

18.2. Extreme Values

Proposition 18.3 (Continuous on Closed Interval $\implies$ Bounded)**.** *Let f be a continuous function on a closed interval $[a, b]$. Then the image $f([a, b])$ is bounded.*

Proof. Assume for the sake of contradiction that f is not bounded. Then for each $n \in \mathbb{N}$ there must be some $x_n \in [a, b]$ where $|f(x_n)| > n$. This sequence $\{x_n\}$ need not be convergent, but it lies in the interval $[a, b]$ so it is bounded, and thus contains a convergent subsequence x_{n_k} by Bolzano Weierstrass. Say $x_{n_k} \to x$. Then since $a \leq x_{n_k} \leq b$ for all k, by the inequalities of limits we see $a \leq x \leq b$ so the limit x lies in the interval $[a, b]$ as well.

But what is the value $f(x)$? Since f is continuous and $x_{n_k} \to x$ we know that

$$f(x_{n_k}) \to f(x)$$

But for each k, x_{n_k} has the property that $f(x_{n_k}) > n_k$ *by definition.* Thus, the sequence $f(x_{n_k})$ is not bounded, and cannot be convergent (since all convergent sequences are bounded). This is a contradiction, as it implies that $f(x)$ is not defined, even though we have assumed f is defined on the entire interval $[a,b]$.

Thus, no such sequence x_n is possible, and so there must be some n where $|f(x)| < n$ for all $x \in [a,b]$. That is, f must be bounded on $[a,b]$. $\square$

Building off this result, one can prove that a continuous function actually *achieves* its upper and lower bounds on any closed interval. This result will play a role several times across the theory of functions and derivatives, so we give it a memorable name: the *extreme value theorem* (as maxima and minima taken collectively are called *extrema*).

Theorem 18.2 (Extreme Value Theorem). *Let f be a continuous function on a closed interval $[a,b]$. Then f achieves a maximum and minimum value: that is, there exists a point p where $f(p) \geq f(x)$ for all $x \in [a,b]$, and a q where $f(q) \leq f(x)$ for all $x \in [a,b]$.*

Proof. We show f achieves a maximum, and leave the minimum case as an exercise. Let f be continuous on $[a,b]$ and let $R = \{f(x) \mid x \in [a,b]\}$ be the set of outputs, or the *range* of f. Since f is bounded we see that R is a bounded subset of $\mathbb{R}$, and so by completeness

$$m = \inf R \qquad M = \sup R$$

must exist. Our goal is to find values $x_m, x_M \in [a,b]$ for which the infimum and supremum are realized:

$$f(x_m) = m \qquad f(x_M) = M$$

Here we show this holds for the supremum, the infimum is left as an exercise below. Since M is the supremum, for any $\epsilon > 0$ we know that $M - \epsilon$ is *not* an upper bound for $R = \{f(x) \mid x \in [a,b]\}$: thus there must be some x where $f(x) > M - \epsilon$. So letting $\epsilon = 1/n$ each n, let x_n be a point where $M - \frac{1}{n} < f(x_n) \leq M$. As $n \to \infty$ we know $M - \frac{1}{n} \to M$ and so by the squeeze theorem we see that $f(x_n) \to M$ as well.

We don't know that the points x_n themselves converge, but we do know that this entire sequence lies inside the closed interval $[a,b]$ so its bounded and Bolzano Weierstrass lets us extract a convergent subsequence $x_{n_k} \to x$. And as $a \leq x_{n_k} \leq b$ it follows that the limit $x \in [a,b]$ as well. Because subsequences of a convgent sequence converge to the same limit, we know that $f(x_{n_k})$ is convergent, and still has limit M. But now we can finally use continuity!

Since f is continuous, we know $\lim f(x_n) = f(\lim x_n)$, and so $M = f(x)$. Thus we managed to find a point $x \in [a,b]$ where $f(x)$ is the supremum: $f(x)$ is an upper bound for all possible values of f on $[a,b]$, which by definition means its the max value! So f achieves a maximum on $[a,b]$. $\square$

Exercise 18.4. Complete the proof by showing a continuous function on a closed interval achieves a minimum.

18.3. Intermediate Values

The intermediate value theorem is the rigorous version of "you can draw the graph of a continuous function without picking up your pencil".

One note: in the statement below we use the phrase y is between $f(a)$ and $f(b)$ as a shorthand to mean that either $f(a) < y < f(b)$ or $f(b) < y < f(a)$ (as we don't know if $f(a)$ or $f(b)$ is larger).

Theorem 18.3 (The Intermediate Value Theorem)**.** *Let f be a continuous function on the interval $[a, b]$, and let y be any number between $f(a)$ and $f(b)$. Then there exists an x between a and b such that $y = f(x)$.*

Proof. Without loss of generality we will assume that $f(a) < f(b)$ so that y lies in the interval $[f(a), f(b)]$ (the other case is analogous, we just instead must write the interval $[f(b), f(a)]$). We wish to find a point $x \in [a, b]$ where $f(x) = y$, so we start by defining the set of points where $f(x)$ is less than *or equal to* y:

$$S = \{x \in [a, b] \mid f(x) \leq y\}$$

This set is nonempty: $a \in S$ as $f(a) < y$ by assumption. And its bounded above by b: if $x \in S$ then $x \in [a, b]$ so $x \leq b$ by definition. Thus, the supremum $\sigma = \sup S$ exists, and $\sigma \in [a, b]$. We will show that $f(\sigma) = y$, by showing both inequalities $f(\sigma) \leq y$ and $f(\sigma) \geq y$.

First, we show $\leq$. Since σ is the supremeum, for each n we know that $\sigma - \frac{1}{n}$ is not an upper bound, and so there must be an point $x_n \in (\sigma - 1/n, \sigma)$ where $f(x_n) \leq y$. The squeeze theorem assures that $x_n \to \sigma$, and the continuity of f assures that $f(x_n)$ converges (since x_n does). But for all n we know $f(x_n) \leq y$, so by the inequalities of limits we also know $\lim f(x_n) = f(\sigma) \leq y$.

Next, we show $\geq$. First note that $\sigma \neq b$ as $f(\sigma) \leq y$ but $f(b) > y$. So, $\sigma < b$ and so after truncating finitely many terms, the sequence $x_n = \sigma + 1/n$ lies strictly between σ and b. Since this sequence is *greater* than the upper bound σ, we know that none of the x_n are in S and so $f(x_n) > y$ by definition, for all n. But as $n \to \infty$ the sequence of x_n's is squeezed to converge to σ, and so by continuity we know

$$f(\sigma) = f(\lim x_n) = \lim f(x_n)$$

Applying the inequalities of limits this time yields the reverse: since for all n we know $f(x_n) > y$, it follows that $\lim f(x_n) \geq y$ so $f(\sigma) \geq y$.

Putting these together we know that $f(\sigma)$ is some number which must simultaneously by $\geq y$ and $\leq y$. The only number satisfying both of these inequalities is y itself, so

$$f(\sigma) = y$$

$\square$

Corollary 18.1. *Continuous image of a closed interval is a closed interval.*

Historically, the intermediate value theorem was one of the reasons for developing much of analysis: mathematicians knew that *whatever the correct formal definition of continuity was, it should certainly imply this!* So, our proof of the intermediate value theorem (which embodies the intuitive notion of continuity) may be seen as evidence that we have chosen good definitions of continuity and convergence: they work as we expect!

Remark 18.1. It may seem at first that the intermediate value theorem is equivalent to continuity: if a function satisfies the intermediate value property, then its continuous. Try to prove it! Where do you get stuck?

Example 18.1. Consider the following function

$$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Then f satisfies the conclusion of the intermediate value theorem on every closed interval, but f is not continuous at 0.

18.3.1. Useful Corollaries

Continuity is a strong constraint on a function, and the behavior of a continuous function at one or more points can often be used to gain information about nearby points. A direct corollary of the intermediate value theorem that is very useful is the special case when $y = 0$:

Corollary 18.2 (Positive to Negative implies Zero)**.** *If f is a continuous function on an interval and it is positive one endpoint and negative on the other, then f has a zero in-between.*

This suggests a means of finding the zeros of a function, which narrows in on them exponentially fast! Called "bisection": find any two points where function changes sign. Divide region in half, evaluate at midpoint. Keep interval with different sign endpoints, repeat.

This argument also suggests simple proofs of various other theroems proving the *existence* of a point in the domain having some specified property. Here we give a classic *fixed point theorem* as an example

Proposition 18.4 (A Fixed Point Theorem)**.** *If $f\colon [0,1] \to [0,1]$ is continuous then there is some $x \in [0,1]$ with $f(x) = x$.*

Proof. Consider the function $g(x) = f(x) - x$. Since $f(0) \geq 0$ we know $g(0) = f(0) - 0 \geq 0$, and as $f(1) \leq 1$ we similarly know that $g(1) = f(1) - 1 \leq 0$. $\square$

Like fixed points above, we can also use the IVT to prove the existence of solutions to various equations. Here, we use it to prove the existence of the square root of 2 - a calculation that took us quite some fiddling around with algebra and the Archimedean property originally!

Example 18.2 (Existence of $\sqrt{2}$)**.** The function $f(x) = x^2$ is continuous on $\mathbb{R}$. But $f(1) = 1$ which is less than 2 and $f(2) = 4$ which is greater than 2. Thus, by the intermediate value theorem there must be some $s \in [1, 2]$ such that $f(s) = s^2 = 2$, so $s = \sqrt{2}$.

Exercise 18.5 (Existence of n^{th} roots.)**.** For every $x \geq 0$ there exists a unique positive number y such that $y^n = x$.

Its worth mentioning one additional corollary of the interemediate value theorem together with the extreme value theorem, which helps us understand the *ranges* of continuous functions

Corollary 18.3. *If f is a continuous function and $I \subset \mathbb{R}$ is a closed interval, then $f(I)$ is an interval.*

Here we allow the degenerate case $[a, a] = \{a\}$ to count as an interval, if f is constant.

18.4. Uniform δ on Closed Intervals

In general the definition of continuity takes place at *each* x individually, so for a fixed ϵ we might find different δ's depending on which point we are at. This can be theoretically bothersome sometimes, as it would be *much easier* to just pick a δ once and for all and use it in an entire problem.

The following theorem shows this is indeed possible

Theorem 18.4 (Continuous + Closed $\implies$ Uniform)**.** *Let f be a continuous function defined on the closed interval $[a, b]$. Then f is in fact* uniformly continuous *on this interval.*

Proof. Assume for the sake of contradiction that f is not uniformly continuous, and fix $\epsilon > 0$. Then there is no fixed δ that works, so for any proposed δ, there must be some a where it fails.

We can use this to produce a sequence: for $\delta = 1/n$ let $a_n \in I$ be a point where this δ fails: there is some x_n within $1/n$ of a_n but $|f(x_n) - f(a_n)| > \epsilon$.

Thus, in fact we have two sequences x_n and a_n! We know very little about either except that they are in a closed interval I, so we can apply Bolzano Weierstrass to get convergent subsequences (we have to be a bit careful here, see the exercise below).

We will call the subsequences X_n and A_n (with capital letters). Now that we know they both converge, we can see that they also have the same limit: (as, by construction $|X_n - A_n| < \frac{1}{n}$). Call that limit L.

Then since f is continuous at L, we know that

$$\lim f(X_n) = f(\lim X_n) = f(L) = f(\lim A_n) = \lim f(A_n)$$

Thus, $\lim f(X_n) - f(A_n) = 0$. However this is impossible, since for all values of n we know $|f(X_n) - f(A_n)| > \epsilon$! This is a contradiction, and thus there must have been some uniform δ that worked all along. $\square$

In proof, use that we can simultaneously apply bolzano weierstrass to two sequences: this appears as an Exercise 9.14 back in the chapter on subsequences. If you didn't do it then, you should prove this for yourself now.

19. ⋆ Uniform Continuity

Continuity is a local property: to verify that a function is continuous, we look near individual points. But what if we want a *global* notion — one that controls the behavior of a function across its entire domain at once?

This brings us to **uniform continuity**, a stronger condition that ensures a single δ works for all points in the domain. While every uniformly continuous function is continuous, the reverse is not always true.

In this chapter, we will: - Define uniform continuity and explore how it differs from ordinary continuity. - Test whether standard operations like addition, multiplication, and inversion preserve uniform continuity. - Prove two deep results that make uniform continuity especially useful: that it ensures *extendability* and automatically holds for continuous functions on closed intervals.

Definition 19.1 (Uniform Continuity: $\epsilon - \delta$)**.** A function f is uniformly continuous on a domain $D \subset \mathbb{R}$ if for every ϵ there exists a δ such that for any $x, y \in D$ with $|x - y| < \delta$, it follows that $|f(x) - f(y)| < \epsilon$.

Here's an example showing how to use the definition, proving x^2 is uniformly continuous on an interval.

Example 19.1. $f(x) = x^2$ is uniformly continuous on the interval $[1, 3]$.

Scratch. Here's some scratch work: let $\epsilon > 0$. Then at any a we see that $|f(x)-f(a)| = |x^2 - a^2| = |x + a||x - a|$. If $|x - a| < \delta$ and we want $|f(x) - f(a)| < \epsilon$, this tells us that we want

$$|x + a|\delta < \epsilon$$

We don't know what x and a are, but we do know they are points in the interval $[1, 3]$! So, the smallest $x + a$ could be is $1 + 1 = 2$, and the biggest is $3 + 3 = 6$. This means that

$$|x + a|\delta \leq 6\delta$$

So, if we can make $6\delta < \epsilon$, we are good! This is totally possible: just set $\delta = \epsilon/6$. Below is the rigorous proof. □

Rigorous. Let $\epsilon > 0$, and set $\delta = \epsilon/6$. Note that for any $a \in [1, 3]$ and any x within δ of a, we know $a \leq 3$ and $x \leq 3$ so $x + a \leq 6$. But this implies that

$$|x^2 - a^2| = |x + a||x - a| \leq 6|x - a| < 6\delta < 6\frac{\epsilon}{6} = \epsilon$$

And so f is uniformly continuous, as this single choice of δ works for every point $a \in [1, 3]$. □

For normal continuity, we had a way to test using sequences. This proved quite useful since we are *so good* at working with sequences these days. There is an analog for uniform continuity as well

Exercise 19.1 (Sequences and Uniform Continuity)**.** A function f is uniformly continuous if and only if for every pair of sequences u_n, v_n in the domain with $\lim u_n - v_n = 0$, then $\lim f(u_n) - f(v_n) = 0$.

Uniform continuity is stricter than regular continuity: there are functions which are continuous but are *not* uniformly continuous. Here we see x^2 is such an example, using the sequence criterion

Example 19.2. The function $f(x) = 1/x$ is continuous, but not uniformly continuous on $(0, 1)$. Looking at the sequence $1/n$ we see $f(1/n) = 1/(1/n) = n$. So, consider the two sequences $s_n = 1/(n + 1)$ and $t_n = 1/n$. These have $s_n - t_n \to 0$ by the limit theorems (as each individually goes to zero) yet $f(s_n) - f(t_n) = (n + 1) - n = 1$ is a constant sequence *not* converging to zero.

The sequence $1/n$ used in this example provides a hint of one way to detect uniformly continuous functions: $1/n$ is Cauchy but $f(1/n)$ was not, and we were able to use this to show f was *not* uniformly continuous.

These examples show that uniform continuity is genuinely stronger than ordinary continuity. In particular, the failure of uniform continuity for f(x) = 1/x on (0,1) highlights that local control is not enough — behavior near the edges matters, and if the function "does something crazy" (here a vertical asymptote) we won't be able to find a *uniform* δ.

Theorem 19.1 (Uniformly Continuity Preserves Cauchy Sequences)**.** *If f is uniformly continuous and x_n is cauchy, then $f(x_n)$ is cauchy.*

Proof. Let x_n be an arbitrary Cauchy sequence in the domain of f, and choose arbitrary $\epsilon > 0$. Then by uniform continuity there is a δ such that for $|x - y| < \delta$ we know $|f(x) - f(y)| < \epsilon$. Since x_n is Cauchy, given this δ we can find an N such that $n, m > N$ implies $|x_n - x_m| < \delta$, and hence $|f(x_n) - f(x_m)| < \epsilon$. But this is precisely the definition of $\{f(x_n)\}$ being a Cauchy sequence, so we are done. □

Great way to check if a function is *not* uniformly continuous: can you find a cauchy seq taken to a non-cauchy sequence?

Example (PICTURE) functions like $\sin(1/x)$ are also not uniformly continuous on $(0, 1)$ even though it is bounded.

WARNING: does not work in reverse: the function $x \mapsto x^2$ takes Cauchy sequences to cauchy seqs but is *not* uniformly continuous.

Definition 19.2 (Cauchy Continuous Functions)**.** A real valued function f on a domain $D \subset \mathbb{R}$ is Cauchy Continuous if for every cauchy sequence $\{d_n\}$ in D, the sequence $f(d_n)$ is also Cauchy.

19.1. Properties of Uniformly Continuous Functions

In a previous chapter, we showed that continuous functions behave well under addition, multiplication, and composition. Now we ask: do these same operations preserve *uniform* continuity? The answers are a little more nuanced. Let's go through them carefully.

Proposition 19.1 (Constant Multiples of Uniformly Continuous Functions)**.** *Let f be uniformly continuous, and $k \in \mathbb{R}$. Then kf is uniformly continuous.*

Proof. If $k = 0$ then kf is the constant zero function, so we ignore that case. For $k \neq 0$, let $\epsilon \geq 0$ consider $\epsilon/|k|$ and take the corresponding uniform δ for f. For $|x - y| < \delta$ we see $|f(x) - f(y)| < \epsilon/|k|$, and so

$$|kf(y) - kf(x)| < |k||f(y) - f(x)| \leq |k|\frac{\epsilon}{|k|} = \epsilon$$

□

Exercise 19.2 (Sums of Uniformly Continuous Functions)**.** Let f and g be uniformly continuous. Then $f + g$ is uniformly continuous.

From these it follows that $f - g$ is uniformly continuous (as its equal to $f + (-1)g$) and $af + bg$ are for any $a, b \in \mathbb{R}$ are uniformly continuous as well. It might be tempting to believe, after seeing the above proofs that all of the limit laws should have analogs for uniform continuity, just as they did for continuity. But this is not true!

Example 19.3 (Reciprocals need not be Uniformly Continuous)**.** The function $y = x$ is uniformly continuous and nonzero on $(0, 1)$ but its reciprocal $f(x) = 1/x$ is not.

Proof. Fx any $\delta > 0$, and note that given any $1/n < \delta$ we have

$$\frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)} < \frac{1}{n} < \delta$$

but applying f,

$$f\left(\frac{1}{n+1}\right) - f\left(\frac{1}{n}\right) = n + 1 - n = 1$$

Thus, fixing any $\epsilon < 1$ there can't be a uniform δ, as its always possible to find points separated by less than δ mapped to points separated by a distance of 1. □

This generalizes directly to reciprocals: if f is uniformly continuous then $1/f$ need not be

Exercise 19.3. Let f be uniformly continuous and bounded away from zero: $f(x) \geq b > 0$ for all x in the domain. Prove that $1/f$ is uniformly continuous.

What about products? Again we need a boundedness assumption:

Exercise 19.4 (Uniform Continuity and Products)**.** Let f and g be uniformly continuous bounded functions with the same domain. Then $f(x)g(x)$ is uniformly continuous.

Proof. Since f, g are bounded we chan choose an $M > 0$ with $|f(x)| < M$ and $|g(x)| < M$ for all x in the domain. Let $\epsilon > 0$ be arbitrary, and using uniform continuity for f, g choose δ_f such that $|x - y| < \delta_f$ implies $|f(x) - f(y)| < \epsilon/2M$ and an analogous δ_g for g. Set $\delta = \min\{\delta_f, \delta_g\}$ and for any x, y with $|x - y| < \delta$ we compute

$$|f(y)g(y) - f(x)g(x)| = |f(y)g(y) - f(y)g(x) + f(y)g(x) - f(x)g(x)|$$

$$\leq |f(y)g(y) - f(y)g(x)| + |f(y)g(x) - f(x)g(x)| = |f(y)||g(y) - g(x)| + |g(x)||f(y) - f(x)|$$

As both $|f|$ and $|g|$ are bounded by M, this is less than or equal to $M(|g(y) - g(x)| + |f(y) - f(x)|)$, and each of these terms is less than $\epsilon/2M$ by hypothesis, so

$$|f(y)g(y) - f(x)g(x)| \leq M\left(\frac{\epsilon}{2M} + \frac{\epsilon}{2M}\right) = \epsilon$$

as required. □

Exercise 19.5. Show that this boundedness assumption is necessary by giving an example of two uniformly continuous functions whose product is *not* uniformly continuous.

Proposition 19.2 (Composition of Uniformly Continuous)**.** *Let f and g be uniformly continuous functions. Then the composition $f \circ g(x) = f(g(x))$ is uniformly continuous.*

Proof. Choose $\epsilon > 0$ and let δ_f be a uniform delta for $f(x)$. Use this to select a uniform δ_g for g, such that whenever $|x - y| < \delta_g$, we have $|g(x) - g(y)| < \delta_f$. This turns out to be the right uniform value for the composition $f \circ g$, as $|g(x) - g(y)| < \delta_f \implies |f(g(x)) - f(g(y))| < \epsilon$. □

Like reciprocals, inverses pose a problem:

Exercise 19.6 (Inverses and Uniform Continuity)**.** Give an example of a uniformly continuous function whose inverse is not uniformly continuous.

Exercise 19.7. Prove that $f(x) = x^2$ is not uniformly continuous on the entire real line, using either the $\epsilon - \delta$ definition or the sequence definition.

19.2. Continuous Extension

So far uniform continuity seems to be a slightly more restrictive definition (requiring one to prove their choice of δ works *everywhere*) with consequently weaker theorems (inverses and products need not be *uniformly* continuous, even though they are continuous, for one). So why would one care about this harder to verify and harder to work with notion of continuity? The main reason is that the *stricter* definition of uniform continuity allows us to prove some (very useful) things which are just not true about the standard version! The chief among these is perhaps the *continuous extension* theorem.

Roughly speaking, if a function is uniformly continuous on an open interval, then we can define its values at the endpoints in a way that makes it continuous on the closed interval. This is not always possible for merely continuous functions, as we saw with f(x) = 1/x on (0,1). Uniform continuity makes all the difference.

Theorem 19.2 (Extending Uniform Continuity to Endpoints)**.** *If $f : (a, b) \to \mathbb{R}$ is uniformly continuous, then there exists a* continuous extension *of $\tilde{f}$ of f to $[a, b]$.*

We could stop to prove this here, but in fact the same technique proves a more general extension theorem of which this is a special case:

Theorem 19.3 (The Continuous Extension Theorem)**.**

Proof. Proof sketch: D dense in X, define $f(x)$ by $\lim f(d_n)$ for $d_n \to x$. Need to check (1) this defines a value, (2) its well defined, independent of sequence.

For (1): if $d_n \to x$ then d_n is convergent, hence Cauchy. f is uniformly continuous so it takes Cauchy sequences to cauchy sequences. Thus $f(x_n)$ is Cauchy, hence convergent.

Next for (2): if $c_n, d_n \to x$ are two such sequences, make the interleaved sequence $c_1, d_1, c_2, d_2, c_3, d_3 \cdots$. This converges to x as well so is Cauchy. Thus applying f yields a cauchy (hence convergent) sequence, and all subsequences have the same limit. Since c_n and d_n are subsequences, we see $f(c_n)$ and $f(d_n)$ converge to the same value. □

Note in the proof above we only used one property of uniformly continuous functions: that they take cauchy sequences to cauchy sequences. So this actually applies more generally, to *Cauchy Continuous* functions.

Corollary 19.1 (Continuous Extension of Cauchy Continuous Functions)**.** *If f is Cauchy continuous on a set D which is dense in X, then there exists a unique continuous extension $\tilde{f}$ of f to X.*

19.3. Continuous on a Closed Interval

The continuous extension theorems provide a first (of several) motivations for being interested in this stronger notion of continuity. Hence its useful to develop some results for telling when a function which known a priori only to be continuous is in fact uniformly continuous. The most useful of these provides a surprisingly simple condition: so long as the domain is a closed interval, continuity and uniform continuity are equivalent!

Theorem 19.4 (Continuous + Closed $\implies$ Uniform)**.** *Let f be a continuous function defined on the closed interval $[a, b]$. Then f is in fact* uniformly continuous *on this interval.*

Proof. Assume for the sake of contradiction that f is not uniformly continuous, and fix $\epsilon > 0$. Then there is no fixed δ that works, so for any proposed δ, there must be some a where it fails.

We can use this to produce a sequence: for $\delta = 1/n$ let $a_n \in I$ be a point where this δ fails: there is some x_n within $1/n$ of a_n but $|f(x_n) - f(a_n)| > \epsilon$.

Thus, in fact we have two sequences x_n and a_n! We know very little about either except that they are in a closed interval I, so we can apply Bolzano Weierstrass to get convergent subsequences (we have to be a bit careful here, see the exercise below).

We will call the subsequences X_n and A_n (with capital letters). Now that we know they both converge, we can see that they also have the same limit: (as, by construction $|X_n - A_n| < \frac{1}{n}$). Call that limit L.

Then since f is continuous at L, we know that

$$\lim f(X_n) = f(\lim X_n) = f(L) = f(\lim A_n) = \lim f(A_n)$$

Thus, $\lim f(X_n) - f(A_n) = 0$. However this is impossible, since for all values of n we know $|f(X_n) - f(A_n)| > \epsilon$! This is a contradiction, and thus there must have been some uniform δ that worked all along. $\square$

In proof, use that we can simultaneously apply bolzano weierstrass to two sequences: this appears as an Exercise 9.14 back in the chapter on subsequences. If you didn't do it then, you should prove this for yourself now.

Exercise 19.8 (Periodic Continuous Functions are Uniformly Continuous)**.** Let f be a periodic continuous function on $\mathbb{R}$. Then f is uniformly continuous.

20. Elementary Functions

> Highlights of this Chapter: we introduce the idea of defining functions by a *Functional Equation* specifying how a function should behave instead of specifying how to compute it. Following this approach, we give rigorous definitions for exponentials logarithms and trigonometric functions, and investigate some of their consequences. With these definitions in hand, we are able to define the field of *Elementary Functions*, familiar from calculus and the sciences.

At the heart of real analysis is the study of functions. But which functions should we study? Polynomials are a natural class built from the field operations, and power series are a natural thing to look at given polynomials and the concept of a limit. But there are many, *many* other functions out there, and we should wonder which among them are worthy of our attention. Looking to history as a guide, we see millennia of use of trigonometric functions, and centuries of use of exponentials and logarithms. Indeed these functions are not only important to the origins of analysis but also to its modern development. In this chapter we will not focus on *how to compute* such functions, but rather on the more pressing question of *how to even define them*: if all we have available to us are the axioms of a complete ordered field how do we rigorously capture aspects of circles in the plane (trigonometry) or continuous growth (exponentials)? The key is the idea of a *functional equation*: something that will let us define a function by how it behaves, instead of by directly specifying a formula to compute it.

20.1. Warm Up: What is Linearity?

We know how to express linear functions already using the field axioms, as maps $f(x) = kx$ for some real number k. To speak of linear functions *functionally* however, we should not give a definition telling us how to compute their values (take the input, and multiply by a fixed constant k) but rather by *what they're for*: by the defining property of linearity.

The most important property of a linear function is that it distributes over addition (think of how we use linear maps, say, in Linear Algebra). So, in the 1800

Definition 20.1 (Cauchy's Functional Equation for Linearity)**.** A function $f : \mathbb{R} \to \mathbb{R}$ satisfies Cauchy's functional equation if for all $x, y \in \mathbb{R}$,

$$f(x+y) = f(x) + f(y)$$

Note it follows for any finite sum $x_1 + x_2 + \cdots + x_n$, one can 'distribute' a solution to Cauchy's functional equation, by induction: $f(x_1+x_2+\cdots+x_n) = f(x_1)+f(x_2)+\cdots+f(x_n)$. Cauchy's idea works: we can completely characterize the concept of *Linear Function* from $\mathbb{R}$ to $\mathbb{R}$ via this functional equation and continuity.

Theorem 20.1 (Characterizing Linear Functions)**.** *If f is a continuous solution to Cauchy's functional equation, then $f(x) = kx$ for some $k \in \mathbb{R}$.*

Exercise 20.1. Prove Theorem 20.1, following the outline below.

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function where $f(x+y) = f(x) + f(y)$.

- Prove that $f(n) = nf(1)$ for all $n \in \mathbb{N}$.
- Extend this to negative integers.
- Show that $f(1/n) = \frac{1}{n}f(1)$ for $n \in \mathbb{N}$ *Hint: use that $\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n} = 1$
- From the above, deduce that for rational $r = p/q$, $p \in \mathbb{Z}$ $q \in \mathbb{N}$ that $f(r) = rf(1)$.
- Now use continuity! If $k = f(1)$, then $f(r) = kr$ on $\mathbb{Q}$...

20.2. Exponentials

Exponential functions occur all across the math and sciences, representing any kind of growth that *compounds multiplicatively* as time *progresses linearly.* THat is, the core feature of exponentials underlying their ubiquity is the *law of exponents* $a^{m+n} = a^m a^n$ turning the addition of m and n into the multiplication of a^m and a^n. Following Cauchy's lead, we will single this out and use it to define a class of functions via *functional equation.*

Definition 20.2 (The Law of Exponents)**.** A function $E : \mathbb{R} \to \mathbb{R}$ satisfies the *law of exponents* if for every $x, y \in \mathbb{R}$

$$E(x+y) = E(x)E(y)$$

An *exponential function* is a continuous nonconstant solution to the law of exponents.

This just rigorously spells out *what we want exponential functions to be.* We still have to prove they exist! But before doing that, we pause to gain some comfort with the functional equation definition, and derive a few basic properties that exponentials must have.

Proposition 20.1. *If E satisfies the law of exponents and evaluates to zero at any point, then E is the zero function.*

Proof. Let E be an exponential function and assume there is some $z \in \mathbb{R}$ such that $E(z) = 0$. Then for any $x \in \mathbb{R}$ we may write $x = x - z + z = (x - z) + z = y + z$ for $y = x - z \in \mathbb{R}$. Evaluating $E(x)$ using the law of exponents,

$$E(x) = E(y + z) = E(y)E(z) = E(y) \cdot 0 = 0$$

□

Proposition 20.2. *Prove that if E is any exponential function, then $E(0) = 1$, and that $E(-x) = 1/E(x)$.*

Proof. The number 0 has the property that $0 + 0 = 0$. Plugging this into the exponential property, we find

$$E(0) = E(0 + 0) = E(0)E(0)$$

By the previous proposition, we know $E(0)$ is nonzero, so we can divide by it, leaving $1 = E(0)$. For the second part, we begin with the identity $x+(-x) = 0$. Exponentiating gives

$$1 = E(0) = E(x + (-x)) = E(x)E(-x)$$

We can then divide by $E(x)$ giving the result $E(-x) = \frac{1}{E(x)}$. □

Exercise 20.2. If $E(x)$ is an exponential and $s \neq 0$ is a real number, then $x \mapsto E(sx)$ is also an exponential function.

20.2.1. Existence

Here we show that exponential functions exist, and fully characterize them. We've already done plenty of work understanding rational and irrational powers, so we can make perfect sense of the expression a^x for arbitrary $a > 1$ and real x. Furthermore in algebra and calculus classes we have *certainly* treated such expressions as exponentials: we've used the law of exponents on their powers without worry! But now in our rigorous mindset there is much more to do, we need to confirm that our rather complicated definition of $x \mapsto a^x$ actually is (1) continuous nonconstant and (2) satisfies the laws of exponents.

To do so we make use of all our previous work with exponents: namely the following facts (the first two are from our inital investigation into numbers and operations, the second two from when we studied *Monotone Convergence*)

- If $r = p/q$ then a^r is defined as $a^{p/q} = \sqrt[q]{a^p}$.
- If r, s are rational numbers $a^{r+s} = a^r a^s$, so powers satisfy the law of exponents on *rational inputs.*

- If r_n is a monotone increasing sequence of rational numbers a^{r_n} converges.
- If r_n is a sequence of positive rationals with $r_n \to 0$ then $a^{r_n} \to 1$.

From these we can prove an important lemma that helps us make a rigorous definition of the function a^x:

Lemma 20.1 (Irrational Powers). *If $x \in \mathbb{R}$ and r_n is any monotone increasing sequence of rational numbers $r_n \nearrow x$, then a^{r_n} converges to the same limit.*

Proof. Let s_n and r_n be two monotone increasing sequences of rationals converging to x. We know (by monotone convergence) that a^{r_n} and a^{s_n} both converge, so let's name their limits $\lim a^{r_n} = R$ and $a^{s_n} = S$. We wish to prove $R = S$.

Defining $z_n = r_n - s_n$, note that $z_n \in \mathbb{Q}$ we can write $r_n = s_n + z_n$, and using the law of exponents for rational numbers,

$$a^{r_n} = a^{s_n + z_n} = a^{s_n} a^{z_n}$$

Applying the limit law for differences we see $\lim z_n = \lim r_n - \lim s_n = x - x = 0$, and so we know (from our earlier fact) that $a^{z_n} \to 1$. Thus all three of the sequences above converge, and we can use the limit law for products:

$$R = \lim a^{r_n} = \lim (a^{s_n} a^{z_n}) = (\lim a^{s_n})(\lim a^{z_n}) = (\lim a^{s_n})(\lim a^{z_n}) = (\lim a^{s_n})(1) = S$$

□

Because the limiting value of a^x does not depend on *which sequence we take*, we can define irrational powers by saying *take any monotone sequence of rationals and compute the limit* without worrying if different people will get different answers.

Definition 20.3 (Raising to the Power of x). Given a positive $a \neq 1$, we define the function $x \mapsto a^x$ via

$$E(x) = \begin{cases} a^x & x \in \mathbb{Q} \\ \lim a^{r_n} & x \notin \mathbb{Q} \text{ for } r_n \in \mathbb{Q},\ r_n \nearrow x \end{cases}$$

Theorem 20.2 (Existence: Powers Are Exponentials). *Exponential functions exist. Precisely, any positive $a \neq 1$, $E(x) = a^x$ is a continuous function which satisfies the law of exponents $a^{x+y} = a^x a^y$ for all $x, y \in \mathbb{R}$.*

We prove these two claims separately. To simplify things, we work with the case $a > 1$: for $a < 1$ one can either perform analogous arguments, or write $a = 1/b$ fo$ $b > 1$, and work with $E(x) = 1/b^x$ where we already fully understand b^x.

Proposition 20.3. *The function a^x satisfies the law of exponents*

Laws of Exponents. Let $x, y \in \mathbb{R}$, we wish to show that $a^{x+y} = a^x a^y$. Note that if both x and y are rational we are done, as we already know that the law of exponents holds for rational powers. So, the interesting case is when at least one is irrational, where the function definition involves *limits.*

Let's continue with the case where both x and y are irrational (we'll see the case where only one is irrational also follows from the same logic). To define a^x and a^y we need to choose monotone sequences of rational numbers $x_n \to x$ and $y_n \to y$: then we have $a^x = \lim a^{x_n}$ and $a^y = \lim a^{y_n}$. Since both of these sequences converge, we can use the limit law for products to conclude

$$\lim a^{x_n} a^{y_n} = (\lim a^{x_n})(\lim a^{y_n}) = a^x a^y$$

But since x_n and y_n are rational numbers, we know the law of exponents holds for them: $a^{x_n} a^{y_n} = a^{x_n + y_n}$ for each n. By the limit law for sums we know that $x_n + y_n \to x + y$, but furthermore its a sequence of rational numbers (since x_n and y_n are) and its monotone increasing (since x_n and y_n are). This means (by definition!) $\lim a^{x_n + y_n} = a^{x+y}$.

Stringing these equalities together yields the law of exponents for x and y:

$$a^{x+y} = \lim a^{x_n + y_n} = \lim (a^{x_n} a^{y_n}) = (\lim a^{x_n})(\lim a^{y_n}) = a^x a^y$$

We can use the same argument when only one of the numbers is rational: if $x \in \mathbb{Q}$ but $y \notin \mathbb{Q}$, we still need to choose a monotone sequence $y_n \to y$ of rationals, but there's an obvious choice of sequence for x: just take the constant sequence $x, x, x, \ldots$: this is rational monotone (because its constant), and converges to x after all! Running the same argument as above with these two sequences yields $a^{x+y} = a^x a^y$ as well. □

Exercise 20.3. Prove the exponential function a^x for $a > 1$ is monotone increasing: that is, if $x < y$, $a^x \leq a^y$.

Hint: we know its monotone on rational inputs, so the interesting cases are again when at least one is irrational (and, the argument for both irrational can be generalized to include the other case). Write down monotone increasing sequences, truncate the sequences until you can insure $x_n < y_n$ for all n, and then apply the limit laws.

We will additionally need (for a later argument) that the exponential is *strictly increasing*: that is, if $x < y$ then $a^x < a^y$ (that is, the equals case is impossible), so we'll prove that now as well.

Lemma 20.2. *If $x < y$ then $a^x < a^y$, when $a > 1$.*

Proof. This is equivalant to showing $1 < a^y/a^x$, which, since we know the laws of exponents hold, means showing $1 < a^{y-x}$. That is, our problem is equivalent to proving for any $z > 0$ the exponential a^z is strictly greater than 1.

First note this is clearly true for rational $z = p/q$, as $a > 1$ implies $a^p > 1^p = 1$, which implies $\sqrt[q]{a^p} > \sqrt[q]{1} = 1$. For irrational z we proceed by choosing a rational sequence $z_n \to z$. By picking an epsilon (say $\epsilon = z$) we can truncate our sequence and after some point, and assume its always positive. So (possibly re-labeling the indices) we can assume without loss of generality $z_n > 0$ for all n. Since z is monotone increasing, we see $z_n \geq z_1$ for all n, and since the exponential is monotone for rational numbers,

$$z_1 \leq z_n \implies a^{z_1} \leq a^{z_n}$$

Thus, by the inequality for limits, we see $a^{z_1} \leq \lim a^{z_n} = a^z$. Since z_1 is rational we know a^{z_1} is strictly greater than 1, so a^z is as well. □

We are now ready to prove continuity. The sequence criterion looks suspiciously similar to our definition of a^x, so this sounds like it might be easy. But as with many things in analysis, there are details to be considered: the definition of our function a^x considers *monotone, rational sequences* whereas the definition of continuity requires we consider *arbitrary sequences.* So, we need to bridge this gap. To do so, it will be useful to have a quick lemma, which relies on similar techinques to the above proof

Continuity. Let $x \in \mathbb{R}$ and $x_n \to x$ be arbitrary. Choose an arbitrary sequence $x_n \to x$, we wish to show that a^{x_n} converges to a^x. We proceed by contradiction, assuming it does not. Our goal is to throw away terms of this sequence until we get something nicer (less arbitrary) to work with.

Negating the definition of convergence, there must be some bad ϵ where for every N there is an $n > N$ where $a^{x_{n_k}}$ differs from a^x by more than ϵ. Taking $N = 1, N = 2, N = 3, \ldots$ we can build a subsequence x_{n_k} where every single term is more than ϵ away from a^x. But we can go even further, recalling that every sequence has a monotone subsequence, we can throw away more terms until we have a subsequence $x_{n_{k_\ell}}$ which is *monotone* and has $a^{x_{n_{k_\ell}}}$ not converging a^x.

Phew! That's a lot of subsequences. Its annoying to carry them all around in print, and so we will just rename things: let's call this sequence y_ℓ. Since the original sequence x_n converged to x and this is a subsequence, we know $y_\ell \to x$ as well. And, its much closer to something we might know about (its monotone, and the definition of a^x requires talking about monotone sequences). The only thing left to confront is rationality. We have no idea if the terms y_ℓ are rational (and they need not be). So we are going to do a very cool trick, to replace this with a different sequence.

By the density of rationals we can find an r_ℓ between each pair $y_\ell, y_{\ell+1}$. This defines a sequence of rational numbers, $y_\ell \leq r_\ell \leq y_{\ell+1}$, which converges to x by the squeeze theorem (since $\lim y_\ell = \lim y_{\ell+1} = x$). This was useful: we learned something about the

r sequence using something we know about the y sequence. But since the sequences are *interleaved*

$$y_1 \leq r_1 \leq y_2 \leq r_2 \leq y_3 \leq r_3 \leq y_4 \leq r_4 \leq y_4 \leq r_5 \leq y_6 \cdots$$

We can also think about the y sequence as trapped between the r sequence: $r_{\ell-1} \leq y_\ell < r_\ell$. Because the exponential is monotone (increasing, for $a > 1$) this implies that $a^{r_{\ell-1}} \leq a^{y_\ell} < a^{r_\ell}$. But *now we know about the convergence of the outer two sequences* $\lim a^{r_\ell} = a^x$ by definition, as $r_n \to x$ is a monotone rational sequence. The same holds for $r_{\ell+1}$ as truncating the first term doesn't change convergence. Thus by the squeeze theorem,

$$\lim a^{y_\ell} = a^x$$

But this is a contradiction! As the terms y_ℓ were *specifically chosen* so that a^{y_ℓ} was always further from a^x than ϵ, so it can't eventually be less than ϵ from it. $\square$

We've done it! We've rigorously confirmed all the calculations we've done from precalculus onwards, involving the law of exponents: this really does hold for the continuous function a^x, even at irrational powers! Before moving onwards, its useful to pause for a minute and put our newfound knowledge to the test, proving a couple other facts about the exponential.

Corollary 20.1. *The exponential function a^x is one-to-one on its entire domain.*

Proof. Let $x \neq y$ be real numbers, we want to show $a^x \neq a^y$.
By trichotomy, we know either $x < y$ or $x > y$. In the first case, by strict monotonicity, $a^x < a^y$, and in the second $a^x > a^y$. That is, in both cases $a^x \neq a^y$, so we are done. $\square$

Exercise 20.4. Prove range of the exponential function a^x is all positive real numbers: that is, for any positive y, show there is some x where $a^x = y$.

Hint: can you find some $n \in \mathbb{N}$ where $a^n > y$? If so, can you modify the idea to get an m with $a^{-m} < y$? Once you have these two values, can you apply a theorem about continuity?

Exercise 20.5 (Convexity of exponentials)**.** Prove that exponential functions are convex (Definition 5.8): their secant lines lie above their graphs.

20.3. Logarithms

We've completely put the theory of exponential functions on a rigorous footing, so its time to do the same for logarithms. We define logarithms similarly to what we did for exponentials, by a *functional equation* telling us what they are for.

Definition 20.4 (The Law of Logarithms)**.** A function L satisfies the *law of logarithms* if for every $x, y > 0$,

$$L(xy) = L(x) + L(y)$$

A *logarithm* is a continuous nonconstant solution to the law of logarithms.

Exercise 20.6. Let $L(x)$ be a logarithm and $r \in \mathbb{Q}$ a rational number. Prove directly from the functional equation that $L(x^r) = rL(x)$.

One might be initially concerned: we don't have a nice candidate function on the rationals that we *know* satisfies this, and we just need to extend: so how are we going to prove the existence of such functions? Happily this case actually turns out to be *much less technical* than it looks - because we can put all the hard work we did above to good use!

Theorem 20.3 (Logarithms Exist, and are Inverses to Exponentials)**.** *Let $E(x)$ be an exponential function. Then its inverse function is a logarithm.*

Proof. Let E be an exponential function, and L be its inverse. Because E is continuous, Theorem 16.5 implies that L is also continuous and nonconstant, so we just need to show L satisfies the law of logarithms. Since the range of E is $(0, \infty)$ this means we must check for any $a, b > 0$ that $L(ab) = L(a) + L(b)$.
With a, b in the range fo E we may find x, y with $E(x) = a$ and $E(y) = b$, and (by the definition of L as the inverse) $L(a) = x$ and $L(b) = y$. By the law of exponents for E we see $ab = E(x)E(y) = E(x + y)$, and as L and E are inverses, $L(E(x + y)) = x + y$. Putting this all together gives what we need:

$$L(ab) = L(E(x)E(y)) = L(E(x + y)) = x + y = L(a) + L(b)$$

□

Definition 20.5. The base of a logarithm L is the real number a such that $L(a) = 1$. That is, the log base a is the inverse of a^x.

20.4. ♦ Trigonometric Functions

Like for the exponential and logarithm functions, to propose a rigorous definition of the trigonometric functions, we require them to satisfy the trigonometric identities. To make a specific choice, we take the *angle difference identities*

Definition 20.6 (Angle Identities)**.** A pair of two functions (c, s) are *trigonometric* if they are a continuous nonconstant solution to the *angle identities*

$$s(x-y) = s(x)c(y) - c(x)s(y)$$

$$c(x-y) = c(x)c(y) + s(x)s(y)$$

Definition 20.7 (Other Trigonometric Functions)**.** Given a trigonometric pair s, c we define the tangent function $t(x) = s(x)/c(x)$, as well as the secant $1/c(x)$, cosecant $1/s(x)$ and cotangent $1/t(x)$.

It may seem strange at first: is this really enough to fully nail down trigonometry? It turns out it is: if s, c satisfy these identities then they actually satisfy all the usual trigonometric identities! Its good practice working with functional equations to confirm some of this, which is laid out in the exercises below. I'll start it off, by confirming at least such functions take the right value at zero.

Lemma 20.3 (Values at Zero)**.** *If s, c are trigonometric, then we can calculate their values at 0:*

$$s(0) = 0 \qquad c(0) = 1$$

Proof. Setting $x = y$ in the first immediately gives the first claim

$$s(0) = s(x-x) = s(x)c(x) - c(x)s(x) = 0$$

Evaluating the second functional equation also at $x = y$

$$c(0) = c(x-x) = c(x)c(x) + s(x)s(x) = c(x)^2 + s(x)^2$$

From this we can see that $c(0) \neq 0$, as if it were, we would have $c(x)^2 + s(x)^2 = 0$: since both $c(x)^2$ and $s(x)^2$ are nonnegative this implies each are zero, and so we would have $c(x) = s(x) = 0$ are constant, contradicting the definition. Now, plug in 0 to what we've derived, and use that we know $s(0) = 0$

$$c(0) = c(0)^2 + s(0)^2 = c(0)^2$$

Finally, since $c(0)$ is nonzero we may divide by it, which gives $c(0) = 1$ as claimed. □

An important corollary showed up during the proof here, when we observed that $c(0) = c(x)^2 + s(x)^2$: now that we know $c(0) = 1$, we see that (c, s) satisfy the Pythagorean identity!

Exercise 20.7 (Pythagorean Identity)**.** If s, c are trigonometric, then for every $x \in \mathbb{R}$

$$s(x)^2 + c(x)^2 = 1$$

Continuing this way, we can prove many other trigonometric identities: for instance, the double angle identity (which will be useful to us later)

Exercise 20.8 (Evenness and Oddness)**.** If s, are trigonometric, then s is odd and c is even:

$$s(-x) = -s(x) \qquad c(-x) = c(x)$$

Exercise 20.9 (Angle Sums)**.** If s, c are trigonometric, then for every $x \in \mathbb{R}$

$$s(x + y) = c(x)s(y) + s(x)c(y)$$

$$c(x + y) = c(x)c(y) - s(x)s(y)$$

Exercise 20.10 (Double Angles)**.** If s, c satisfy the angle sum identities, then for any $x \in \mathbb{R}$,

$$s(2x) = 2s(x)c(x)$$

Another useful identity we'll need is the 'Half Angle Identites':

Lemma 20.4. *If s, c are trigonometric functions, then*

$$c(x)^2 = \frac{1 + c(2x)}{2}$$

Proof. Using the angle sum identity we see

$$c(2x) = c(x)c(x) - s(x)s(x) = c(x)^2 - s(x)^2$$

Then applying the pythagorean identity

$$\begin{aligned} c(2x) &= c(x)^2 - s(x)^2 \\ &= c(x)^2 - (1 - c(x)^2) \\ &= 2c(x)^2 - 1 \end{aligned}$$

Re-arranging yields the claimed identity. □

Exercise 20.11. If s, c are trigonometric functions then

$$s(x)^2 = \frac{1 - c(2x)}{2}$$

Just like for exponentials and logs we don't expect this to pick out a *unique* pair of functions, but rather there many be many solutions to the angle identities (corresponding to *different units* we could measure angles with)

Exercise 20.12. Prove that if $s(x), c(x)$ are a trigonometric pair then so are $s(kx), c(kx)$ for any constant $k > 0$.

To prove the existence of trigonometric functions, we'll follow a similar path to exponentials: we'll propose a pair of functions, and confirm they are continuous and nonconstant, and satisfy the trig identities. This is one of the options of the final project, for those interested!

Part V.

Derivatives

- In Chapter 21 we define the derivative as a limit of difference quotients, and investigate the basic properties of differentiable functions
- In Chapter 22 we prove the Mean Value Theorem, a cornerstone result in real analysis and give some applications; from understanding maxima and minima to L'Hospitals rule.
- In Chapter 23 we investigate the differentiability of power series
- In Chapter 24 we use calculus to single out the natural exponential, natural logarithm, and natural units (radians) for the trigonometric functions.

21. Definition & Properties

> Highlights of this Chapter: we prove many foundational theorems about the derivative that one sees in an early calculus course. We see how to take the derivative of scalar multiples, sums, products, quotients and compositions. We also compute - directly from the definition - the derivative of exponential functions. This leads to an important discovery: there is a unique *simplest*, or *natural exponential*, whose derivative is itself. This is the origin of e in Analysis.

Finally - on to some calculus! Here we will define the derivative, and study its properties. This may sound daunting at first, remembering back to the days of calculus when it all seemed so new and advanced. But hopefully, after so much exposure to sequences and series during this course, the rigorous notion of a derivative will feel more just like a nice application of what we've learned, than a whole new theory.

21.1. Difference Quotients

The derivative is defined to capture the *slope of a graph at a point.* Elementary algebra tells us we can compute the slope of a line given two points as rise over run, and so we can compute the slope of a secant line of a function between the points a, t as

$$\frac{f(t) - f(a)}{t - a}$$

The derivative is the *limit* of this, as $t \to a$:

Definition 21.1 (The Derivative)**.** Let f be a function defined on an open interval containing a. Then f is differentiable at a if the following limit of difference quotients exists. In this case, we define the limiting value to be the *derivative of f at a.*

$$f'(a) = Df(a) = \lim_{t \to a} \frac{f(t) - f(a)}{t - a}$$

Exercise 21.1 (Equivalent Formulation)**.** Prove that we may alternatively use the following limit definition to calculate the derivative:

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

Example 21.1. The function $f(x) = x^2$ is differentiable at $x = 2$.

This is a classic problem from calculus 1, whose argument is already pretty much rigorous! We wish to compute the limit

$$\lim_{x \to 2} \frac{x^2 - 4}{x - 2}$$

So, we choose an arbitrary sequence x_n with $x_n \neq 2$ but $x_n \to 2$ and compute

$$\lim \frac{x_n^2 - 4}{x_n - 2} = \lim \frac{(x_n + 2)(x_n - 2)}{x_n - 2} = \lim x_n + 2$$

Where the arithmetic is justified since $x_n \neq 2$ for all n by definition, so everything is defined. But now, as $x_n \to 2$ we can just use the limit laws to see

$$\lim x_n + 2 = 2 + 2 = 4$$

Since x_n was arbitrary, this holds for all such sequences, so the limit exists and equals 4. Because this limit defines the derivative, we have that f is differentiable at 2 and

$$f'(2) = 4$$

Exercise 21.2. Compute the derivative of $f(x) = x^3$ at an arbitrary point $a \in \mathbb{R}$, directly from the definition and show $f'(a) = 3a^2$.

As defined above, the derivative is a *limit* $t \to a$, which depends on values of t both greater than and less than a. But sometimes its useful to have a notion of the derivative that only cares about one sided limits (for instance, when computing the slope at the end of an interval). We give the analogous definition below

Definition 21.2 (One Sided Derivatives)**.** Let f be a function defined at a; then its 1-sided derivatives are defined by the following limits, when they exist

$$D_+ f(a) = \lim_{t \to a^+} \frac{f(t) - f(a)}{t - a}$$
$$D_- f(a) = \lim_{t \to a^-} \frac{f(t) - f(a)}{t - a}$$

This definition, together with our previous work on limits, implies that a function f is differentiable if and only if its two one sided derivatives exist and are equal. This is useful in practice, for instance in showing the non-differentiability of the absolute value:

Exercise 21.3. Show that $f(x) = |x|$ is not differentiable at $x = 0$.

21.2. Derivative as a Function

So far we have been discussing the derivative *at a point* as a number; the result of a limiting process. But we can let this point vary, and produce a *function* taking in x and outputting the derivative at x:

Definition 21.3 (The Function f')**.** Let f be a function, and suppose that the derivative of f exists at each point of a set $D \subset \mathbb{R}$. Then we may define a function $f' : D \to \mathbb{R}$ by

$$f' : x \mapsto f'(x) = \lim_{t \to x} \frac{f(t) - f(x)}{t - x}$$

If f' is continuous, f is called *continuously differentiable* on D.

For example, $f(x) = x^3$ is continuously differentiable on $\mathbb{R}$ since by Exercise 21.2 we see its derivative is the function $x \mapsto 3x^2$, and this is a polynomial: we proved all polynomials are continuous.

Since the derivative of a function yields another function, we can look at *iterating this process* to produce higher derivatives

Definition 21.4 (nth Derivatives)**.** Given a differentiable function f, the second derivative f'' is defined as the derivative of f'. A function is twice differentiable at x if

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

exists. Continuing inductively, we define the n^{th} derivative of a function at a as the derivative of the $n - 1^{st}$ derivative of f at a.

We will use the prime notation for small numbers of derivatives, like $f'(x)$, $f''(x)$ and $f'''(x)$. For higher derivatives it is traditional to denote via the number of derivatives in parentheses: $f^{(2)} = f''$, $f^{(3)} = f'''$ and so on; so $f^{(47)}$ for the 47th derivative of f.

Exercise 21.4 (A Difference Quotient for 2nd Derivative)**.** If f is twice differentiable at a, show that

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

Find a limit depending only on f (not f' or f'') which computes the third derivative

Its useful to have a notation for functions which admit k derivatives, we say a function is C^k if you can differentiate it k times (but not necessarily $k + 1$ times). And, we call a function *smooth* if you can differentiate it n times for any $n \in \mathbb{N}$. The set of smooth functions is denoted C^∞.

21.2.1. Continuity

Before jumping in we prove one small oft-useful result often not mentioned in a calculus class, relating differentiability to continuity.

Theorem 21.1 (Differentiable implies Continuous). *Let f be differentiable at $a \in \mathbb{R}$. Then f is continuous at a.*

Proof. Since f is differentiable at a, we know the limit of the difference quotient is finite

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

We also know that $\lim_{x\to a}(x-a) = 0$\$ So, using the limit theorems we may multiply these together and get what we want. Precisely, let $x_n \to a$ be any sequence with $x_n \neq a$ for all n. Then we have

$$\begin{aligned}
0 &= (0)(f'(a)) \\
&= (\lim x_n - a)\left(\lim \frac{f(x_n)-f(a)}{x_n - a}\right) \\
&= \lim\left((x_n - a)\frac{f(x_n)-f(a)}{x_n - a}\right) \\
&= \lim\left(f(x_n) - f(a)\right)
\end{aligned}$$

Thus $\lim(f(x_n) - a) = 0$ so by the limit theorems we see $\lim f(x_n) = a$. Since x_n was arbitrary with $x_n \neq a$ this holds for any such sequence, we see that f is continuous at a using the sequence definition. □

Remark 21.1. There is a little gap not explicitly spelled out at the end of the proof above, that we should fill in now (to assure ourselves this style of reasoning always works). We just proved that for sequences $x_n \neq a$ the property we want holds, but continuity requires this fact for *all arbitrary sequences*. How do we bridge this gap? Let $y_n \to a$ be an arbitrary sequence: then we split into the subsequences $x_n \neq a$ and the subsequence of all terms $= a$. If either of these is finite, we can just truncate the original sequence at a point past which all terms are of one or the other: each of these has $\lim f(x_n) = f(a)$ so we are done. In the case that both are infinite, we just use that we have separated our sequence into a union of two subsequences, each with the same limit! Thus the overall limit exists.

Thus continuous functions must be differentiable, but what can we say about the derivative itself? If a function is *everywhere differentiable* must the derivative itself be continuous? In fact not, as the following example shows

Example 21.2. While its hard to imagine a function that is differentiable at every point but *not continuously differentiable* such things exist. For example

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Its possible to find a formula for $f'(x)$ when $x \neq 0$, and show that $\lim_{x\to 0} f'(x)$ does not exist (we will do this later). However one can also calculate directly the derivative *at zero*: and find $f'(0) = 0$. This means $\lim_{x\to 0} f'(x) \neq f'(\lim_{x\to 0} x)$ as one side does not exist and the other is zero: thus f' is not continuous at 0.

Exercise 21.5. For $f(x)$ as above in Example 21.2, calculate $f'(0)$ directly using the limit definition. (Perhaps surprisingly, all you need to know about the sine function here is that it is bounded between -1 and 1!)

21.3. Field Operations

Here we prove the 'derivative laws' of Calculus I:

21.3.1. Sums and Multiples

Theorem 21.2 (Differentiating Constant Multiples). *Let f be a function and $c \in \mathbb{R}$. Then if f is differentiable at a point $a \in \mathbb{R}$ so is cf, and*

$$(cf)'(a) = c\,(f'(a))$$

Proof. Let's use the difference quotient with $a + h_n$ to change things up: Let $h_n \to 0$ be arbitrary, and we wish to compute the limit

$$\lim \frac{cf(a+h_n) - cf(a)}{h_n}$$

By the limit laws we can pull out the constant c, and the remainder converges to $f'(a)$, as f is assumed to be differentiable at a.

$$= c \lim \frac{f(a+h_n) - f(a)}{h_n} = cf'(a)$$

Because this is true for all sequences $h_n \to 0$ with $h_n \neq 0$, the limit exists, and equals $cf'(a)$. □

Theorem 21.3 (Differentiating Sums)**.** *Let f, g be functions which are both differentiable at a point $a \in \mathbb{R}$. Then $f + g$ is also differentiable at a, and*

$$(f+g)'(a) = f'(a) + g'(a)$$

Exercise 21.6. Prove the differentiability rule for sums.

21.3.2. Products and Quotients

Theorem 21.4 (Differentiating Products)**.** *Let f, g be functions which are both differentiable at a point $a \in \mathbb{R}$. Then fg is differentiable at a and*

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a)$$

Proof. Let f, g be differentiable at $a \in \mathbb{R}$, and choose an arbitrary sequence $a_n \to a$. Then we wish to compute

$$\lim \frac{f(a_n)g(a_n) - f(a)g(a)}{a_n - a}$$

To the numerator we add $0 = f(a_n)g(a) - f(a_n)g(a)$ and regroup with algebra:

$$\begin{aligned}
&= \lim \frac{f(a_n)g(a_n) - f(a_n)g(a) + f(a_n)g(a) - f(a)g(a)}{a_n - a} \\
&= \lim \frac{f(a_n)g(a_n) - f(a_n)g(a)}{a_n - a} + \frac{f(a_n)g(a) - f(a)g(a)}{a_n - a}
\end{aligned}$$

Using the limit laws, we can take each of these limits individually so long as they exist (which we will show they do). But even more, note that the first term has a common factor of $f(a_n)$ in the numerator that can be factored out, and the second a common factor of $g(a)$. Thus, by the limit laws, we see

$$= (\lim f(a_n))\left(\lim \frac{g(a_n) - g(a)}{a_n - a}\right) + g(a)\left(\frac{f(a_n) - f(a)}{a_n - a}\right)$$

Because f is differentiable at a, its continuous at a, and so we know $\lim f(a_n) = f(a)$. The other two limits above converge to the derivatives $f'(a)$ and $g'(a)$ respectively. Thus, alltogether we find the resulting limit to be

$$f(a)g'(a) + f'(a)g(a)$$

As this was the result for an arbitrary sequence $a_n \to a$ with $a_n \neq a$, it must be the same for all sequences, meaning the limit exists, and

$$(f \cdot g)'(a) = f(a)g'(a) + f'(a)g(a)$$

□

Exercise 21.7. Let f be a function and $a \in \mathbb{R}$ be a point such that $f(a) \neq 0$ and f is differentiable at a. Prove that $1/f$ is also differentiable at a and

$$\left(\frac{1}{f}\right)'(a) = \frac{-f'(a)}{f(a)^2}$$

Theorem 21.5 (Differentiating Quotients)**.** *Let f, g be a functions which are differentiable at a point $a \in \mathbb{R}$ and assume $g(a) \neq 0$. Then the function f/g is also differentiable at a and*

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

Exercise 21.8. Use the Reciprocal Rule and Product Rule to prove the quotient rule.

21.4. Compositions and Inverses

21.4.1. The Chain Rule

Theorem 21.6 (The Chain Rule)**.** *If $g(x)$ is differentiable at $a \in \mathbb{R}$ and $f(x)$ is differentiable at $g(a)$ then the composition $f \circ g$ is differentiable at a, with*

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

Wish this Worked! We are taking the derivative at a, so let $x_n \to a$ wtih $x_n \neq a$ be arbitrary. Then the limit defining $[f(g(a))]'$ is

$$\lim \frac{f(g(x_n)) - f(g(a))}{x_n - a}$$

We multiply the numerator and denominator of this fraction by \$$g(x_n) - g(a)$ and regroup:

$$\begin{aligned}\frac{f(g(x_n))-f(g(a))}{x_n-a} &= \frac{f(g(x_n))-f(g(a))}{x_n-a}\frac{g(x_n)-g(a)}{g(x_n)-g(a)} \\ &= \lim \frac{f(g(x_n))-f(g(a))}{g(x_n)-g(a)}\frac{g(x_n)-g(a)}{x_n-a}\end{aligned}$$

Because g is continuous at a, we know $g(x_n) \to a$, and because f is differentiable at $g(a)$ we recognize the first term here as the *limit defining f' at* $g(a)$! Since the second term is the limit defining the derivative of g, both of these exist by our assumptions, and so by the limit theorems we can compute

$$\begin{aligned}&= \left(\lim \frac{f(g(x_n))-f(g(a))}{g(x_n)-g(a)}\right)\left(\lim \frac{g(x_n)-g(a)}{x_n-a}\right) \\ &= f'(g(a))g'(a)\end{aligned}$$

□

Unfortunately, this proof fails at one crucial step! Wile we do know that $x_n - a \neq 0$ (in the definition of $\lim_{x\to a}$, we only choose sequences $x_n \to a$ with $x_n \neq a$) we do *not* know that the other denominator $g(x_n) - g(a)$ is nonzero.

If this problem could only happen *finitely many times* it would be no trouble - we could just truncate the beginning of our sequence and rest assured we had not affected the value of the limit. But functions - even differentiable functions - can be pretty wild. The function $x^2 \sin(1/x)$ (from Example 21.2) ends up equaling zero infinitely often in any neighborhood of zero! So such things are a real concern.

Happily the fix - while tedious - is straightforward. It's given below.

Exercise 21.9. We define the auxiliary function $d(y)$ as follows:

$$d(y) = \begin{cases} \frac{f(y)-f(g(a))}{y-g(a)} & y \neq g(a) \\ f'(g(a)) & y = g(a) \end{cases}$$

This function equals our problematic difference quotient most of the time, but equals the quantity we *want it to be* when the denominator is zero.

Prove that d is continuous at $g(c)$ and we may use d in place of the difference quotient in our computation: that for all $x \neq a$, the following equality holds:

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

Given this, the original proof is rescued:

Proof. We are taking the derivative at a, so let $x_n \to a$ with $x_n \neq a$ be arbitrary. Then the limit defining $[f(g(a))]'$ is (by the exercise)

$$\lim \frac{f(g(x_n)) - f(g(a))}{x_n - a} = \lim d(g(x_n)) \frac{g(x_n) - g(a)}{x_n - a}$$

Because d is continuous at $g(a)$ and $g(x_n) \to g(a)$ we know $d(g(x_n)) \to d(g(a)) = f'(g(a))$. And, as g is differentiable at a we know the limit of the difference quotient exists. Thus, by the limit laws we can separate them and

$$= (\lim d(g(x_n))) \left(\frac{g(x_n) - g(a)}{x_n - a}\right) = f'(g(a))g'(a)$$

□

21.4.2. Differentiating Inverses

Theorem 21.7 (Differentiating Inverses). *Let f be an invertible function and $a \in \mathbb{R}$ a point where $f(a) = b$. Assume f is differentiable at a with $f'(a) \neq 0$. Then its inverse function f^{-1} is differentiable at b, and*

$$(f^{-1})'(b) = \frac{1}{f'(a)}$$

One may be tempted to prove this using the chain rule, by the following argument: since $f \circ f^{-1}(x) = x$ we differentiate to yield $(f \circ f^{-1}(x))' = 1$ and apply the chain rule to the left hand side, resulting in

$$f'\left(f^{-1}(x)\right)(f^{-1})'(x) = 1$$

Solving for $(f^{-1})'$ and plugging in $x = b$ yields the result. However a more careful review shows doesn't actually do what we think: in applying the chain rule, we've implicitly assumed that f^{-1} is invertible; which is part of what we want to prove! (This proof *does* go through when we already know f^{-1} to be differentiable, but we are unfortunately not often already in possession of that knowledge). Below we give a direct proof of the theorem from the limit definition, fixing this oversight:

Proof. We attempt to compute the limit defining the derivative for f^{-1}: $\lim_{y \to b} \frac{f^{-1}(y) - f^{-1}(b)}{y - b}$. To compute such a limit we choose an arbitrary sequence $y_n \to y$ with $y_n \neq y$ and evaluate

$$\lim \frac{f^{-1}(y_n) - f^{-1}(b)}{y_n - b}$$

By definition $b = f(a)$, and for each n there is a *unique* x_n such that $y_n = f(x_n)$: making these substitutions yields

$$\lim \frac{f^{-1}(f(x_n)) - f^{-1}(f(a))}{f(x_n) - f(a)}$$

The composition $f^{-1} \circ f$ is the identity since they are inverse functions so $f^{-1}(f(x_n)) = x_n$ and $f^{-1}(f(a)) = a$. Making these additional substitutions our limit statement becomes

$$\lim \frac{x_n - a}{f(x_n) - f(a)}$$

By assumption f is differentiable at a and $f'(a) \neq 0$, so we know that

$$f'(a) = \lim \frac{f(x_n) - f(a)}{x_n - a}$$

The limit we are interested in is the reciprocal of this, and as the limit value is nonzero by assumption, the limit laws imply

$$\lim \frac{x_n - a}{f(x_n) - f(a)} = \lim \frac{1}{\frac{f(x_n)-f(a)}{x_n - a}} = \frac{1}{\lim \frac{f(x_n)-f(a)}{x_n-a}} = \frac{1}{f'(a)}$$

Since the sequence y_n was arbitrary, this argument holds for any such sequence. Thus the limit defining $(f^{-1})'(b)$ exists, and $(f^{-1})'(b) = \frac{1}{f'(a)}$. □

Exercise 21.10. Compute the derivative of $y = \sqrt{x}$ using this idea.

21.5. ⧫ The Power Rule

Perhaps the most memorable fact from Calculus I is the power rule, that $(x^n)' = nx^{n-1}$. In this short section, we prove the power level at various levels of generality, starting with natural number exponents and proceeding to arbitrary real exponents.

Exercise 21.11 (Power Rule: Integer Exponents)**.**

We can use the chain rule, and the the functional equation for roots to differentiate $nth roots as well:

Proposition 21.1. *If $R(x) = x^{1/n}$ is the n^{th} root function, then*

$$R'(x) = \frac{1}{n} x^{\frac{1}{n}-1}$$

Proof. The definition of the n^{th} root function is that $R(x)^n = x$. We differentiate this equation with the chain rule, using that n is a natural number exponent:

$$(R(x)^n)' = nR(x)^{n-1}R'(x)$$

The other side was x, whose derivative is 1. Thus,

$$nR(x)^{n-1}R'(x) = 1$$

and, solving for R' yields

$$\begin{aligned} R'(x) &= \frac{1}{nR(x)^{n-1}} = \frac{1}{n(x^{1/n})^{n-1}} = \frac{1}{nx^{\frac{n-1}{n}}} \\ &= \frac{1}{n}x^{\frac{1}{n}-1} \end{aligned}$$

□

Exercise 21.12 (Power Rule: Rational Exponents)**.** Run a similar argument to the n^{th} root case to prove that if $r > 0$ is rational, then x^r is differentiable and $(x^r)' = rx^{r-1}$.

When it comes to arbitrary real exponents one can use their definition as *limits of rational powers*, and work to differentiate such a limit. This is possible but requires an exchange of limits, so needs care. Another method is to use the work we've already put into understanding exponentials and logarithms to help us out!

21.6. Problems

The pasting lemma has a differentiable analog, which shows exactly when gluing two pieces (like the absolute value) is differentiable, and when its not.

Exercise 21.13. Let f, g be two continuous and differentiable functions with $a \in \mathbb{R}$ a point such that $f(a) = g(a)$. Prove that the piecewise function

$$h(x) = \begin{cases} f(x) & x \leq a \\ g(x) & x > a \end{cases}$$

is differentiable at a *if and only if* $f'(a) = g'(a)$. (recall we saw such a function is always continuous at a in **?@exr-pasting-lemma**).

Exercise 21.14 (Differentiable, but The Derivative is Not Continuous)**.** While its hard to imagine a function that is differentiable at every point but *not continuously differentiable* such things exist. For example

$$f(x)=\begin{cases} x^2\sin\left(\frac{1}{x^2}\right) & x\neq 0\\ 0 & x=0\end{cases}$$

Assume for the sake of this problem that $\sin(x)$ is a differentiable function on the entire real line, and prove that $f(x)$ is differentiable at every nonzero point, using the product/chain rules.

At $x=0$ this method fails, but we can compute $f'(0)$ directly using the limit definition. Do this, and show you get zero. (Perhaps surprisingly, all you need to know about the sine function here is that it is bounded between -1 and 1!)

22. Theorems & Applications

Highlights of this Chapter: we study the relationship between the behavior of a function and its derivative, proving several foundational results in the theory of differentiable functions:

- Fermat's Theorem: A differentiable function has derivative zero at an extremum.
- Rolle's Theorem: if a differentiable function is equal at two points, it must have zero derivative at some point in-between.
- The Mean Value Theorem: the average slope of a differentiable function on an interval is realized as the instantaneous slope at some point inside that interval.

The Mean Value theorem is really the star of the show, and we follow it with several important applications

That the derivative (rate of change) should be able to detect local extrema is an old idea, even predating the calculus of Newton and Leibniz. Though certainly realized earlier in certain cases, it is Fermat who is credited with the first general theorem (so, the result below is often called *Fermat's theorem*) We will have more to say about extrema later in the chapter, but this theorem is so useful we prove it first, so it's available for our use throughout.

Theorem 22.1 (Finding Local Extrema (Fermat's Theorem))**.** *Let f be a function with a local extremum at m. Then if f is differentiable at m, we must have $f'(m) = 0$.*

Proof. Without loss of generality we will assume that m is the location of a local minimum (the same argument applies for local maxima, except the inequalities in the numerators reverse). As f is differentiable at m, we know that both the right and left hand limits of the difference quotient exist, and are equal.

First, some preliminaries that apply to both right and left limits. Since we know the limit exists, it's value can by computed via any appropriate sequence $x_n \to m$. Choosing some such sequence we investigate the difference quotient

$$\frac{f(x_n) - f(m)}{x_n - m}$$

Because m is a local minimum, there is some interval (say, of radius ϵ) about m where $f(x) \geq f(m)$. As $x_n \to m$, we know the sequence eventually enters this interval (by the definition of convergence) thus for all sufficiently large n we know

$$f(x_n) - f(m) \geq 0$$

Now, we separate out the limits from above and below, starting with $\lim_{x\to m^-}$. If $x_n \to m$ but $x_n < m$ then we know $x_n - m$ is negative for all n, and so

$$\frac{f(x_n) - f(m)}{x_n - m} = \frac{\text{pos}}{\text{neg}} = \text{neg}$$

Thus, for all n the difference quotient is ≤ 0, and so the limit must be as well! That is,

$$\lim_{x\to m^-} \frac{f(x) - f(m)}{x - m} \leq 0$$

Performing the analogous investigation for the limit from above, we now have a sequence $x_n \to m$ with $x_n \geq m$. This changes the sign of the denominator, so

$$\frac{f(x_n) - f(m)}{x_n - m} = \frac{\text{pos}}{\text{pos}} = \text{pos}$$

Again, if the difference quotient is ≥ 0 for all n, we know the same is true of the limit.

$$\lim_{x\to m^+} \frac{f(x) - f(m)}{x - m} \geq 0$$

But, by our assumption that f is differentiable at m we know both of these must be equal! And if one is ≥ 0 and the other ≤ 0 the only possibility is that $f'(m) = 0$. $\square$

22.1. Mean Values

One of the most important theorems relating f and f' is the *mean value theorem.* This is an excellent example of a theorem that is intuitively obvious (from our experience with reasonable functions) but yet requires careful proof (as we know by know many functions have non-intuitive behavior). Indeed, when I teach calculus I, I often paraphrase the mean value theorem as follows:

> If you drove 60 miles in one hour, then at some point you must have been driving 60 miles per hour

How can we write this mathematically? Say you drove D miles in T hours. If $f(t)$ is your *position* as a function of time*, and you were driving between $t = a$ and $t = b$ (where $b - a = T$), your average speed was

$$\frac{D}{T} = \frac{f(b) - f(a)}{b - a}$$

To then say *at some point you were going D miles per hour implies that there exists some $t^\star$ between a and b where the instantaneous rate of change - the derivative - is equal to this value. This is exactly the Mean Value Theorem:

Theorem 22.2 (The Mean Value Theorem)**.** *If f is a function which is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), then there exists some $x^\star \in (a, b)$ where*

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

Note: The reason we require differentiability only ont he *interior* of the interval is that the two sided limit defining the derivative may not exist at the endpoints, (if for example, the domain of f is only $[a, b]$).

In this section we will prove the mean value theorem. It's simplest to break the proof into two steps: first the special case were $f(a) = f(b)$ (and so we are seeking $f'(x^\star = 0)$), and then apply this to the general version. This special case is often useful in its own right and so has a name: *Rolle's Theorem.*

Theorem 22.3 (Rolle's Theorem)**.** *Let f be continuous on the closed interval $[a, b]$ and differentiable on (a, b). Then if $f(b) = f(a)$, there exists some $x^\star \in (a, b)$ where $f'(x^\star) = 0$.*

Proof. Without loss of generality we may take $f(b) = f(a) = 0$ (if their common value is k, consider instead the function $f(x) - k$, and use the linearity of differentiation to see this yields the same result).

There are two cases: (1) f is constant, and (2) f is not. In the first case, $f'(x) = 0$ for all $x \in (a, b)$ so we may choose any such point. In the second case, since f is continuous, it achieves both a maximum and minimum value on $[a, b]$ by the extreme value theorem. Because f is nonconstant these values are distinct, and so at least one of them must be nonzero. Let $c \in (a, b)$ denote the location of either a (positive) absolute max or (negative) absolute min.

Then, $c \in (a, b)$ and for all $x \in (a, b)$, $f(x) \leq f(c)$ if c is the absolute min, and $f(x) \geq f(c)$ if its the max. In both cases, c satisfies the definition of a *local extremum.* And, as f is differentiable on (a, b) this implies $f'(c) = 0$, as required. $\square$

Now, we return to the main theorem:

Of the Mean Value Theorem. Let f be a function satisfying the hypotheses of the mean value theorem, and L be the secant line connecting $(a, f(a))$ to $(b, f(b))$. Computing this line,

$$L = f(a) + \frac{f(b) - f(a)}{b - a}(x - a)$$

Now define the auxiliary function $g(x) = f(x) - L(x)$. Since $L(a) = f(a)$ and $L(b) = f(b)$, we see that g is zero at both endpoints. Further, since both L and f are continuous on $[a, b]$ and differentiable on (a, b), so is g. Thus, g satisfies the hypotheses of Rolle's theorem, and so there exists some $\star \in (a, b)$ with

$$g(\star) = 0$$

But differentiating g we find

$$\begin{aligned} 0 &= f'(\star) - L'(\star) \\ &= f'(\star) - \frac{f(b) - f(a)}{b - a} \end{aligned}$$

Thus, at $\star$ we have $f'(\star) = \frac{f(b) - f(a)}{b - a}$ as claimed □

Exercise 22.1. Verify the mean value theorem holds for $f(x) = x^2 + x - 1$ on the interval $[4, 7]$.

22.2. MVT and Function Behavior

Proposition 22.1 (Zero Derivative implies Constant)**.** *If f is a differentiable function where $f'(x) = 0$ on an interval I, then f is constant on that interval.*

Proof. Let a, b be any two points in the interval: we will show that $f(a) = f(b)$, so f takes the same value at all points. If $a < b$ we can apply the mean value theorem to this pair, which furnishes a point $c \in (a, b)$ such that

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

But, $f'(c) = 0$ by assumption! Thus $f(b) - f(a) = 0$, so $f(b) = f(a)$. □

Corollary 22.1 (Functions with the Same Derivative)**.** *If f, g are two functions which are differentiable on an interval I and $f' = g'$ on I, then there exists a $C \in \mathbb{R}$ with*

$$f(x) = g(x) + C$$

Proof. Consider the function $h(x) = f(x) - g(x)$. Then by the differentiation laws,

$$h'(x) = f'(x) - g'(x) = 0$$

as we have assumed $f' = g'$. But now **?@prp-derivative-zero-implies-const** implies that h is constant, so $h(x) = C$ for some C. Substituting this in yields

$$f(x) = g(x) + C$$

□

Definition 22.1. Let f be a function. If F is a differentiable function with the same domain such that $F' = f$, we say F is an *antiderivative* of f.

Corollary 22.2 (Antiderivatives differ by a Constant)**.** *Any two antiderivatives of a function f differ by a constant. Thus, the collection of* all possible antiderivatives *is described choosing any particular antiderivative F as*

$$\{F(x) + C \mid C \in \mathbb{R}\}$$

This is the familiar $+C$ from Calculus!

We can use the theory of derivatives to understand when a function is increasing / decreasing and convex/concave, which prove useful in classifying the extrema of functions among other things.

Proposition 22.2 (Monotonicity and the Derivative)**.** *If f is is continuous and differentiable on $[a, b]$, then $f(x)$ is monotone increasing on $[a, b]$ if and only of $f'(x) \geq 0$ for all $x \in [a, b]$.*

As this is an if and only if statement, we prove the two claims separately. First, we assume that $f' \geq 0$ and show f is increasing:

Proof. Let $x < y$ be any two points in the interval $[a, b]$: we wish to show that $f(x) \leq f(y)$. By the Mean Value Theorem, we know there must be some point $\star \in (x, y)$ such that

$$f'(\star) = \frac{f(y) - f(x)}{y - x}$$

But, we've assumed that $f' \geq 0$ on the entire interval, so $f'(\star) \geq 0$. Thus $\frac{f(y)-f(x)}{y-x} \geq 0$, and since $y - x$ is positive, this implies

$$f(y) - f(x) \geq 0$$

That is, $f(y) \geq f(x)$. Note that we can extract even more information here than claimed: if we know that f' is *strictly greater than 0* then following the argument we learn that $f(y) > f(x)$, so f is *strictly monotone increasing.* □

Next, we assume f is increasing and show $f' \geq 0$:

Proof. Assume f is increasing on $[a, b]$, and let $x \in (a, b)$ be arbitrary. Because we have assumed f is differentiable, we know that the right and left limits both exist and are equal, and that either of them equals the value of the derivative. So, we consider the right limit

$$f'(x) = \lim_{t \to x^+} \frac{f(t) - f(x)}{t - x}$$

For any $t > x$ we know $f(t) \geq f(x)$ by the increasing hypothesis, and we know that $t - x > 0$ by definition. Thus, for all such t this difference quotient is nonnegative, and hence remains so in the limit:

$$f'(x) \geq 0$$

□

Exercise 22.2. Prove the analogous statement for negative derivatives: $f'(x) \leq 0$ on $[a, b]$ if and only if $f(x)$ is monotone decreasing on $[a, b]$.

22.3. Classifying Extrema

We can leverage our understanding of function behavior to classify the maxima and minima of a differentiable function. By Fermat's theorem we know that if the derivative exists at such points it must be zero, motivating the following definition:

Definition 22.2 (Critical Points)**.** A critical point of a function f is a point where either (1) f is not differentiable, or (2) f is differentiable, and the derivative is zero.

Note that not all critical points are necessarily local extrema - Fermat's theorem only claims that extrema are critical points - not the converse! There are many examples showing this is not an if and only if:

Example 22.1. The function $f(x) = x^3$ has a critical point at $x = 0$ (as the derivative is zero), but does not have a local extremum there. The function $g(x) = 2x + |x|$ has a critical point at 0 (because it is not differentiable there) but also does not have a local extremum.

If one is only interested in the *absolute max and min* of the function over its entire domain, this already provides a reasonable strategy, which is one of the early highlights of Calculus I.

Theorem 22.4 (Finding Global Extrema). *Let f be a continuous function defined on a closed interval I with finitely many critical points. Then the absolute maximum and minimum value of f are explicitly findable via the following procedure:*

- *Find the value of f at the endpoints of I*
- *Find the value of f at the points of non-differentiability*
- *Find the value of f at the points where $f'(x) = 0$.*

The absolute max of f is the largest of these values, and the the absolute min is the smallest.

Proof. Because I is a closed interval and f is continuous, we are guaranteed by the extreme value theorem that f achieves both a maximum and minimum value. Let these be max, min respectively, realized at points M, m with

$$f(M) = \text{max} \qquad f(m) = \text{min}$$

Without loss of generality, we will consider M (the same argument applies to m).

First, M could be at one of the endpoints of f. If it is not, then M lies in the interior of I, and there is some small interval (a, b) containing M totally contained in the domain I. Since M is the location of the global max, we know for all $x \in I$, $f(x) \leq f(M)$. Thus, for all $x \in (a, b)$, $f(x) \leq f(M)$ so M is the location of a local max.

But if M is the location of a local maximum, if f is differentiable there by Fermat's theorem we know $f'(M) = 0$. Thus, M must be a critical point of f (whether differentiable or not).

Thus, M occurs in the list of critical points and endpoints, which are the points we checked. □

Oftentimes one is concerned with the more fine-grained information of trying to classify *specific* extrema as (local) maxes or mins, however. This requires some additional investigation of the behavior of f near the critical point

Proposition 22.3 (Distinguishing Maxes and Mins). *Let f be a continuously differentiable function on $[a, b]$ and $c \in (a, b)$ be a critical point where $f'(x) < 0$ for $x < c$ and $f'(x) > 0$ if $x > 0$, for all x in some small interval about c.*

Then c is a local minimum of f.

Proof. By the above, we know that $f'(x) < 0$ for $x < c$ implies that f is monotone decreasing for $x < c$: that is, $x < c \implies f(x) \geq f(c)$. Similarly, as $f'(x) > 0$ for $x > 0$, we have that f is increasing, and $c < x \implies f(c) \leq f(x)$.

Thus, for x on either side of c we have $f(x) \geq f(c)$, so c is the location of a local minimum. □

This is even more simply phrased in terms of the *second derivative*, as is common in Calculus I.

Theorem 22.5 (The Second Derivative Test)**.** *Let f be a twice continuously differentiable function on $[a, b]$, and c a critical point. Then if $f''(c) > 0$, the point c is the location of a local minimum, and if $f''(x) > 0$ then c is the location of a local maximum.*

Proof. We consider the case that $f''(c) > 0$, the other is analogous. Since f'' is continuous and positive at c, we know that there exists a small interval $(c - \delta, c + \delta)$ about c where f'' is positive (by **?@prp-continuous-positive-neighborhood**).

Thus, by **?@prp-pos-deriv-increasing**, we know on this interval that f' is an increasing function. Since $f'(c) = 0$, this means that if $x < c$ we have $f'(x) < 0$ and if $x > c$ we have $f'(x) > 0$. That is, f' changes from negative to positive at c, so c is the location of a local minimum by **?@cor-max-min-first-deriv**. □

22.4. Contraction Maps

We can use what we've learned about derivatives and the mean value theorem to also produce a simple test for finding contraction maps.

Proposition 22.4 (Contraction Mappings)**.** *If f is continuously differentiable and $|f'| < 1$ on closed interval then f is a contraction map.*

Proof. Let f have a continuous derivative f' which satisfies $|f'(x)| < 1$ for all x in a closed interval I. Because f' is continuous and $|x|$ is continuous, so is the composition $|f'|$, and thus it achieves a maximum value on I (Theorem 18.2); call this maximum M, and note that $M < 1$ by our assumption.

Now let $x < y \in I$ be arbitrary. By the Mean Value Theorem there is some $c \in [x, y]$ such that

$$f(y) - f(x) = f'(c)(y - x)$$

Taking absolute values and using that $|f'(c)| \leq M$ this implies

$$|f(y) - f(x)| \leq M|y - x|$$

Since x, y were arbitrary this holds for all such pairs, and so the distance between x and$ y decreases by a factor of at least M, which is strictly less than 1. Thus f is a contraction map! □

We know contraction maps to be extremely useful as they have a unique fixed point, and iterating from any starting value produces a sequence which rapidly converges to that fixed point. Using this differential condition its easy to check if a function is a contraction mapping, and thus easy to rigorously establish the existence of certain convergent sequences.

As a good example, we give a re-proof of the convergence of the Babylonian procedure to $\sqrt{2}$

Example 22.2. The function $f(x) = \frac{x+\frac{2}{x}}{2}$ is a contraction map on the interval $[1, 2]$. The fixed point of this map is $\sqrt{2}$∈[1,2]$, thus the sequence $1, f(1), ff(1), fff(1), \ldots$ converges to $\sqrt{2}$.

To prove this, note that if $x = (x + \frac{2}{x})/2$ then $x^2 = 2$ whose only positive solution is $\sqrt{2}$, thus it remains only to check that f is a contraction. Computing its derivative;

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

On the interval $[1, 2]$ the function $1/x^2$ lies in $[1/4, 1]$ and so f' lies in the interval $[-1/2, 1/4]$, and $|f'|$ lies in $[1/4, 1/2]$: thus $|f'|$ is bounded above by $1/2$ and is a contraction map!

22.5. ⋆ Newton's Method

Netwon's method is a recipe for numerically finding zeroes of a function $f(x)$. It works iteratively, by taking one guess for a zero and producing a (hopefully) better one, using the geometry of the derivative and linear approximations. The procedure is simple to derive: given a point a we can calculate the tangent line to f at a

$$\ell(x) = f(a) + f'(a)(x - a)$$

and since this tangent line should be a good approximation of f near a, if a is near the a of f, we can approximate this zero by solving not for $f(x) = 0$ (which is hard, if f is a complicated function) but $\ell(x) = 0$ (which is easy, as ℓ is linear). Doing so gives

$$x =$$

Definition 22.3 (Newton Iteration)**.** Let f be a differentiable function. Then *Newton iteration* is the recursive procedure

$$x \mapsto x - \frac{f(x)}{f'(x)}$$

Starting from some x_0 this defines a recursive sequence $x_{n+1} = x_n - f(x_n)/f'(x_n)$

This is an extremely useful calculational procedure in practice, so long as you can cook up a function that is zero at whatever point you are interested in. To return to a familiar example, to calculate $\sqrt{a}$ one might consider the function $f(x) = x^2 - a$, or to find a solution to $\cos(x) = x$, one may consider $g(x) = x - \cos(x)$.

Exercise 22.3. Show the sequence of approximates from newtons method for $\sqrt{2}$ starting at $x_0 = 2$ is precisely the babylonian sequence.

We already have several proofs this sequence for $\sqrt{2}$ converges, so we know that Newton's method works as expected in at least one instance. But we need a general proof. Below we offer a proof of the special case of a *simple zero*: were $f(x)$ crosses the axes like x rather than running tangent to it like x^2:

Definition 22.4 (Simple Zero)**.** A continuously differentiable function f has a simple zero at c if $f(c) = 0$ but $f'(c) \neq 0$.

Theorem 22.6 (Newton's Method)**.** *Let f be a continuously twice-differentiable function with a simple zero at c. Then there is some δ such that applying newton iteration to any starting point in $I = (c - \delta, c + \delta)$ results in a sequence that converges to c.*

Proof. Our strategy is to show that there is an interval on which the Newton iteration $N(x) = x - f(x)/f'(x)$ is a contraction map.

Since c is a simple zero we know $f'(c) \neq 0$ and without loss of generality we take $f'(c) > 0$. Since f is continuously twice differentiable f' is also continuous, meaning there is some $a > 0$ where f' is positive on the entire interval $(c - a, c + a)$. On this interval we may compute the derivative of the Newton map

$$N'(x) = 1 - \frac{f'f' - f''f}{(f')^2} = \frac{ff''}{(f')^2}$$

Since f, f' and f'' are all continuous and f' is nonzero on this interval, N' is continuous. As $f(c) = 0$ we see $N'(c) = 0$, so using continuity for any $\epsilon > 0$ there is some $b > 0$ where $x \in (c - b, c + b)$ implies $|N'(x)| < \epsilon$.

Thus, choosing any $\epsilon < 1$ and taking $\delta = \min\{a, b\}$ we've found an interval $(c-\delta, c+\delta)$ where the derivative of N is strictly bounded away from 1: thus by Proposition 22.4

N is a contraction map on this interval, and so iterating N from any starting point produces a sequence that converges to the unique fixed point of N (Theorem 11.1). This fixed point $\star$ satisfies

$$N(\star) = \star - \frac{f(\star}{f'(\star)} = \star$$

which after some algebra simplifies to

$$f(\star) = 0$$

Since $f(c) = 0$ and $f'(x)$ is positive on the entire interval by construction, f is increasing and so $f(x) < 0$ for $x < c$ and $f(x) > 0$ for $x > c$. That is, f has a *unique* zero on this interval, so $\star = c$ and our sequence of Newton iterates converges to c as desired. □

The structure of this proof tells us that Netwon's method is actually quite efficient: a contraction map which contracts by $\epsilon < 1$ creates a cauchy sequence that converges *exponentially fast* (like ϵ^n). And in our proof, we see continuity of N' lets us set *any* $\epsilon < 1$ and get an interval about c where convergence is exponential in ϵ. These intervals are nested, and so as x gets closer and closer to c the convergence of Newton's method gets *better and better*: its always exponentially fast but the *base* of the exponential improves as we close in.

Exercise 22.4. Provide an alternative proof of Newton's method when f is convex: if c is a simple zero and $x_0 > c$ show the sequence of Newton iterates is a monotone decreasing sequence which is bounded below, and converges to the c via Monotone Convergence.

22.6. ⧫ L'Hospital's Rule

L'Hospital's rule is a very convenient trick for computing tricky limits in calculus: it tells us that when we are trying to evaluate the limit of a quotient of continuous functions and 'plugging in' yields the undefined expression $0/0$ we can attempt to find the limit's value by differentiating the numerator and denominator, and trying again. Precisely:

Theorem 22.7 (L'Hospital's Rule)**.** *Let f and g be continuous functions on an interval containing a, and assume that both f and g are differentiable on this interval, with the possible exception of the point a.*

Then if $f(a) = g(a) = 0$ and $g'(x) \neq 0$ for all $x \neq a$,

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = L \qquad \textit{implies} \qquad \lim_{x\to a} \frac{f(x)}{g(x)} = L$$

Sketch.

- Show that for any x, we have

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

- For any x, use the MVT to get points c, k such that $f'(c) = \frac{f(x)-f(a)}{x-a}$ and $g'(k) = \frac{g(x)-g(a)}{x-a}$.
- Choose a sequence $x_n \to a$: for each x_n, the above furnishes points c_n, k_n: show these sequences converge to a by squeezing.
- Use this to show that the sequence $s_n = \frac{f'(c_n)}{g'(k_n)}$ converges to L, using our assumption $\lim_{x\to a} \frac{f'}{g'} = L$.
- Conclude that the sequence $\frac{f(x_n)}{g(x_n)} \to L$, and that $\lim_{x\to a} \frac{f(x)}{g(x)} = L$ as claimed.

Hint: Use the $\epsilon - \delta$ definition of a functional limit our assumption $\lim_{x\to a} \frac{f'(x)}{g'(x)} = L$ to help: for any ϵ, theres a δ where $|x-a| < \delta$ implies this quotient is within ϵ of L. Since $c_n, k_n \to a$ can you find an N beyond which $f'(c_n)/g'(k_n)$ is always within ϵ of L? □

Exercise 22.5. Fill in the details of the above proof sketch.

23. Power Series

Highlights of this Chapter: we prove to marvelous results about power series: we show that they are differentiable (and get a formula for their derivative), and we *also* prove a formula about how to approximate functions well with a power series, and in the limit get a *power series representation* of a known function, in terms of its derivatives at a single point.

23.1. Differentiating Term-By-Term

The goal of this section is to prove that power series are differentiable, and that we can differentiate them term by term. That is, we seek to prove

$$\left(\sum_{k\geq 0} a_k x^k\right)' = \sum_{k\geq 0}(a_k x^k)' = \sum_{k\geq 1} k a_k x^{k-1}$$

Because a derivative is defined as a limit, this process of bringing the derivative inside the sum is really an exchange of limits: and we know the tool for that,*Dominated Convergence*! This applies quite generally so we give a *general formulation* and then apply it to power series

Theorem 23.1. *Consider a infinite sum $\sum_k f_k(x)$ of functions which (1) converges on a domain D (2) has each f_k differentiable on D. If there is a sequence M_k such that*

- *M_k with $|f_k'(x)| < M_K$, for all $x \in D$*
- *The sum $\sum M_k$ is convergent.*

Then, the sum $\sum_k f_k'(x)$ is convergent, and

$$\left(\sum_k f_k(x)\right)' = \sum_k f_k'(x)$$

Proof. Recall the limit definition of the derivative (Definition 21.1):

$$\left(\sum_k f_k(x)\right)' = \lim_{y\to x}\frac{\sum_k f_k(y) - \sum_k f_k(x)}{y-x}$$

Writing each sum as the limit of finite sums, we may use the limit theorems (Theorem 7.3,Theorem 7.2) to combine this into a single sum

$$\lim_{y\to x}\frac{\lim_N\sum_{k=0}^N f_k(y)-\lim_N\sum_{k=0}^N f_k(x)}{y-x}=\lim_{y\to x}\lim_N\sum_{k=0}^N\frac{f_k(y)-f_k(x)}{y-x}$$

And now, rewriting the limit of partial sums as an infinite sum, we see

$$\left(\sum_k f_k(x)\right)'=\lim_{y\to x}\sum_k\frac{f_k(y)-f_k(x)}{y-x}$$

If we are justified in switching the limit and the sum via Dominated Convergence, this becomes

$$\sum_k\lim_{y\to x}\frac{f_k(y)-f_k(x)}{y-x}=\sum_k f_k'(x)$$

which is exactly what we want. Thus, all we need to do is justify that the conditions of Dominated Convergence are satisfied, for the terms appearing here. To be precise, this is a limit of functions, and we evaluate these by showing they exist for arbitrary sequences $y_n\to x$ with $y_n\neq x$. Choosing such a sequence and plugging in, we see we are really considering the sequence of terms $\lim\sum_k\frac{f_k(y_n)-f_k(x)}{y_n-x}$.

Dominated convergence tells us we need to bound these terms $\frac{f_k(y_n)-f_k(x)}{y_n-x}$ above by some M_k. As part of the theorem hypothesis, we are given that there exists an M_k bounding the derivative of f_k on D, so we just need to show that these suffice. For any $x\neq y_n$ then $\frac{f_k(y_n)-f_k(x)}{y_n-x}$ measures the slope of the secant line of f_k between x and y_n, so by the Mean Value Theroem (Theorem 22.2) there is some c_n between x and y_n with

$$\left|\frac{f_k(y_n)-f_k(x)}{y_n-x}\right|=|f_k'(c_n)|$$

Since $|f_k'(c)|\leq M_k$ by assumption (as $c\in D$), M_k is a bound for this difference quotient, as required.

Now recall our *other* assumption on the M_k: that $\sum M_k$ converges! This means we can apply dominated convergence bringing the limit inside, and

$$\lim_n\sum_k\frac{f_k(y_n)-f_k(x)}{y_n-x}=\sum_k\lim_n\frac{f_k(y_n)-f_k(x)}{y_n-x}$$

Since f_k is differentiable and $y_n\to x$, by definition this limit converges to the derivative $f_k'(x)$. Thus, the limit of our sums is actually equal to $\sum_k f_k'(x)$. And, as $y_n\to x$

was arbitrary, this holds for *all such sequences*. This means the limit defininig the derivative exists, and putting it all together, that $\left(\sum_k f_k(x)\right)' = \sum_k f_k'(x)$ as required.

□

Now we look to apply this to the specific case of power series $\sum_k a_k x^k$ within their intervals of convergence. The proof is much the same spirit as for *continuity* (where we also used dominated convergence), where we provide bounds M_k by looking at a *larger point* that remains in the interval of convergence. To do so, we need to understand the convergence of the power series of termwise derivatives. This is very similar to a previous homework problem (where you considered termwise *antiderivatives*) so we again leave as an exercise:

Exercise 23.1. Assume that $\sum_k a_k x^k$ has radius of convergence R. Show that $\sum_k k a_k x^{k-1}$ has the same radius of convergence.

Using this, we can put everything together.

Theorem 23.2. *Let $f = \sum_{k\geq 0} a_k x^k$ be a power series with radius of convergence R. Then for $x \in (-R, R)$:*

$$f'(x) = \sum_{k\geq 1} k a_k x^{k-1}$$

Proof. Let $x \in (-R, R)$ be arbitrary. Since x lies strictly within the interval of convergence, we may choose some closed interval $\subset (-R, R)$ containing x. For concreteness we take $[-y, y]$ for some $y < R$, which we will use for the domain D when applying Theorem 23.1.

Getting to work verifying the assumptions of this theorem, our series converges on D (as D is a subset of the radius of convergence). Our individual functions f_k of the sum are just the k^{th} term of the series $f_k(x) = a_k x^k$. These are differentiable (as they are constant multiples of the monomial x^k) with derivatives $f_k' = k a_k x^{k-1}$. The bounds M_k we seek are numbers which are greater in absolute value than this derivative on the domain $D = [-y, y]$. Choose some value $z > y$ within the radius of convergence (say, $a = (R + y)/2$). Then for all $x \in D$ we have $|x| < z$ and so

$$|x|^{k-1} < z^{k-1} = |z^{k-1}| \implies |k a_k x^{k=-1}| < |k a_k z^{k-1}|$$

So we may take $M_k = |k a_k z^{k-1}|$. But since the series of termwise derivatives has the same radius of convergence as the original series, and z is within the radius of convergence, we know $\sum k a_k z^{k-1}$ converges absolutely! That is, $\sum M_k$ converges, and we are done. □

Example 23.1. We know the geometric series converges to $1/(1-x)$ on $(-1,1)$:

$$\sum_{k\geq 0} x^k = \frac{1}{1-x}$$

Differentiating term by term yields a power series for $1/(1-x)^2$:

$$\begin{aligned}
\frac{1}{(1-x)^2} &= \left(\frac{1}{1-x}\right)' \\
&= \left(\sum_{k\geq 0} x^k\right)' \\
&= \sum_{k\geq 0} x^k \\
&= \sum_{k\geq 1} kx^{k-1} \\
&= 1 + 2x + 3x^2 + 4x^3 + \cdots
\end{aligned}$$

The fact that power series are differentiable on their entire radius of convergence puts a strong constraint on which sort of functions can ever be written as the limit of such a series.

Example 23.2. The absolute value $|x|$ is not expressible as a power series.

But this applies much more powerfully than even this: we can show that a power series must be *infinitely differentiable* at each point of its domain!

Corollary 23.1 (Power Series are Smooth)**.** *Proceed by induction on N. We know if $\sum_k a_k x^k$ is a power series it is at least $N = 1$ times differentiable on the same radius of convergence, by our big result above. Now assume it is N times differentiable. Because we can differentiate power series term by term, the N^{th} derivative is* also *a power series, which has the same radius of convergence as the original.*

But now we can apply our main theorem again: this power series is differentiable, with the same radius of convergence! Thus our original function is $N+1$ times differentiable, completing the induction.

23.2. Power Series Representations

While power series are interesting in their own right, our main purpose for them is to *compute functions we already care about.* In this section we use their differentiability to provide tools to do so.

Definition 23.1 (Power Series Representation)**.** A *power series representation* of a function f at a point a is a power series p where $p(x) = f(x)$ on some neighborhood of a.

Proposition 23.1 (Candidate Series Representation)**.** *Let f be a smooth real valued function whose domain contains a neighborhood of 0, and let $p(x) = \sum_{k\geq 0} a_k x^k$ be a power series which equals f on some neighborhood of zero. Then, the power series p is uniquely determined:*

$$p(x) = \sum_{k\geq 0} \frac{f^{(k)}(0)}{k!} x^k$$

Proof. Let $f(x)$ be a smooth function and $p(x) = \sum_{k\geq 0} a_k x^k$ be a power series which equals f on some neighborhood of zero. Then in particular, $p(0) = f(0)$, so

$$\begin{aligned} f(0) &= \lim_N (a_0 + a_1 x + a_2 x^2 + \cdots + a_N x^N) \\ &= \lim_N (a_0 + 0 + 0 + \cdots + 0) \\ &= a_0 \end{aligned}$$

Now, we know the first coefficient of p. How can we get the next? Differentiate!

$$p'(x) = \left(\sum_{k\geq 0} a_k x^k\right)' = \sum_{k\geq 0} (a_k x^k)' = \sum_{k\geq 1} k a_k x^{k-1}$$

Since $f(x) = p(x)$ on some small neighborhood of zero and the derivative is a limit, $f'(0) = p'(0)$. Evaluating this at 0 will give the constant term of the power series p'

$$\begin{aligned} f'(0) &= \lim_N (a_1 + 2a_2 x + 3a_3 x^2 \cdots + N a_N x^{N-1}) \\ &= \lim_N (a_1 + 0 + 0 + \cdots + 0) \\ &= a_1 \end{aligned}$$

Continuing in this way, the second derivative will have a multiple of a_2 as its constant term:

$$p''(x) = 2a_2 + 3 \cdot 2 \cdot a_3 x + 4 \cdot 3 \cdot a_4 x^2 + \cdots$$

And evaluating the equality $f''(x) = p''(x)$ at zero yields

$$f''(0) = 2a_2, \qquad \text{so} \qquad a_2 = \frac{f''(0)}{2}$$

This pattern continues indefinitely, as f is infinitely differentiable. The term a_n arrives in the constant term after n differentiations (as it was originally the coefficient of x^n), at which point it becomes

$$a_n x^n \mapsto n a_n x^{n-1} \mapsto n(n-1)a_n x^{n-2} \mapsto \cdots \mapsto n(n-1)(n-2)\cdots 3\cdot 2\cdot 1 a_n$$

As the constant term of $p^{(n)}$ this means $p^{(n)}(0) = n! a_n$, and so using $f^{(n)}(0) = p^{(n)}(0)$,

$$a_n = \frac{f^{(n)}(0)}{n!}$$

In each case there was no choice to be made, so long as $f = p$ in any small neighborhood of zero, the unique formula for p is

$$p(x) = \sum_{k\geq 0} \frac{f^{(k)}(0)}{k!} x^k$$

□

This candidate series makes it very easy to search for power series representations of known smooth functions: there's only one series to even consider! This series is usually named after Brook Taylor, who gave their general formula in 1715.

Definition 23.2 (Taylor Series)**.** For any smooth function $f(x)$ we define the *Taylor Polynomial* (centered at 0) of degree N to be

$$p_N(x) = \sum_{0\leq k\leq N} \frac{f^{(k)}(0)}{k!} x^k$$

In the limit as $N \to \infty$, this defines the *Taylor Series* $p(x)$ for f.

We've seen for example, that the geometric series $\sum_{k\geq 0} x^k$ is a power series representation of the function $1/(1-x)$ at zero: it actually converges on the entire interval $(-1, 1)$. There are many reasons one may be interested in finding a power series representation of a function - and the above theorem tells us that if we were to search for one, there is a single natural candidate. If there is *any power series representation*, its this one!

So the next natural step is to study this representation: does it actually converge to $f(x)$?

23.2.1. Taylor's Error Formula

To *prove* that our series actually do what we want, we are going to need some tools relating a functions derivatives to its values. Rolle's Theorem / the Mean Value Theorem does this for the first derivative, and so we present a generalization here the *polynomial mean value theorem*, which does so for n^{th} derivatives.

Proposition 23.2 (A Generalized Rolle's Theorem). *Let f be a function which is $n+1$ times differentiable on the interior of an interval $[a,b]$. Assume that $f(a)=f(b)=0$, and further that the first n derivatives at a are zero:*

$$f(a)=f'(a)=f''(a)=\cdots=f^{(n)}(a)=0$$

Then, there exists some $c\in(a,b)$ where $f^{(n+1)}(c)=0$.

Proof. Because f is continuous and differentiable, and $f(a)=f(b)$, the original Rolle's Theorem implies that there exists some $c\in(a,b)$ where $f'(c_1)=0$. But now, we know that $f'(a)=f'(c_1)=0$, so we can apply Rolle's theorem to f' on $[a,c_1] to get a point $c_2\in(a,c_1)$ with $f''(c_2)=0$.

Continuing in this way, we get a $c_3\in(a,c_2)$ with $f^{(3)}(c)=0$, all the way up to to a $c_n\in(a,c_{n-1})$ where $f^n(c_n)=0$. This leaves one more application of Rolle's theorem possible, as we assumed $f^{(n)}(a)=0$, so we get a $c\in(a,c_n)$ with $f^{(n+1)}(c)=0$ as claimed. □

Proposition 23.3 (A Polynomial Mean Value Theorem). *Let $f(x)$ be an $n+1$-times differentiable function on $[a,b]$ and $h(x)$ a polynomial which shares the first n derivatives with f at zero:*

$$f(a)=h(a),\quad f'(a)=h'(a),\ldots,\quad f^{(n)}(a)=p^{(n)}(a)$$

Then, if additionally $f(b)=h(b)$, there must exist some point $c\in(a,b)$ where

$$f^{(n+1)}(c)=h^{(n+1)}(c)$$

Proof. Define the function $g(x)=f(x)-h(x)$. Then all the first n derivatives of g at $x=a$ are zero (as f and h had the same derivatives), and furthermore $g(b)=0$ as well, since $f(b)=h(b)$. This means we can apply the generalized Rolle's theorem and find a $c\in(a,b)$ with

$$g^{(n+1)}(c)=0$$

That is, $f^{(n+1)}(c)=h^{(n+1)}(c)$. □

Theorem 23.3 (Taylor Remainder)**.** *Let $f(x)$ be an $n+1$-times differentiable function, and $p_n(x)$ the degree n Taylor polynomial $p(x) = \sum_{0\le k\le n} \frac{f^{(k)}(0)}{k!}x^k$.*

Then for any fixed $b \in \mathbb{R}$, we have

$$f(b) = p_n(b) + \frac{f^{(n+1)}(c)}{(n+1)!}b^{n+1}$$

For some $c \in [0, b]$.

Proof. Fix a point b, and consider the functions $f(x)$ and $p_n(x)$ on the interval $[0, b]$. These share their first n derivatives at a, but $f(b) \neq p_n(b)$: in fact, it is precisely this error we are trying to quantify.

We need to modify p_n in some way without affecting its first n derivatives at zero. One natural way is to add a multiple of x^{n+1}, so define

$$q(x) = p_n(x) + \lambda x^{n+1}$$

for some $\lambda \in \mathbb{R}$, where we choose λ so that $f(b) = q(b)$. Because we ensured $q^{(k)}(0) = f^{(k)}(0)$ for $k \le n$, we can now apply the polynomial mean value theorem to these two functions, and get some $c \in (0, b)$ where

$$f^{(n+1)}(c) = q^{(n+1)}(c)$$

Since p_n is degree n its $n+1^{st}$ derivative is zero, and

$$q^{(n+1)}(x) = 0 + \left(\lambda x^{n+1}\right)^{(n+1)} = (n+1)!\lambda$$

Putting these last two observations together yields

$$f^{(n+1)}(c) = (n+1)!\lambda \implies \lambda = \frac{f^{(n+1)}(c)}{(n+1)!}$$

As $q(b) = f(b)$ by construction, this in turn gives what we were after:

$$f(b) = p_n(b) + \frac{f^{(n+1)}(c)}{(n+1)!}b^{n+1}$$

□

23.2.2. Series Centered at $a \in \mathbb{R}$

All of our discussion (and indeed, everything we will need about power series for our course) dealt with defining a power series based on derivative information at zero. But of course, this was an arbitrary choice: one could do exactly the same thing based at any point $a \in \mathbb{R}$.

Theorem 23.4. *Let f be a smooth function, defined in a neighborhood of $a \in \mathbb{R}$. Then there is a unique power series which has all the same derivatives as f at a:*

$$p(x) = \sum_{k \geq 0} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

And, for any N the error between f and the N^{th} partial sum is quantified as

$$f(x) - p_N(x) = \frac{f^{(N+1)}(\xi)}{(N+1)!}(x-a)^{N+1}$$

For some $c \in [a, x]$.

Exercise 23.2. Prove this.

24. Elementary Functions

In this section we look at how to find derivatives of functions which are defined not explicitly, but by the *functional equations* defining exponentials, logarithms and trigonometric functions.

24.1. Exponentials

Proposition 24.1. *Let $E(x)$ be an exponential function. Then E is differentiable on the entire real line, and*

$$E'(x) = E'(0)E(x)$$

First we show that this formula holds so long as E is actually differentiable at zero. Thus, differentiability at a single point is enough to ensure differentiability everywhere *and* fully determine the formula!

Proof. Let $x \in \mathbb{R}$, and $h_n \to 0$. Then we compute $E'(x)$ by the following limit:

$$E'(x) = \lim \frac{E(x+h_n) - E(x)}{h_n}$$

Using the property of exponentials and the limit laws, we can factor an $E(x)$ out of the entire numerator:

$$= \lim \frac{E(x)E(h_n) - E(x)}{h_n} = E(x) \lim \frac{E(h_n) - 1}{h_n}$$

But, $E(0) = 1$ so the limit here is actually the *derivative of E at zero$!

$$E'(x) = E(x)E'(0)$$

□

Next, we tackle the slightly more subtle problem of showing that E is in fact differentiable at zero. This is tricky because *all we have assumed* is that E is continuous and satisfies the law of exponents: how are we going to pull differentiability out of this? One trick is two parts (1) show the right and left hand limits defining the derivative exist, and (2) show they're equal.

In fact, $E'(0)$ log *of a (Cite*

Proof. Because E is convex (Exercise 20.5) so the difference quotient is monotone increasing and so the limit $\lim_{x\to 0^-}$ exists (as a sup) and $\lim_{x\to 0^+}$ exists (as an inf). Now that we know each of these limits exist, let's show they are equal using the definition: To compute the lower limit, we can choose any sequence approaching 0 from below: let h_n be a positive sequence with $h_n \to 0$, then $-h_n$ will do:

$$\lim_{h\to 0^-} \frac{E(h)-1}{h} = \lim \frac{E(-h_n)-1}{-h_n}$$

As $E(-h_n) = 1/E(h_n)$, we compute

$$\begin{aligned}
\lim \frac{E(-h_n)-1}{-h_n} &= \lim \frac{\frac{1}{E(h_n)}-1}{-h_n} \\
&= \lim \frac{1-E(h_n)}{-h_n}\frac{1}{Eh_n} \\
&= \lim \frac{E(h_n)-1}{h_n}\frac{1}{E(h_n)}
\end{aligned}$$

But, since E is continuous (by definition) and $E(0) = 1$ the limit theorems imply

$$\lim \frac{1}{E(h_n)} = \frac{1}{\lim E(h_n)} = \frac{1}{E(\lim h_n)} = \frac{1}{E(0)} = 1$$

Thus,

$$\begin{aligned}
&\lim \left(\frac{E(h_n)-1}{h_n}\frac{1}{E(h_n)}\right) \\
&= \left(\lim \frac{E(h_n)-1}{h_n}\right)\left(\lim \frac{1}{E(h_n)}\right) \\
&= \lim \frac{E(h_n)-1}{h_n}
\end{aligned}$$

But this last limit evaluates exactly to the *limit from above* since $h_n > 0$ and $h_n \to 0$. Stringing all of this together, we finally see

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

As both one sided limits exist and are equal the entire limit exists: E is differentiable at 0. □

When studying the functional equations for logs and exponentials we saw there is not one solution but a whole family of them. While the functional equation itself gave no preference to any exponential over any other, the derivative selects an obvious natural candidate:

Definition 24.1. We write $\exp(x)$ for the exponential function which has $\exp'(0) = 1$. This exponential satisfies the simple differential identity

$$\exp'(x) = \exp(x)$$

Note that by the chain rule we know such a thing exists so long as *any* exponential exists. If $E(x)$ is any exponential then $E(x/E'(0))$ has derivative 1 at $x = 0$!

24.1.1. A Series Representation

To work with the natural exponential efficiently, we need to find a *formula* that lets us compute it. And this is exactly what power series are good at! However, the theory of power series is a little tricky, as we saw in the last chapter. Not every function has a power series representation, but *if* a function does, there's only one possibility:

Proposition 24.2. If *the natural exponential has a power series representation, then it is*

$$p(x) = \sum_{k\geq 0} \frac{x^k}{k!}$$

Proof. We know the only candidate series for a function $f(x)$ is $\sum_{k\geq 0} \frac{f^{(k)}(0)}{k!}x^k$, so for exp this is

$$p(x) = \sum_{k\geq 0} \frac{\exp^{(k)}(0)}{k!}x^k$$

However, we know that $\exp' = \exp$ and so inductively $\exp^{(k)} = \exp$, and so

$$\exp^{(k)}(0) = \exp(0) = 1$$

Thus

$$p(x) = \sum_{k\geq 0} \frac{1}{k!}x^k$$

□

Since p is a power series, this really means that *the limit of its partial sums equals* $\exp(x)$, or

$$\forall x \in \mathbb{R} \quad \exp(x) = \lim_{N} p_N(x)$$

For any finite partial sum p_N, we know that it is not exactly equal to $\exp(x)$ (as this finite sum is just a polynomial!). Thus there must be some error term $R_N = \exp - p_N$, or

$$\exp(x) = p_N(x) + R_N(x)$$

This is helpful, as we know from the previous chapter how to calculate such an error, using the Taylor Error Formula: for each fixed $x \in \mathbb{R}$ and each fixed $N \in \mathbb{N}$, there is some point $c_N \in [0, x]$ such that

$$R_N(x) = \frac{\exp^{(N+1)}(c_N)}{(N+1)!} x^{N+1}$$

And, to show the power series becomes the natural exponential in the limit, we just need to show this error tends to zero!

Proposition 24.3. *As $N \to \infty$, for any $x \in \mathbb{R}$ the Taylor error term for the exponential goes to zero:*

$$R_N(x) \to 0$$

Proof. Fix some $x \in \mathbb{R}$. Then for an arbitrary N, we know

$$R_N(x) = \frac{\exp^{(N+1)}(c_N)}{(N+1)!} x^{N+1}$$

where $c_N \in [0, x]$ is some number that we don't have much control of (as it came from an existence proof: Rolle's theorem in our derivation of the Taylor error). Because we don't know c_N explicitly, its hard to directly compute the limit and so instead we use the squeeze theorem:

We know that exp is an increasing function: thus, the fact that $0 \le c_N \le x$ implies that $1 = \exp(0) \le \exp(c_N) \le \exp(x)$, and multiplying this inequality through by $x^{N+1}(N+1)!$ yields the inequality

$$\frac{x^{N+1}}{(N+1)!} \le R_N(x) = \exp(c_N)\frac{x^{N+1}}{(N+1)!} \le \exp(x)\frac{x^{N+1}}{(N+1)!}$$

(Here I have assumed that $x \ge 0$: if $x < 0$ then the inequalities reverse for even values of N as x^{N+1} is negative and we are multiplying through by a negative number. But this does not affect the fact that the error term $R_N(x)$ is still sandwiched between the two.)

So now our problem reduces to showing that the upper and lower bounds converge to zero. Since $\exp(x)$ is a constant (remember, N is our variable here as we take the limit), a limit of both the upper and lower bounds comes down to just finding the limit

$$\lim_N \frac{x^{N+1}}{(N+1)}$$

But this is just the $N+1$st term of the power series $p(x) = \sum_{n\geq 0} x^n/n!$ we studied above! And since this power series converges, we know that as $n \to \infty$ its terms must go to zero (the divergence test). Thus

$$\lim_N \frac{x^{N+1}}{(N+1)} = 0 \qquad \lim_N \exp(x)\frac{x^{N+1}}{(N+1)} = 0$$

and so by the squeeze theorem, $R_N(x)$ converges and

$$\lim_N R_N(x) = 0$$

$\square$

Now we have all the components together at last: we know that exp exists, we have a candidate power series representation, that candidate converges, and the error between it and the exponential goes to zero!

Theorem 24.1. *The natural exponential is given by the following power series*

$$\exp(x) = \sum_{k\geq 0} \frac{x^k}{k!}$$

Proof. Fix an arbitrary $x \in \mathbb{R}$. Then for any N we can write

$$\exp(x) = p_N(x) + R_N(x)$$

For p_N the partial sum of $p(x) = \sum_{k\geq 0} x^k/k!$ and $R_N(x)$ the error. Since we have proven both p_N and R_N converge, we can take the limit of both sides using the limit theorems (and, as $\exp(x)$ is constant in N, clearly $\lim_N \exp(x) = \exp(x)$):

$$\begin{aligned}
\exp(x) &= \lim_N (p_N(x) + R_N(x)) \\
&= \lim_N p_N(x) + \lim_N R_N(x) \\
&= p(x) + 0 \\
&= \sum_{k\geq 0} \frac{x^k}{k!}
\end{aligned}$$

$\square$

Its incredible in and of itself to have such a simple, explicit formula for the natural exponential. But this is just the beginning: this series actually gives us a means to express *all* exponentials:

Theorem 24.2. *Let $E(x)$ be an arbitrary exponential function. Then E has a power series representation on all of $\mathbb{R}$ which can be expressed for some real nonzero c as*

$$E(x) = \sum_{n\geq 0} \frac{c^n}{n!}x^n$$

Proof. Because E is an exponential we know E is differentiable, and that $E'(x) = E'(0)E(x)$ for all x. Note that $E'(0)$ is nonzero; else we would have $E'(x) = 0$ constantly, and so $E(x)$ would be constant. Set $c = E'(0)$.

Now, inductively take derivatives at zero:

$$E'(0) = c \qquad E''(0) = c^2 \qquad E^{(n)}(0) = c^n$$

Thus, *if* E has a power series representation it must be

$$\sum_{n\geq 0} \frac{c^n}{n!}x^n = \sum_{n\geq 0} \frac{1}{n!}(cx)^n$$

This is just the series for exp evaluated at cx: since exp exists and is an exponential, so is this function (as its defined just by a substitution). So there is such an exponential. □

24.1.2. The Number e

Recalling our work with irrational exponents, we know that exponentials are *powers*: if E is an exponential with $E(1) = a$, then we may write $E(x) = a^x$ for any $x \in \mathbb{R}$ (defined as a limit of rational exponents). So, our special exponential exp comes with a special number as its base.

Definition 24.2. We denote by the letter e the base of the exponential $\exp(x)$: that is, $e = \exp(1)$, and

$$\exp(x) = e^x$$

What is this natural base? We can estimate its value using the power series representation for exp, and the Taylor error formula.

Proposition 24.4. *The base of the natural exponential is between 2 and 3.*

Proof. The series defining e is all positive terms, so we see that e is greater than any partial sum. Thus

$$2 = 1 + 1 = \frac{1}{0!} + \frac{1}{1!} < \sum_{k\geq 0} \frac{1}{k!} = e$$

so we have the lower bound. To get the upper bound, we need to come up wtih a computable upper bound for our series. This turns out to be not that difficult: as the factorial grows so quickly, we can produce many upper bounds by just fining something that grows slower than the reciprocal and summing up their reciprocals. For instance, when $k \geq 2$

$$k(k-1) \leq k!$$

and so,

$$e = \sum_{k\geq 0} \frac{1}{k!} = 1 + 1 + \sum_{k\geq 2} \frac{1}{k!} \leq 1 + 1 + \sum_{k\geq 2} \frac{1}{k(k-1)}$$

But this upper bound now is our favorite telescoping series! After a rewrite with partial fractions, we directly see that it sums to 1. Plugging this in,

$$e < 1 + 1 + 1 = 3$$

□

How can we get a better estimate? Since we do have a convergent infinite series just sitting here defining e for us, the answer seems obvious - why don't we just sum up more and more terms of the series? And of course - that is *part* of the correct strategy, but it's missing one key piece. If you add up the first 10 terms of the series and you get some number, how can you know how accurate this is?

Just because the first two digits are 2.7, who is to say that after adding a million more terms (all of which are positive) it won't eventually become 2.8? To give us any confidence in the value of e we need a way of measuring how far off any of our partial sums could be.

Our usual approach is to try and produce sequences of upper and lower estimates: nested intervals of error bars to help us out. But here we have only one sequence (and producing even a single upper bound above was a bit of work!) so we need to look elsewhere. It turns out, the correct tool for the job is the Taylor Error formula once more!

Proposition 24.5. *Adding up the first N terms of the series expansion of e results in a an estimate of the true value accurate to within* $3/(N+1)!$.

Proof. The number e is defined as $\exp(1)$, and so using $x = 1$ we are just looking at the old equation

$$\exp(1) = p_N(1) + R_N(1)$$

Where $R_N(1) = \exp(c_N)\frac{1^{N+1}}{(N+1)!}$ for $c_N \in [0,1]$. Since exp is increasing, we can bound $\exp(c_N)$ below by $\exp(0) = 1$ and above by $\exp(1) = e$, and e above by 3: thus

$$\frac{1}{(N+1)!} \leq R_N(x) \leq \frac{3}{(N+1)!}$$

And so, the difference $|e - p_N(1)| = |R_N(1)|$ is bounded above by the upper bound $3/(N+1)!$ □

This gives us a readily computable, explicit estimate. Precisely adding up to the $N =$ 5th term of the series yields

$$1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} \approx 2.71666\ldots$$

with the total error between this and e is less than $\frac{3}{6!} = \frac{1}{240} = 0.0041666\ldots.$ Thus we can be confident that the first digit after the decimal is a 7, as $2.7176 - 0.0041 = 2.7135 \leq e \leq 2.7176 + 0.0041 = 2.7217$.

Adding up five more terms, to $N = 10$ gives

$$1 + 1 + \frac{1}{2} + \frac{1}{3!} + \cdots + \frac{1}{10!} = 2.71828180114638\ldots$$

now with a maximal error of $3/11! = 0.000000075156\ldots.$ This means we are now absolutely confident in the first six digits:

$$e \approx 2.718281$$

Pretty good, for only having to add eleven fractions together! Thats the sort of calculation one could even manage by hand.

24.2. Logarithms

Having done all this work for exponentials, we can immediately confirm that logarithms are also differentiable, and give a simple formula.

Proposition 24.6. *Let L be a logarithm function, then L is differentiable and*

$$L'(x) = \frac{L'(1)}{x}$$

Proof. Let L be a logarithm, with inverse the exponential $E(x)$. We know that E is differentiable and $E'(x) \neq 0$ as its a constant multiple of the everywhere-positive $E(x)$ itself. Thus by Theorem 21.7 the function L is also differentiable. Choosing $b > 0$ and setting $L(b) = a$ gives

$$L'(b) = \frac{1}{E'(a)} = \frac{1}{E'(0)E(a)} = \frac{1}{E'(0)b}$$

Where the last equality follows as $L(b) = a$ implies $E(a) = b$

Now we apply the theorem on differentiability of inverses one more time to remove the mention of E in the final answer, and express everything in terms of the logarithm itself. Since $E(0) = 1$ and $L(1) = 0$, we have $L'(1) = \frac{1}{E'(0)}$ and substituting this in gives the claimed form. $\square$

Thus every logarithm has as its derivative some multiple of $1/x$. While the functional equation doesn't distinguish between these different logarithms, calculus finds one of them most natural: when this constant is equal to 1!

Definition 24.3 (The Natural Log)**.** We write log for the logarithm function which has $\log'(1) = 1$. This logarithm satisfies the simple differential identity

$$\log'(x) = \frac{1}{x}$$

Furthermore the two notions of "naturalness" picking out a logarithm and an exponential are compatible with one another!

Corollary 24.1. *The natural exponential and natural log are inverses of one another.*

We will make much use of this pair of special functions, and their exeedingly simple differentiation rules. As a first application, we give a re-proof of the power rule avoiding difficult limiting arguments:

Theorem 24.3 (The General Power Rule)**.** *If $a \in \mathbb{R}$ and $f(x) = x^a$. Then f is differentiable for all $x > 0$, and*

$$(x^a)' = ax^{a-1}$$

Proof. Let exp be the natural exponential, and log be the natural log. Then $\exp(\log(x)) = x$, and so $\exp(\log(x^a)) = x^a$. Using the property of logarithms and powers (**?@cor-log-exponent**) this simplifies

$$x^n = \exp(\log(x^a)) = \exp(a\log(x))$$

By the chain rule,

$$\begin{aligned}
[\exp(a\log(x))]' &= \exp(a\log(x))\,[a\log(x)]' \\
&= \exp(a\log(x))a\log'(x) \\
&= \exp(a\log(x))a\frac{1}{x}
\end{aligned}$$

But, recalling that $\exp(a\log(x)) = \exp(\log(x^a)) = x^a$ this simplifies to

$$= x^a a\frac{1}{x} = ax^{a-1}$$

□

Unfortunately our newfound tool does not apply so well to giving a formula for the logarithm: power series are always defined on some symmetric interval $(-r, r)$ about 0, but the domain of the logarithm is $(0, \infty)$. Thus there is no simple power series that will equal $\log(x)$!. We will come up with formulas to compute the logarithm later on, first as an integral; and then as a series (that converges only for some values of x).

24.3. Trigonometric Functions

Recall our definition of trigonometric functions: two functions $s(x), c(x)$ are *trigonometric* if they satisfy the angle difference identities

$$s(x-y) = s(x)c(y) - c(x)s(y) \qquad c(x-y) = c(x)c(y) + s(x)s(y)$$

In one option for the final project in this class, you prove that such functions exist by taking a proposed power series, and showing directly that it satisfies these identities. Our goal in this section is to build up the work that leads to *proposing that power series* via a sequence of exercises

Exercise 24.1. Assume that a pair $s(x), c(x)$ of trigonometric functions are differentiable at zero (so $s'(0)$, $c'(0)$ exist). Use the functional equations defining them to show that they are differentiable at every $x \in \mathbb{R}$.

Hint: Write out the difference quotients, as a limit $h \to 0$ for $f(x+h) - f(x)$. Additionally, we know s, c are continuous by definition, and have previously shown $c(0) = 1$, $s(0) = 0$.

We will not *prove* that s and c are differentiable at zero here (though that can be done, much like we did for the exponential above!). Since our goal is simply to *propose* a power series,

Exercise 24.2. Prove that *if* $s(x), c(x)$ are differentiable at zero, then $c'(0) = 0$ and $s'(0) \neq 0$.
Then deduce the following differentiation laws:

$$s'(x) = \lambda c(x) \qquad c'(x) = -\lambda s(x)$$

for λ some nonzero constant.

Hint: recall the definition that s, c are nonconstant and continuous. Also recall Fermat's Theorem.

This singles out an obvious 'best' set of trigonometric functions, setting the arbitrary constant parameter that appears equal to one!

Definition 24.4 (The Natural Trigonometric Functions)**.** The natural trigonometric functions are the pair $\sin(x), \cos(x)$ with

$$\sin'(x) = \cos(x) \qquad \cos'(x) = -\sin(x)$$

Exercise 24.3. Show that *if the natural trigonometric functions have a series representation* they must be

$$\sin(x) = \sum_{n\geq 0} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

$$\cos(x) = \sum_{n\geq 0} \frac{(-1)^n}{(2n)1} x^{2n}$$

As part of the optional final project, you show these functions do satisfy the angle identities, and are periodic. This defines a natural number associated to trigonometry, much as we had the natural number e associated to exponentiation.

Definition 24.5 (π)**.** The natural trigonometric functions are periodic, and we define their half period as π.

That is, π is the smallest positive number such that $\sin(x + 2\pi) = \sin(x)$ and $\cos(x + 2\pi) = \cos(x)$ for all $x \in \mathbb{R}$.

Exercise 24.4. Prove that $\pi/2$ is the smallest positive zero of $\cos(x)$.

Exercise 24.5 (Approximating π with Newton's Method). The first zero of $\cos(x)$ is $\pi/2$, so one might hope to use Newton's method to produce an approximation for π. Show the sequence

$$x_{n+1} = N(x_n) = x_n + \frac{\cos(x_n)}{\sin(x_n)}$$

starting at $x_0 = 1$ converges to $\pi/2$, and use a calculator to compute the first couple terms.

This of course is not very satisfying as we had to use a calculator to find values of sin and cos! But we know enough to approximate these values with a series expansion.

Exercise 24.6. How many terms of the series expansions of sin, cos are needed to evaluate at $x = 1$ to within 0.0001? Use this many terms of the series expansion to approximate the terms appearing in the first two iterations of Newtons method

$$1, N(1), N(N(1))$$

What is your approximate value for π resulting from this?

24.4. ⋆ Existence of Exponentials: an Alternative Proof

Our argument above used that we had *previously confirmed* the existence of exponential functions, together with the Taylor Error formula to find a series representation. But as often happens, the amount of new technology we have developed along the way gives a new self-contained means of both proving the existence of exponentials, and constructing their series in one stroke! We give this alternative argument here.

The idea essentially turns some of our previous reasoning on its head: we start by looking at solutions to the equation $y' = y$ and (1) show they satisfy the law of exponents, then (2) construct an explicit solution as a power series. First, a helpful lemma about this differential equation:

Proposition 24.7. *Let f, g be two solutions to the differential equation $y' = y$. Then they are constant multiples of one another.*

Proof. Consider the function $h(x) = \frac{f(x)}{g(x)}$. Differentiating with the quotient rule,

$$h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \tag{24.1}$$
$$= \frac{f(x)g(x) - f(x)g(x)}{g(x)^2} \tag{24.2}$$
$$= \frac{0}{g(x)^2} \tag{24.3}$$
$$= 0 \tag{24.4}$$

Thus $h'(x) = 0$ for all x, which implies $h = f/g$ is a constant function, and g is a constant multiple of f as claimed. □

Now we're ready for the main theorem:

Theorem 24.4. *Let g be any differentiable function which solves $g' = g$ and has $g(0) = 1$. Then g is an exponential.*

Proof. Let $g : \mathbb{R} \to \mathbb{R}$ solve $Y' = Y$ and satisfy $g(0) = 1$. We wish to show that $g(x + y) = g(x)g(y)$ for all $x, y \in \mathbb{R}$.

So, fix an arbitrary y, and consider each of these separately, defining functions $L(x) = g(x + y)$ and $R(x) = g(x)g(y)$.
Differentiating,

$$\begin{aligned} L'(x) &= (g(x+y))' \\ &= g(x+y)(x+y)' \\ &= g(x+y) \\ &= L(x) \end{aligned}$$

$$\begin{aligned} R'(x) &= (g(x)g(y))' \\ &= (g(x))'g(y) \\ &= g(x)g(y) \\ &= R(x) \end{aligned}$$

Thus, both L and R satisfy the differential equation $Y' = Y$. Our previous proposition implies they are constant multiples of one another,

$$\frac{L(x)}{R(x)} = k \qquad \forall x \in \mathbb{R}$$

To find this constant we evaluate at $x = 0$ where (using $g(0) = 1$) we have

$$L(0) = g(0 + y) = g(y)$$
$$R(0) = g(0)g(y) = g(y)$$

They are equal at 0 so the constant is 1:

$$\frac{L(x)}{R(x)} = \frac{L(0)}{R(0)} = \frac{g(y)}{g(y)} = 1$$
$$\implies L = R$$

But these two functions are precisely the left and right side of the law of exponents for g. Thus their equality is equivalent to g sayisfying the law of exponents for this fixed value of y:

$$\forall x,\ L(x) = g(x + y) = g(x)g(y) = R(x)$$

As y was arbitrary, this holds for all y, and g is an exponential. □

This proof does not establish the *existence* of a solution to this equation, it only says *if you have a solution* then its an exponential. But we may now use the theory of power series to directly construct a solution!

Proposition 24.8. *The series $E(x) = \sum_{n\geq 0} \frac{x^n}{n!}$ satisfies $E'(x) = E(x)$ and $E(0) = 1$. Thus, it defines an exponential function.*

Proof. This series converges on the entire real line via the ratio test (as checked above). Thus it defines a continuous and differentiable function on $\mathbb{R}$, which can be differentiated term-by-term (**?@thm-pseries-differentiation**) to yield

$$\begin{aligned}
E(x) &= \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots + \frac{x^n}{n!} + \cdots\right)' \\
&= (1)' + (x)' + \left(\frac{x^2}{2}\right)' + \left(\frac{x^3}{6}\right)' + \cdots + \left(\frac{x^n}{n!}\right)' + \cdots \\
&= 0 + 1 + x + \frac{3x^2}{6} + \cdots + \frac{nx^{n-1}}{n!} + \cdots \\
&= 1 + x + \frac{x^2}{2} + \cdots + \frac{x^{n-1}}{(n-1)!} + \cdots \\
&= E(x)
\end{aligned}$$

Finally, plugging in zero yields $E(0) = 1 + 0 + \frac{0^2}{2!} + \cdots = 1$, finishing the argument. $\square$

Part VI.

Integrals

- In Chapter 25 we discuss the area problem and give a set of axioms to characterize the operation of integration.
- In Chapter 26 we give a particular limiting construction (the Darboux Integral) and prove it satisfies the axioms
- In **?@sec-int-theory** we investigate further properties of the Darboux Integral beyond the axioms; showing that it is a linear operation and all continuous functions are integrable

25. Axioms

Highlights of this Chapter: we give an *axiomatic* definition of integration, and use these axioms to propose a *formula for computing integrals* using sums.

The *integral* is meant to measure the (net) area. When f is positive, for instance, we learn in calculus that $\int_a^b f dx$ should be the area under f between a and b. That is, it should be the area of the region $\mathscr{R} = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b \;\; 0 \leq y \leq f(x)\}$. But how does one measure area? Perhaps surprisingly, this turns out to be much more difficult than it sounds, and a full resolution took until the beginnings of the 20th century, with the advent of *measure theory*.

n b Happily, we do not need the full generalities of this theory to introduce the single variable integration theories one first meets in analysis, and we can do something conceptually simpler. Following the example we set with our introduction to the elementary functions, we seek an *axiomatic description* of what integration is 'about', before we demand a procedure to calculate it. In this chapter we carry this out: proposing a simple set of axioms that anything worthy of being called an integral must satisfy, and then use these to produce a formula which can calculate integrals in every case we will need.

25.1. Characterizing Integration

If we are to propose axioms for integration, we must first think carefully about what we expect an integral to *be*. Looking back to calculus, we recall integration to be a procedure taking a function f and a closed interval $[a, b]$ and producing a number, which we denote $\int_a^b f(x)\,dx$. That is, at its most basic level, an integral is a *function* taking an interval and a function defined on that interval to a number. To remind ourselves of the calculus notation (while still maintaining something distinct, to help us stay formal), we write this function as

$$\{f, I\} \mapsto \int_I f \in \mathbb{R}$$

Fixing the interval I, we can think of $\int_I \{\}$ as a map assigning real valued functions on I to numbers, and fixing $f : \mathbb{R} \to \mathbb{R}$, we may think of $\int_{\{\}} f$ as a map assigning intervals to numbers (the integral of f over that interval). From each of these perspectives there are some obvious properties that anything worthy of being called an integral must satisfy:

- **It gets rectangles right**: If $f(x) = k \in \mathbb{R}$ is a constant function, its graph is a horizontal line, which encloses a rectangle over an interval $I = [a, b]$. But we know the area of a rectangle should be its base times its height! So $\int_{[a,b]} k$ should be $k(b - a)$.
- **It gets inequalities right** If $f(x) \leq g(x)$ on an interval I, the graph of f lies fully underneath the graph of g, so the *net area* under f cannot be greater than the area under g. That is, $f \leq g$ on I should imply $\int_I f \leq \int_I g$.
- **Area is additive** Given an interval $I = [a, b]$ one should be able to find the net area under f by breaking the interval into pieces, finding the area under f on each piece, and adding up the results. This is simplest when we consider just two pieces: for $c \in [a, b]$ we should have $\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f$.

Since these are properties we clearly want an integral to have, we might wish to take them directly as *axioms*: and basically that is what we will do! But there is a slight subtlety we need to contend with: the collection of all real valued functions contains some wild beasts, and we shouldn't be so hasty as to assume that it makes sense to measure the area under the curve of *every function* (indeed, for our theory, it will turn out that the *area under the function which is 1 on the rationals and zero on the irrationals, is undefined). So, instead of insisting these hold for all functions, we will formalize an integral as declaring a subset of functions to be integrable, and only imposing these axioms on that subset.

Definition 25.1 (Axiomatic Integration)**.** For any closed interval $J = [a, b]$ we denote by $\mathcal{I}(J)$ the set of integrable functions on J. Then a collection of functions $\mathcal{I}(J) \to \mathbb{R}$ is an *integral*, and denoted

$$f \mapsto \int_J f$$

if it satisfies the following axioms:

- If $k \in \mathbb{R}$ then $f(x) = k$ is an element of $\mathcal{I}([a, b])$ for any interval $[a, b]$ and

$$\int_{[a,b]} k = k(b - a).$$

- If $f, g \in \mathcal{I}([a, b])$ and $f(x) \leq g(x)$ for all $x \in [a, b]$ then

$$\int_{[a,b]} f \leq \int_{[a,b]} g$$

- If $[a,b]$ is an interval and $c\in(a,b)$, then $f\in\mathcal{I}([a,b])$ if and only if $f\in\mathcal{I}([a,c])$ and $f\in\mathcal{I}([c,b])$. Furthermore, in this case their values are related by

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f$$

Note these axioms do not aim to *uniquely* specify an integral, but rather to delineate properties that *anything worthy of being called an integral* must have. Over the past two centuries, there has been quite a lot of work done studying the possible *different integrals* - all the different functions that satisfy this definition. However through all this work a beautiful story has emerged: for all their differences, all the various constructions give exactly the same answers for the continuous functions, and those answers can be calculated directly from the axioms themselves! Its this streamlined, abstract thread that we will pursue in this course.

25.1.1. ♦ Improper Integrals

We have axiomatized the integral for bounded functions on closed intervals, but the definition can be naturally extended to unbounded intervals and (certain) unbounded functions via limits.

Definition 25.2 (Improper Integrals: Unbounded Intervals)**.** The integral of a bounded function f on a ray $[a,\infty)$ is defined as a limit of its integrals over growing closed intervals

$$\int_{[a,\infty)} f := \lim_{b\to\infty}\int_{[a,b]} f$$

with the analogous definition for rays $(-\infty,b]$. The integral over the entire real line is defined by taking each endpoint to $\pm\infty$ separately

$$\int_{\mathbb{R}} f := \lim_{\substack{a\to-\infty\\ b\to\infty}}\int_{[a,b]} f$$

That is, both orders $\lim_{a\to-\infty}\lim_{b\to\infty}$ and $\lim_{b\to\infty}\lim_{a\to-\infty}$ exist and are equal.

Definition 25.3 (Improper Integrals: Unbounded Functions)**.** If f is defined on $(a,b]$ and integrable on each subinterval $[t,b]$ for $t>a$, we define the improper integral on $(a,b]$ as a limit

$$\int_{(a,b]} f = \lim_{t\to a^+}\int_{[t,a]} f$$

Similarly for functions unbounded on $[a,b)$ but bounded on each $[a,t]\subset[a,b)$, we define $\int_{[a,b)} f = \lim_{t\to b^-}\int_{[a,t]} f$.

If a function defined on $[a, b]$ is unbounded in a neighborhood of some $c \in (a, b)$ but is integrable on every subinterval missing c, we say the integral on $[a, b]$ exists if and only if the integral on $[a, c)$ and $(c, b]$ both exist, in which case we define it as equal to their sum.

25.2. The Integral as a Function

WHile our definition strictly only gives the value of integrals on one interval at a time, the axioms let us do a bit of work and define the *Integral Function* $F : [a, b]$ of an integrable function $f : [a, b]$ and study its properties.

Proposition 25.1 (The Integral as a Function)**.** *If $f \in \mathcal{I}([a, b])$ is an integrable function, then there exists a function $F : [a, b] \to \mathbb{R}$ defined by*

$$F(x) = \int_{[a,x]} f$$

Proof. This is just subdivision at work: for any $x \in [a, b]$ we may write

$$[a, b] = [a, x] \cup [x, b]$$

. Then Axiom III implies that f is integrable on $[a, x]$, and so the number $\int_{[a,x]} f$ is defined. This assignment describes a real valued function

$$x \mapsto \int_{[a,x]} f$$

□

We can learn much from the axioms about this integral function: they imply that when a function is integrable, its integral is continuous!

Theorem 25.1. *If $f \in \mathcal{I}([a, b])$ is a bounded integrable function, then its integral $F(x) = \int_{[a,x]} f$ is* continuous.

Proof. Let f be integrable and bounded by M on $[a, b]$, and set $F(x) = \int_{[a.x]} f$. Begin by choosing an $\epsilon > 0$. We will prove something even strong than asked - that f is *uniformly continuous* by finding a $\delta > 0$ where if $|y - x| < \delta$ we have $|F(y) - F(x)| < \epsilon$. Let's unpack this a bit: if $x < y$ are two points of $[a, b]$,

$$F(y) - F(x) = \int_{[a,y]} f - \int_{[a,x]} f$$

But subdivision (Axiom III) implies

$$\begin{aligned} F(y) &= \int_{[a,y]} f \\ &= \int_{[a,x]} f + \int_{[x,y]} f \\ &= F(x) + \int_{[x,y]} f \end{aligned}$$

Thus $F(y) - F(x)$ is just the integral of f on the subinterval $[x, y] \subset [a, b]$. Because f is bounded by M we know $-M \leq f(x) \leq M$. By subdivsion, f is then integrable on every sub-interval $I \subset [a, b]$, and by comparison (Axiom II) this implies

$$-M|I| \leq \int_I f \leq M|I|$$

So, we choose $\delta = \epsilon/M$. This immediately yields what we want, as if $|y - x| < \delta$,

$$-\epsilon = -M\delta < -M|y - x| \leq \int_{[x,y]} f \leq M|y - x| < M\delta = \epsilon$$

Thus $|F(y) - F(x)| = \left| \int_{[x,y]} f, dx \right| < \epsilon$. □

Remark 25.1. Of course, the proven result is not *really* stronger than what was asked, since we began on a closed interval, and we know that continuous on a closed interval implies uniformly continuous.

However, if you look carefully at the proof you see we nowhere used that the original domain was a closed interval! So what we have really proven is that the area function $F(x) = \int_{[a,b]} f$ is *uniformly continuous* anytime f is bounded!

As defined, the integral is only a function of x for x greater than the chosen starting endpoint. While this is what is desired in many applications, its also useful to be able to extend the definition to make sense for x below the starting point as well. This is the integral often met in calculus, which we call the *oriented integral* as it changes sign when the interval from a to x is traced backwards.

Definition 25.4 (The Oriented Integral). Given a function f which is integrable on the interval between a and b, the *oriented integral* of f is denoted $\int_a^b f$ and equals

$$\int_a^b f = \begin{cases} \int_{[a,b]} f & a \leq b \\ -\int_{[b,a]} f & b < a \end{cases}$$

Corollary 25.1. *Given any function f which is integrable on an interval $[a, b]$, for any $c \in [a, b]$ the oriented integral defines a function $[a, b] \to \mathbb{R}$ by*

$$x \mapsto \int_c^x f$$

25.3. How to Compute?

Just as we can use these axioms to *prove theorems* about any possible integral, we can also use it to *compute values*. We content ourselves with a simple example here, to illustrate, and follow with a more general discussion.

25.3.1. Integrating f(x)=x

Given the function $f(x) = x$, lets temporarily *assume* that we have some integral $\int$ satisfying the axioms above, and also that this integral considers x to be integrable on the interval $[0, 1]$. From here, it turns out the axioms unambiguously determine its value, allowing us to prove

Theorem 25.2. *If $\int$ is* any *integral for which $f(x) = x$ is integrable, then $\int_{[0,1]} x = \frac{1}{2}$.*

We proceed in steps. First, note that on the interval $[0, 1]$ we know (by definition) that $f(x) = x$ is between 0 and 1. So, by axiom 2

$$0 \leq x \leq 1 \implies \int_{[0,1]} 0 \leq \int_{[0,1]} x \leq \int_{[0,1]} 1$$

The upper and lower bounds here are constants, and so we can evaluate their integrals by axiom 1, giving

$$0 \leq \int_{[0,1]} x \leq 1$$

This of course is a pretty terrible estimate; but we can easily use the same ideas to do better! Indeed, since we've assumed that x is integrable on $[0, 1]$ axiom 3 ensures us that its also integrable on $I_1 = [0, 1/2]$ and $I_2 = [1/2, 1]$ (or any other subintervals, for that matter). On these smaller intervals, we have better understanding of the behavior of $f(x)$, and thus better estimates:

$$x \in I_1 \implies 0 \leq f(x) = x \leq 1/2 \qquad x \in I_2 \implies 1/2 \leq f(x) = x \leq 1$$

To each (constant!) bound we can apply axiom 1 to integrate, it and then apply axiom 2 to ensure the inequality is preserved. Thus

$$0=\int_{I_1}0\leq\int_{I_1}x\leq\int_{I_1}\frac{1}{2}=\frac{1}{4}$$
$$\frac{1}{4}=\int_{I_2}\frac{1}{2}\leq\int_{I_2}x\leq\int_{I_2}1=\frac{1}{2}$$

Adding these two intequalities gives $\frac{1}{4}\leq\int_{I_1}x+\int_{I_2}x\leq\frac{3}{4}$, but then using axiom 2 we recognize that the sum in the middle is just the *subdivision* of $\int_{[0,1]}x$ at $1/2$. Thus we have

$$\frac{1}{4}\leq\int_{[0,1]}x\leq\frac{3}{4}$$

Of course we can do better. If we divide $[0,1]$ into N intervals of equal length $1/N$ (say $I_i=[(i-1)/N,i/N]$) then by using the subdivision axiom inductively, we find $f(x)=x$ is integrable on each, and

$$\int_{[0,1]}x=\int_{I_1}x+\int_{[1/n,1]}x=\int_{I_1}x+\int_{I_2}x+\int_{[2/n,1]}x=\cdots=\sum_{i=1}^{N}\int_{I_i}x$$

On each of these intervals I_i we can easily bound $f(x)=x$:since its monotone increasing, its smallest value is its left endpoint and its largest value is its right endpoint: $(i-1)/N\leq x\leq i/N$. Thus, by axioms 1 and 2

$$\frac{i-1}{N^2}=\int_{I_i}\frac{i-1}{N}\leq\int_{I_i}x\leq\int_{I_i}\frac{i}{N}=\frac{i}{N^2}$$

Summing up these inequalities, and recalling $\int_{[0,1]}x=\sum_{i=0}^{N}\int_{I_i}x$, we find

$$\sum_{i=1}^{N}\frac{i-1}{N^2}\leq\int_{[0,1]}x\leq\sum_{i=1}^{N}\frac{i}{N^2}$$

This inequality must hold for *all values of* N, which turns out to be enough to completely fix the value of $\int_{[0,1]}x$

Exercise 25.1 (Integrating x).

- Call the lower estimate L_N and the upper estimate U_N. Prove that as $N \to \infty$, so long as one of these sums converges so does the other, and their values are equal. Thus, the constant sequence $\int_{[0,1]} x$ is squeezed between L_N and U_N, so in the limit must also take their common value!

- Next, prove that $U_N \to \frac{1}{2}$ as $N \to \infty$ *Hint: use previous homework, where we did summation by parts to find a formula for $\sum_{i=1}^{N} i = 1 + 2 + 3 \cdots + N$.

Throughout this entire calculation we've only used the axioms, and the assumption that x is integrable. Thus, we've proven a pretty strong result: *no matter how you try to precisely define integrals, there is an unambiguous choice for the value of* $\int_{[0,1]} x$. *If its defined at all, it must equal exactly* $1/2$.

Exercise 25.2 (Integrating x, Part II)**.** Extend the above result to show that if x is integrable on $[a, b]$, then

$$\int_{[a,b]} x = \frac{b^2}{2} - \frac{a^2}{2}$$

(Its OK if you assume in your proof that $0 < a < b$ to cut down on worrying about negative numbers, the proof of the general case is not much more work)

Hint: first, generalize the work we did together in the book above, from the interval $[0, 1]$ *to a general interval* $[0, c]$ *and prove that* $\int_{[0,c]} x = \frac{c^2}{2}$. Then use the fact that you know this for all $c > 0$ and the subdivision axiom to get what you want.*

Exercise 25.3 (Integrating x^2)**.** Prove that if x^2 is integrable on $[0, c]$ that its value must be

$$\int_{[0,c]} x^2 = \frac{c^3}{3}$$

Use this to deduce that for any interval $[a, b]$, (feel free to just do the case $a, b \geq 0$)

$$\int_{[a,b]} x^2 = \frac{b^3}{3} - \frac{a^3}{3}$$

Hint: follow the similar process to what we did above: using the axioms to bound by sums, and then using the summation by parts formula from earlier in the course to calculate the limit

25.3.2. The Darboux Integral

We can take this kind of reasoning even farther, and propose a means of calculating integrals of arbitrary functions, whenever they are forced to exist by the axioms. The main idea is the same: to estimate an integral we use subdivision to break the domain into smaller and smaller pieces, and then use inequalities to get better and better estimates of f on each piece. It will be useful to give a name to such a subdivision of the interval: a *partition.*

Definition 25.5 (Partitions)**.** A partition of the interval $I = [a, b]$ is a finite ordered set $P = \{t_0, t_1, \ldots, t_n\}$ with $a = t_0 < t_1 < \ldots < t_{n-1} < t_N = b$.

- N is called the *length* of the partition
- We write $P_i = [t_i, t_{i+1}]$ for the i^{th} interval of P, and $|P_i| = (t_{i+1} - t_i)$ for its width.

- The *maxwidth* of P is the maximal width of the P's intervals, $\mathrm{maxwidth}(P) = \max_{0\leq i<N}\{|P_i|\}$.
- The set of all partitions on a fixed interval I is denoted $\mathscr{P}_I$.

$$\mathscr{P}_I = \{P \,:\, P \text{ is a partition of } I\}$$

On each partition, we can try to find bounds on the value of our function f. We no longer know the lower bound will occur at the left side and the upper bound at the right side, or even that the lower and upper bounds are even achieved by some points in the domain (we do know this when f is continuous, by the extreme value theorem, but we can't say much for a general f). But, even when the max and min do not exist the infimum and supremum always do (for any boounded nonempty set,by completeness).

Definition 25.6 (Upper and Lower Sums)**.** Let f be a function, and P a partition of the closed interval I. For each segment $P_i = [t_i, t_{i+1}]$, we define

$$m_i = \inf_{x\in P_i}\{f(x)\} \qquad M_i = \sup_{x\in P_i}\{f(x)\}$$

We then define the *upper sum* $U_I(f, P)$ and the *lower sum* $L_I(f, P)$ as

$$L_I(f, P) = \sum_{0\leq i<N} m_i|P_i|$$

$$U_I(f, P) = \sum_{0\leq i<N} M_i|P_i|$$

s

Using the subdivision axiom on our partition, we see that *if* f is integrable, we must have

$$\sum_{0\leq i<N} m_i|P_i| \leq \sum_{0\leq i<N}\int_{I_i} f \leq \sum_{0\leq i<N} M_i|P_i|$$

This is great in that given *any* partition we can get some nice bounds on the possible values of our integral $\int_I f$, but we can't focus on a single partition and need to think more generally. Let's look at each inequality separately. What does it mean that $L_I(f,P)$ is less than or equal to the integral, *for any partition*? Turning this around, we are saying that $\int_I f$ is an *upper bound* for the lower sums. So this upper bound can't be less than the *least upper bound*, meaning

$$\sup_{P\in\mathcal{P}_I} L_I(f,P) \leq \int_I f$$

Similarly, the second inequality tells us that $\int_I f$ is a *lower bound* for the set of all possible upper sums. It must then be less than or equal to the *greatest lower bound*, so

$$\int_I f \leq \inf_{P\in\mathcal{P}_I} U_I(f,P)$$

These quantities prove extremely useful estimates, so we will give them a name:

Definition 25.7 (Upper and Lower Integrals)**.** Let f be a function on the closed interval I. Then we define the *upper integral* $U_I(f)$ and the *lower integral* $L_I(f)$ as

$$U(f) = \inf_{P\in\mathcal{P}_I}\{U_I(f,P)\}$$

$$L(f) = \sup_{P\in\mathcal{P}_I}\{L_I(f,P)\}$$

It may happen that for a given function f, the upper integral and lower integral are not equal to one another. In this case, the best bounds we could think to construct from the axioms (looking over *all possible partitions*) aren't enough to nail down a value for the integral of the function.

Exercise 25.4. Prove the characteristic function of the rationals has $U(\chi) = 1$ and $L(\chi) = 0$ on the interval $[0,1]$.

But sometimes we will find that $U(f) = L(f)$. Since we know $L(f) \leq \int_I f$ and $\int_I f \leq U(f)$ whenver the integral exists, this then uniquely specifies the value: *if f is integrable under any possible definition of the integral, its value must be this common quantity $U(f) = L(f)$ we computed here.* Taking this one step further actually provides a really reasonable potential *definition* of an integral: given a function f we compute the upper and lower sums: if they are not equal, we declare the function is not integrable. But if they are equal, we define the integral to be their common value (as we must, as this is the only option!). This definition is due to Darboux, and is called the Darboux integral.

Definition 25.8 (Darboux Integral)**.** Let f be a function on the closed interval I. Then f is *Darboux-Integrable* on I if $U(f) = L(f)$, and we define the integral to be this common value:

$$\int_{[a,b]} f = U(f) = L(f)$$

Because we used the axioms (and only the axioms) to come up with this construction, its perhaps not surprising that the resulting thing actually does satisfy the axioms, so is an example of an integral. But since we are working with limit-like quantities (infima and suprema) we should to be careful and actually check nothing goes wrong. This is the content of the next (optional) chapter.

26. ★ Construction

> This is a technical chapter, where we prove that there really are integrals: by giving a precise construction that satisfies the axioms!

We have introduced axioms for integration, and gotten a bit comfortable using these axioms to compute things. For example, in the last chapter we proved the following facts with relatively straightforward arguments:

- *If f is integrable* and bounded, then $F(x) = \int_{[a,x]} f$ is continuous.
- *If x is integrable* then $\int_{[a,b]} x = \frac{b^2}{2} - \frac{a^2}{2}$
- *If x^2 is integrable* then $\int_{[a,b]} x^2 = \frac{b^3}{3} - \frac{a^3}{3}$

These theorems all are of the same form: they're conditional on *if an integral exists*, then we know stuff about it. This of course is kind of disappointing: but also to be expected: we've given the axioms for an integral but we haven't shown anything actually satisfies these axioms yet! This chapter fills this gap, by recalling our natural candidate for an integral, and then proving it actually does satisfy the axioms. The arguments are all very reminiscent of the calculations we did for x and x^2 but more abstract, and while straightforward they will not be a focus of ours, here at the end of the course. This is a great chapter to read if you're the kind of mathematician who has been nervously following the work in the previous chapter / final homework, worrying that maybe its all for nothing because *maybe* integrals don't exist at all! But if this hasn't been a big worry of yours, you'll miss nothing by skipping over this chapter.

We recall first the definition of the Darboux Integral. Let f be a function on the closed interval I, and write $U(f) = \inf_{P \in \mathscr{P}_I}\{U_I(f,P)\}$ and $L(f) = \sup_{P \in \mathscr{P}_I}\{L_I(f,P)\}$ for the *upper and lower sums*, with $L_I(f,P) = \sum_{0 \le i < N} m_i |P_i|$ and $U_I(f,P) = \sum_{0 \le i < N} M_i |P_i|$ for any partition P. Then f is *Darboux-Integrable* on I if $U(f) = L(f)$, and we define the integral to be this common value: $\int_{[a,b]}^{\mathscr{D}} f = U(f) = L(f)$. We will prove that this definition satisfies the axioms: that is,

- For every $k \in \mathbb{R}$ then $f(x) = k$ is Darboux integrable for any interval $[a,b]$ and $\int_{[a,b]}^{\mathscr{D}} k = k(b-a)$.
- If f, g are Darboux integrable on I and $f(x) \le g(x)$ for all $x \in I$ then $\int_I^{\mathscr{D}} f \le \int_I^{\mathscr{D}} g$

- If $[a,b]$ is an interval and $c \in (a,b)$, then f is Darboux integrable if and only if it is Darboux integrable on both $[a,c]$ and $[c,b]$. Furthermore, in this case their values are related by $\int_{[a,b]}^{\mathscr{D}} f = \int_{[a,c]}^{\mathscr{D}} f + \int_{[c,b]}^{\mathscr{D}} f$.

The proof takes place in several steps, but we begin by developing a basic theory of partitions of intervals, which are the crucial defining features of our sum.

26.1. Working with Partitions

Definition 26.1 (Partitions)**.** A partition of the interval $I = [a,b]$ is a finite ordered set $P = \{t_0, t_1, \dots, t_n\}$ with $a = t_0 < t_1 < \dots < t_{n-1} < t_N = b$.

- N is called the *length* of the partition
- We write $P_i = [t_i, t_{i+1}]$ for the i^{th} interval of P, and $|P_i| = (t_{i+1} - t_i)$ for its width.

- The *maxwidth* of P is the maximal width of the P's intervals, $\text{maxwidth}(P) = \max_{0 \le i < N}\{|P_i|\}$.
- The set of all partitions on a fixed interval I is denoted $\mathscr{P}_I$.

$$\mathscr{P}_I = \{P \,:\, P \text{ is a partition of } I\}$$

The goal of this section is to prove the seemingly obvious fact $L_I(f) \le U_I(f)$. This takes more work than it seems at first because of the definitions of $L_I(f)$ as a supremum and $U_I(f)$ as an infimum, but proves an invaluable tool in analyzing integrability.

Definition 26.2 (Refining Partitions)**.** A partition Q is a refinement of a partition P if Q contains all the points of P (that is, $P \subset Q$).

Proposition 26.1 (The Refinement Lemma)**.** *If Q is a refinement of the partition P on a closed interval I, then for any bounded function f the following inequalities hold*

$$L_I(f,P) \le L_I(f,Q) \le U_I(f,Q) \le U_I(f,P)$$

Proof. Here we give the argument for lower sums, the analogous case for upper sums is asked in Exercise 26.1. Since $P \subset Q$ and both P, Q are finite sets we know Q contains finitely many more points than P. Here we will show that if Q contains exactly one more point than P, that the claim holds; the general case follows by induction.

In this case we may write $Q = P \cup \{z\}$, where z lies within the partition $P_k = [t_k, t_{k+1}]$. Thus, $Q_k = [t_k, c]$ for the left half after subdivision, and $Q_{k+1} = [c, t_{k+1}]$ for the right half. Outside of P_k, the two partitions are identical, so their difference is given only by the difference of their values on P_k:

$$L_I(f,Q) - L_I(f,P) =$$

$$\left(\inf_{x\in Q_k}\{f(x)\}\,|Q_k| + \inf_{x\in Q_{k+1}}\{f(x)\}\,|Q_{k+1}|\right) - \left(\inf_{x\in P_k}\{f(x)\}\,|P_k|\right)$$

Since both Q_k and Q_{k+1} are *subsets* of P_k, the infimum over each of them is at its smallest the infimum over the whole set. This implies

$$\begin{aligned}
&\inf_{x\in Q_k}\{f(x)\}\,|Q_k| + \inf_{x\in Q_{k+1}}\{f(x)\}\,|Q_{k+1}| \\
&\geq \inf_{x\in P_k}\{f(x)\}|Q_k| + \inf_{x\in P_k}\{f(x)\}|Q_{k+1} \\
&= \inf_{x\in P_k}\{f(x)\}\,(|Q_k| + |Q_{k+1}) \\
&= \inf_{x\in P_k}\{f(x)\}|P_k|
\end{aligned}$$

Thus, the first term in the difference above is bigger than the second, so the overall difference is positive. Thus $L_I(f,Q) - L_I(f,P) \geq 0$ and so as claimed,

$$L_I(f,Q) \geq L_I(f,P)$$

□

Exercise 26.1. Following the structure above, prove that if Q refines P, that

$$U_I(f,Q) \leq U_I(f,P)$$

Proposition 26.2 (Upper/Lower Sum Inequality)**.** *Lower sums are always smaller than upper sums, independent of partition. That is, if P, Q be two arbitrary partitions of a closed interval I, for any bounded function f,*

$$L_I(f,P) \leq U_I(f,Q)$$

Proof. Let P and Q be two arbitrary partitions of the interval I, and consider the partition $P \cup Q$. This contains both P and Q as subsets, so is a *common refinement* of both.

Using our previous work, this implies

$$L(f,P) \leq L(f,P\cup Q) \qquad U(f,P\cup Q) \leq U(f,Q)$$

We also know that for the partition $P \cup Q$ itself,

$$L(f,P\cup Q) \leq U(f,P\cup Q)$$

Taken together these produce the the string of inequalities

$$L(f,P) \leq L(f,P\cup Q) \leq U(f,P\cup Q) \leq U(f,Q)$$

From which immediately follows that $L(f,P) \leq U(f,Q)$, as desired. □

Proposition 26.3 (Upper/Lower Integral Inequality). *Let I be any closed interval and f a bounded function on I. Then the lower integral is less than or equal to the upper integral,*

$$L_I f \leq U_I f.$$

Proof. Recall that $U(f)$ is the *infimum* over all partitions of the upper sums.
Let P be an arbitrary partition. By **?@prp-upper-lower-on-different-partitions** we know the upper sum with respect to *any partition whatsoever* is greater than or equal to $L(f,P)$, so $L(f,P)$ is a *lower bound* for the set of all upper sums.

Thus, the infimum of the upper sums - the *greatest* of all lower bounds - must be at greater or equal to this specific lower bound,

$$L(f,P) \leq \inf_{Q\in\mathscr{P}}\{U(f,Q)\} = U(f)$$

But this holds for *every partition P*. That means this number $U(f)$ is actually an *upper bound* for the set of all $L(f,P)$. And so, it must be greater than or equal to the *least upper bound* $L(f)$:

$$L(f) \leq U(f)$$

□

Corollary 26.1. *To show a function f is integrable, it suffices to show that* $U(f) \leq L(f)$.

(To see this, recall in general that $L_I(f) \leq U_I(f)$ from **?@prp-lower-int-leq-upper-int**. So, if $U_I f \leq L_I f$ then in fact they are equal, which is the definition of f being integrable.)

26.2. Integrability Criteria

Here we prove a very useful condition to test if a function is integrable, by finding sufficient partitions.

Theorem 26.1. *Let f be a bounded function on a closed interval I. Then f is integrable if for every $\epsilon > 0$ there exists a partition P of I such that*

$$U(f,P) - L(f,P) < \epsilon$$

Here we prove one direction of this theorem, namely that if such partitions exist for all $\epsilon > 0$ then f is integrable. We prove the converse below.

Proof. Let $\epsilon > 0$, and assume there is a partition P with

$$U_I(f,P) - L_I(f,P) < \epsilon$$

Then, recalling $L_I(f,P) \leq L_I(f)$ and $U_I(f) \leq U_I(f,P)$ by definition, we chain these together with $L_I(f) \leq U_I(f)$ to get

$$L_I(f,P) \leq L_I(f) \leq U_I(f) \leq U_I(f,P)$$

Thus, the interval $[L_I(f), U_I(f)]$ is contained within the interval $[L_I(f,P), U_I(f,P)]$ which has length $< \epsilon$. Thus its length must also be less than ϵ:

$$0 \leq U_I(f) - L_I(f) \leq \epsilon$$

But ϵ was arbitrary! Thus the only possibility is that $U_I(f) - L_I(f) = 0$, and so the two are equal, meaning f is integrable as claimed. □

Now we prove the second direction of Theorem 26.1: the proof is reminiscent of the triangle inequality, though without absolute values (as we know terms of the form $U - L$ are always nonnegative already)

Proof. Assume that f is integrable, so $L_I(f) = U_I(f)$. Since $U_I(f)$ is the greatest lower bound of all the upper sums, for any $\epsilon > 0$, $U_I(f) + \frac{\epsilon}{2}$ is not a lower bound: that is, there must be some partition P_1 where

$$U_I(f,P_1) < U_I(f) + \frac{\epsilon}{2}$$

Similarly, since $L_I(f)$ is the least upper bound of the lower sums, there must be some partition P_2 with

$$L_I(f,P_2) > L_I(f) - \frac{\epsilon}{2}$$

Now, define $P = P_1 \cup P_2$ to be the common refinement of these two partitions, and observe that

$$\begin{aligned} U_I(f,P) - L_I(f,P) &\leq U_I(f,P_1) - L_I(f,P_2) \\ &< \left(U_I(f) + \frac{\epsilon}{2}\right) - \left(L_I - \frac{\epsilon}{2}\right) \\ &= U_I(f) - L_I(f) + \epsilon \\ &= \epsilon \end{aligned}$$

Where the last inequality uses $L_I(f) = U_I(f)$. Thus, for our arbitrary ϵ we found a partition on which the upper and lower sums differ by less than that, as claimed. □

And finally, we provide an even stronger theorem than ϵ-integrability, that lets us prove a function is integrable *and* calculate the resulting value, by taking the limit of carefully chosen sequences of partitions. More precisely, we want to consider any sequence of partitions that's *getting finer and finer*:

Definition 26.3 (Shrinking Partitions)**.** A sequence $P_n \in \mathcal{P}_I$ of partitions is said to be *shrinking* if the corresponding sequence of max-widths converges to 0.

We often abbreviate the phrase P_n *is a shrinking sequence of partitions* by $P_n \to 0$.

Theorem 26.2 (Integrability & Shrinking Partitions)**.** *Let f be a function on the interval I, and assume that P_n, P_n' are two sequences of shrinking partitions such that*

$$\lim L_I(f, P_n) = \lim U_I(f, P_n')$$

Then, f is integrable on I and $\int_I f$ is equal to this common value.

Proof. Call this common limiting value X. As $L_I f$ is defined as a supremum over all lower sums

$$\begin{aligned}
\lim L_I(f, P_n) &\leq \sup_{\{n \in \mathbb{N}\}} \{L_I(f, P_n)\} \\
&\leq \sup_{P \in \mathcal{P}_I} \{L_I(f, P)\} \\
&= L_I(f)
\end{aligned}$$

Similiarly, as $U_I(f)$ is the *infimum* over all upper sums, we have

$$\lim U_I(f, P_n') \geq U_I(f)$$

By **?@prp-lower-int-leq-upper-int** we know $L_I(f) \leq U_I(f)$, which allows us to string these inequalities together:

$$\lim L_I(f, P_n) \leq L_I(f) \leq U_I(f) \leq \lim U_I(f, P_n')$$

Under the assumption that these two limits are equal, all four quantities in this sequence must be equal, and in particular $L_I(f) = U_I(f)$. Thus f is integrable, and its value coincides with the limit of either of these sequences of shrinking partitions, as claimed. □

26.3. Verification of Axioms

With these tools in hand we verify the axioms of integration hold for the Darboux integral. For readability, we write $\int$ instead of $\int^{\mathscr{D}}$ throughout.

Proposition 26.4 (Integrability of Constants)**.** *Let $f(x) = k$ be a constant function, and $[a,b]$ an interval. Then k is Darboux integrable on $[a,b]$ and*

$$\int_{[a,b]} k = k(b-a)$$

Proof. For any partition P, we have

$$M_i = \sup_{x\in P_i}\{f(x)\} = k = \inf x \in P_i\{f(x)\} = m_i$$

as f is constant. Thus,

$$U(f,P) = \sum_{P_i\in P} M_i|P_i| = k\sum_{P_i\in P}|P_i| = k(b-a)$$

$$L(f,P) = \sum_{P_i\in P} m_i|P_i| = k\sum_{P_i\in P}|P_i| = k(b-a)$$

The upper and lower sums are *constant*, independent of partition, and so their respective infima/suprema are also constant, equal to this same value. Thus k is integrable, and the integral is also this common value

$$\int_{[a,b]} k = k(b-a)$$

□

Proposition 26.5 (Integration and Inequalities)**.** *Let f, g be Darboux integrable functions on $[a,b]$ and assume that $f(x) \leq g(x)$ for all $x \in [a,b]$. Then*

$$\int_{[a,b]} f \leq \int_{[a,b]} g$$

Proof. The constraint $f \leq g$ implies that on any partition P we have

$$L(f,P) \leq L(g,P)$$

Or, equivalently $L(g,P) - L(f,P) \geq 0$. Taking the supremum over all P of this set of nonnegative numbers yields a nonnegative number, so

$$\sup_{P\in\mathscr{P}_{[a,b]}}\{L(g,P) - L(f,P)\} \geq 0$$

$$L(g) - L(f) \geq 0 \implies L(f) \leq L(g)$$

But since we've assumed f and g are integrable we know that $L(f) = U(f) = \int_{a,b} f$ and $L(g) = U(g) = \int_{[a,b]} g$. Thus

$$\int_{[a,b]} f \leq \int_{[a,b]} g$$

□

Proposition 26.6 (Integration and Subdivision)**.** *Let $[a,b]$ be an interval and $c \in (a,b)$. Then a function f defined on $[a,b]$ is Darboux-integrable on this interval if and only if it is Darboux integrable on both $[a,c]$ and $[c,b]$. Furthermore, when defined these three integrals satisfy the identity*

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f$$

Proof. First, assume that f is integrable on $[a,b]$. By **?@thm-epsilon-integrability**, this means for any $\epsilon > 0$ there exists a partition P where $U(f,P) - L(f,P) < \epsilon$. Now consider the refinement $P_c = P \cup \{c\}$. By the refinement lemma,

$$L(f,P) \leq L(f,P_c) \leq U(f,P_c) \leq U(f,P)$$

Thus $U(f,P_c) - L(f,P_c) < \epsilon$ as well. Next we take this partition and divide it into partitions of each subinterval $P_1 = P_c \cup [a,c]$ and $P_2 = P_c \cup [c,b]$. By simply re-grouping the finite sums, we see

$$L(f,P_c) = L(f,P_1) + L(f,P_2) \qquad U(f,P_c) = U(f,P_1) + U(f,P_2)$$

And, by the definitions of upper and lower sums, for each we know $U(f,P_i) - L(f,P_i) \geq 0$. All that remains to insure the integrability of f on $[a,c]$ and $[c,b]$ is to show that these differences are individually less than ϵ. But this is immediate, as for example,

$$\begin{aligned}
U(f,P_1) - L(f,P_1) &\leq U(f,P_1) - L(f,P_1) + (U(f,P_2) - L(f,P_2)) \\
&= (U(f,P_1) + U(f,P_2)) - (L(f,P_1) + L(f,P_2)) \\
&= U(f,P_c) - L(f,P_c) \\
&\leq \epsilon
\end{aligned}$$

and the same argument applies to $U(f,P_2) - L(f,P_2)$. □

Next we assume integrability on the two subintervals, and prove integrability on the whole interval.

Proof. Let $\epsilon > 0$ and by our integrability assumptions choose partitions P_1 of $[a, c]$ and P_2 of $[c, b]$ such that

$$U(f, P_i) - L(f, P_i) \leq \frac{\epsilon}{2} \qquad i \in \{1, 2\}$$

Now, their union $P = P_1 \cup P_2$ is a partition of $[a, b]$, and re-grouping the finite sums, we see

$$L(f, P) = L(f, P_1) + L(f, P_2) \qquad U(f, P) = U(f, P_1) + U(f, P_2)$$

Thus,

$$\begin{aligned}
U(f, P) - L(f, P) &= (U(f, P_1) + U(f, P_2)) - (L(f, P_1) + L(f, P_2)) \\
&= (U(f, P_1) - L(f, P_1)) + (U(f, P_2) - L(f, P_2)) \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon
\end{aligned}$$

□

So, we see that integrability on $[a, b]$ is equivalent to integrability on $[a, c]$ and $[c, b]$. Finally, we need to show in the case where all three integrals are defined, the subdivision equality actually holds.

Proof. Let P be any partition of the interval $[a, b]$ and define the usual suspects:

$$P_c = P \cup \{c\} \qquad P_1 = P_c \cup [a, c] \qquad P_2 = P_c \cup [c, b]$$

We need three pieces of data. First, the inequalities relating integrals to upper and lower sums

$$L(f, P_1) \leq \int_{[a,c]} f \leq U(f, P_1) \qquad L(f, P_2) \leq \int_{[c,b]} f \leq U(f, P_2)$$

Second, the inequalities of refinements:

$$L(f, P) \leq L(f, P_c) \leq U(f, P_c) \leq U(f, P)$$

and third, the relationships between P_1, P_2 and P_c:

$$L(f, P_c) = L(f, P_1) + L(f, P_2) \qquad U(f, P_c) = U(f, P_1) + U(f, P_2)$$

Putting all of these together, we get both lower and upper estimates for the sum of the integrals over the subdivision:

$$L(f,P) \leq L(f,P_c) = L(f,P_1) + L(f,P_2) \leq \int_{[a,c]} f + \int_{[c,b]} f$$

$$\int_{[a,c]} f + \int_{[c,b]} f \leq U(f,P_1) + U(f,P_2) = U(f,P_c) \leq U(f,P)$$

And concatenating these inequalities gives the overall bound, for any *arbitrary* partition P:

$$L(f,P) \leq \int_{[a,c]} f + \int_{[c,b]} f \leq U(f,P)$$

Thus, the sum of these integrals lies between the upper and lower sum of f on $[a,b]$ for *every partition*. As f is integrable, we know there is a *single number with this property*, and that number is by definition the integral. Thus

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f$$

□

Phew! We've successfully verified all three axioms for the Darboux integral. Taken together, these prove that our construction really is an integral!

Corollary 26.2. *The equality of upper and lower sums satisfies the axioms of integration, and thus the Darboux Integral really does define an integral.*

27. ⋆ Examples

> In this optional chapter we integrate familiar functions directly from the definition. I've written this merely to illustrate its *possible* (inspired by a challenge posed by past students to me), not because its *useful*: this entire chapter is rendered entirely superflouous following our proof of the fundamental theorem of calculus!

Recall our definition of *axiomatically integrable* says that f is integrable on $[a, b]$ only if *every possible definition of $\int$ satisfying the axioms agrees on the value of $\int_{[a,b]} f$. This is quite a slippery concept to work with, so we developed the concepts of upper and lower sums to help us out.

27.1. Powers

Here

Proposition 27.1 (Integrating $f(x) = x$). *Let $[a, b]$ be any closed interval in $\mathbb{R}$. Then $f(x) = x$ is integrable on $[a, b]$ and*

$$\int_{[a,b]} x = \frac{b^2 - a^2}{2}$$

Proof. Start with $[0, b]$, then look at $0 < a < b$ using interval subdivision. To show x is integrable, it is enough to find a sequence P_n of shrinking partitions where $\lim L(f, P_n) = \lim U(f, P_n)$. Then $U(f) = L(f)$ necessarily.

For each n, let P_n be the evenly spaced partition with n subintervals, of width $\Delta_n = (b - a)/n$. Since $f(x) = x$ is monotone increasing, we know on each subinterval $[t_{i-1}, t_i]$ that

$$m_i = t_{i-1} = (i - 1)\Delta_n \qquad M_i = t_i = i\Delta_n$$

Thus, the upper and lower sums for these partitions are

$$\begin{aligned} L(x, P_n) &= \sum_{1 \le i \le n} m_i \Delta_n = (i - 1)\Delta_n \Delta_n \\ &= \Delta_n^2 \left(0 + 1 + 2 + \cdots + (n - 1)\right) \end{aligned}$$

$$\begin{aligned} U(x, P_n) &= \sum_{1 \leq i \leq n} M_i \Delta_n = i\Delta_n\Delta_n \\ &= \Delta_n^2 (1 + 2 + \cdots + n) \end{aligned}$$

These are nearly identical formulae: the upper sum is just one term longer than the lower sum and so their difference is

$$U(x, P_n) - L(x, P_n) = n\Delta_n^2 = n\frac{b^2}{n^2} = \frac{b^2}{n}$$

As $n \to \infty$ this converges to zero: thus, if either the upper *or* lower sum converges, then both do, and both converge to the same value by the limit theorems. For example, if we prove $U(f, P_n)$ converges then

$$\begin{aligned} \lim L(x, P_n) &= \lim (U(x, P_n) - U(f, P_n) + L(x, P_n)) \\ &= \lim U(x, P_n) - \lim(U(x, P_n) - L(x, P_n)) \\ &= \lim U(xs, P_n) + 0s \end{aligned}$$

□

So, we focus on just proving that $U(x, P_n)$ converges and finding its value.

Exercise 27.1. Use the sum of the first n integers we have previously derived to prove that $\lim U(x, P_n) = \frac{b^2}{2} \lim(1 + 1/n) = \frac{b^2}{2}$.

Thus x is integrable on $[0, b]$ and

$$\int_{[0,b]} x = \frac{b^2}{2}$$

Knowing this, we complete the case for a general positive interval $[a, b]$ with $0 < a < b$ by subdivision:

Exercise 27.2. Show that

$$\int_{[a,b]} x = \frac{b^2 - a^2}{2}$$

*Hint: do $0 < a < b$ first, then deal with the case where one or both may be negative, with another subdivision.

Proposition 27.2 (Integrating $f(x) = x^2$)**.** *Let* $[a, b]$ *be any closed interval in* $\mathbb{R}$*. Then* $f(x) = x^2$ *is integrable on* $[a, b]$ *and*

$$\int_{[a,b]} x^2\,dx = \frac{b^3 - a^3}{3}$$

Exercise 27.3. Following the same technique as above, show that x^2 is integrable on $[a, b]$:

- First, restrict yourself to intervals of the form $[0, b]$ for $b > 0$.
- Use the monotonicity of x^2 on these intervals to explicitly write out upper and lower sums.
- Use the following identity on sums of squares from elementary number theory to compute their value

$$\sum_{1 \leq k \leq N} k^2 = \frac{N(N+1)(2N+1)}{6}$$

- Explain how to generalize this to intervals of the form $[a, 0]$ for $a < 0$, and finally to general intervals $[a, b]$ for any $a < b \in \mathbb{R}$ using subdivision.

27.2. Exponentials

Here's a quite long calculation showing that it's possible to integrate exponential functions directly from first principles. The length of this calculation alone is a good selling point for the fundamental theorem of calculus! There are several facts about exponentials we will need from our previous investigations; listed here for ease of reference.

- Exponentials are always nonzero
- Exponentials are strictly increasing, or strictly decreasing
- Exponentials are differentiable everywhere

Proposition 27.3 (Integrating Exponentials)**.** *Let* E *be an exponential function, and* $[a, b]$ *an interval. Then* E *is integrable on* $[a, b]$ *and*

$$\int_{[a,b]} E = \frac{E(b) - E(a)}{E'(0)}$$

Proof. We will show the argument for E an increasing exponential (its base $E(1) > 1$): an identical argument applies to decreasing exponentials (only switching U and L in the computations below).

To show $E(x)$ is integrable, it is again enough to find a sequence P_n of shrinking partitions where $\lim L(f, P_n) = \lim U(f, P_n)$. Indeed - for each n, let P_n denote the evenly spaced partition of $[a, b]$ with widths $\Delta_n = (b-a)/n$

$$P_n = \{a, a+\Delta_n, a+2\Delta_n, \cdots, a+n\Delta_n = b\}$$

We will begin by computing the lower sum. Because E is continuous, it achieves a maximum and minimum value on each interval $P_i = [t_i, t_{i+1}]$. And, since E is monotone increasing, this value occurs at the leftmost endpoint. Thus,

$$\begin{aligned} L(E, P_n) &= \sum_{0 \le i < n} \inf_{P_i}\{E(x)\}|P_i| \\ &= \sum_{0 \le i < n} E(t_i)\Delta_n \\ &= \sum_{0 \le i < n} E(a+i\Delta_n)\Delta_n \end{aligned}$$

Using the law of exponents for E we can simplify this expression somewhat:

$$\begin{aligned} E(a+i\Delta_n) &= E(a)E(i\Delta_n) \\ &= E(a)E(\Delta_n + \Delta_n + \cdots + \Delta_n) \\ &= E(a)E(\Delta_n)E(\Delta_n)\cdots E(\Delta_n) \\ &= E(a)E(\Delta_n)^i \end{aligned}$$

Plugging this back in and factoring out the constants, we see that the summation is actually a partial sum of a geometric series:

$$\begin{aligned} \sum_{0 \le i < n} E(a+i\Delta_n)\Delta_n &= \sum_{0 \le i < n} E(a)E(\Delta_n)^i\Delta_n \\ &= E(a)\Delta_n \sum_{0 \le i < n} E(\Delta_n)^i \end{aligned}$$

Having previously derived the formula for the partial sums of a geometric series, we can write this in closed form:

$$\sum_{0 \le i < n} E(\Delta_n)^i = \frac{1 - E(\Delta_n)^n}{1 - E(\Delta_n)}$$

But, we can simplify even further! Using again the laws of exponents we see that $E(\Delta_n)^n$ is the same as $E(n\Delta_n)$, and $n\Delta_n$ is nothing other than the width of our entire interval, so $b-a$. Thus the numerator becomes $1-E(b-a)$, and putting it all together yields a simple expression for $L(E, P_n)$:

$$L(E, P_n) = E(a)\Delta_n \frac{1 - E(b-a)}{1 - E(\Delta_n)}$$

Some algebraic re-arrangement is beneficial: first, note that by the laws of exponents we have

$$\begin{aligned} E(a)(1 - E(b-a)) &= E(a) - E(b-a)E(a) \\ &= E(a) - E(b) \end{aligned}$$

Thus for every n we have

$$L(E, P_n) = (E(a) - E(b)) \frac{\Delta_n}{1 - E(\Delta_n)}$$

We are interested in the limit as $n \to \infty$: by the limit laws we can pull the constant $E(a) - E(b)$ out front, and only concern ourselves with the fraction involving Δ_n. There's one final trick: look at the negative reciprocal of this fraction:

$$\frac{-1}{\frac{\Delta_n}{1-E(\Delta_n)}} = \frac{E(\Delta_n) - 1}{\Delta_n}$$

Because we know $E(0) = 1$ for all exponentials, this latter term is none other than the *difference quotient defining the derivative* for E! Since we have proven E to be differentiable, we know that evaluating this along any sequence converging to zero yields the derivative at zero. And as $\Delta_n \to 0$ this implies

$$\lim \frac{E(\Delta_n) - E(0)}{\Delta_n} = E'(0)$$

Thus, our original limit $\Delta_n/(1 - E(\Delta_n))$ is the negative reciprocal of this, and

$$\begin{aligned}
\lim L(E, P_n) &= \lim\,(E(a) - E(b))\,\frac{\Delta_n}{1 - E(\Delta_n)} \\
&= (E(a) - E(b))\lim\frac{\Delta_n}{1 - E(\Delta_n)} \\
&= (E(a) - E(b))\frac{-1}{E'(0)} \\
&= \frac{E(b) - E(a)}{E'(0)}
\end{aligned}$$

Phew! That was a lot of work! Now we have to tackle the upper sum. But luckily this will not be nearly as bad: we can reuse most of what we've done! Since E is monotone increasing, we know that the maximum on any interval occurs at the rightmost endpoint, so

$$\begin{aligned}
U(E, P_n) &= \sum_{0 \leq i < n} \sup_{P_i}\{E(x)\}|P_i| \\
&= \sum_{0 \leq i < n} E(t_{i+1})\Delta_n \\
&= \sum_{0 \leq i < n} E(a + (i+1)\Delta_n)\Delta_n
\end{aligned}$$

Comparing this with our previous expression for $L(E, P_n)$, we see (unsurprisingly) its identical except for a shift of $i \mapsto i + 1$. The law of exponents turns this additive shift into a multiplicative one:

$$\begin{aligned}
U(E, P_n) &= \sum_{0 \leq i < n} E(a + (i+1)\Delta_n)\Delta_n \\
&= \sum_{0 \leq i < n} E(\Delta_n)E(a + i\Delta_n)\Delta_n \\
&= E(\Delta_n)\sum_{0 \leq i < n} E(a + i\Delta_n)\Delta_n \\
&= E(\Delta_n)L(E, P_n)
\end{aligned}$$

Thus, $U(E, P_n) = E(\Delta_n)L(E, P_n)$ for every n. Since E is continuous,

$$\lim E(\Delta_n) = E(\lim \Delta_n) = E(0) = 1$$

And, as $L(E, P_n)$ converges (as we proved above) we can apply the limit theorem for products to get

$$\begin{aligned}\lim U(E,P_n) &= \lim(E(\Delta_n)L(E,P_n))\\ &= (\lim E(\Delta_n))\,(\lim L(E,P_n))\\ &= \lim L(E,P_n)\\ &= \frac{E(b)-E(a)}{E'(0)}\end{aligned}$$

Thus, the limits of our sequence of upper and lower bounds are equal! And, by the argument at the beginning of this proof, that squeezes $L(E)$ and $U(E)$ to be equal as well. Thus, E is integrable on $[a,b]$ and its value is what we have squeezed:

$$\int_{[a,b]} E = \frac{E(b)-E(a)}{E'(0)}$$

□

Corollary 27.1 (Integrating the Natural Exponential)**.** *On any interval $[a,b]$ the natural exponential is integrable, and*

$$\int_{[a,b]} \exp = \exp(b) - \exp(a)$$

27.3. Trigonometric Functions

Theorem 27.1. *For $x \in [0,\pi/2]$, the sine function is integrable and*

$$\int_{[0,x]} \sin = 1 - \cos(x)$$

Proof. On the interval $[0,x]$, we use the sequence of evenly spaced shrinking partitions P_n of width $\Delta = x/n$, and prove integrability by showing $\lim L(\sin,P_n) = \lim U(\sin,P_n)$. Because sin is monotonically increasing on $[0,\pi/2]$ on any subinterval $I=[a,b]$ that $m=\sin a$ and $M=\sin b$. Thus

$$L(\sin,P_n) = \sum_{i=1}^{n} \sin\left((i-1)\Delta\right)\Delta$$

$$U(\sin,P_n) = \sum_{i=1}^{n} \sin\left(i\Delta\right)\Delta$$

Using $\sin(0)=0$ we see the sums agree except for the final term of U, meaning

$$U(\sin, P_n) - L(\sin, P_n) = \sin(n\Delta)\Delta = \sin(x)\frac{x}{n}$$

As x is a fixed constant this tends to zero as $n \to \infty$, so sin is integrable on $[0, x]$ and we can compute its value as the limit of either the upper or lower sum.
We use the identity for $\sum_{1\leq k\leq n} \sin kx$ proven in **?@exr-summing-angles**:

$$U(\sin, P_n) = \sum_{1\leq i\leq n} \sin(i\Delta)\Delta = \frac{\sin\left(\frac{n}{2}\Delta\right)\sin\left(\frac{n+1}{2}\Delta\right)}{\sin\left(\frac{\Delta}{2}\right)}\Delta$$

Substituting back $\Delta = x/n$ and re-arranging,

$$U(\sin, P_n) = \frac{\sin\left(\frac{x}{2}\right)\sin\left(\frac{n+1}{n}\frac{x}{2}\right)}{\sin\left(\frac{x}{2n}\right)}\frac{x}{n} = \frac{\sin\left(\frac{x}{2}\right)\sin\left(\frac{n+1}{n}\frac{x}{2}\right)}{\frac{\sin(x/2n)}{x/n}}$$

We evaluate the limit as $n \to \infty$ using the limit laws. The numerator is immediate

$$\sin\left(\frac{x}{2}\right)\sin\left(\frac{n+1}{n}\frac{x}{2}\right) \mapsto \sin\left(\frac{x}{2}\right)\sin\left(\frac{x}{2}\right)$$

using that $\frac{n+1}{n} \to 1$ and the continuity of sin. For the denominator, we use the fact that $\frac{\sin x}{x} \to 1$ (**?@cor-sinc**) to see

$$\frac{\sin\left(\frac{x}{2n}\right)}{x/n} = \frac{1}{2}\frac{\sin\left(\frac{x}{2n}\right)}{\frac{x}{2n}} \to \frac{1}{2}$$

Thus

$$\lim U(\sin, P_n) = \frac{\sin\left(\frac{x}{2}\right)\sin\left(\frac{x}{2}\right)}{\frac{1}{2}} = 2\sin^2\frac{x}{2}$$

Using the half-angle identity Exercise 20.11, we can rewrite this

$$\lim U(\sin, P_n) = 2\frac{1-\cos(x)}{2} = 1 - \cos x$$

As we've already shown sin to be integrable, this limit of upper sums over a sequence of shrinking partitions gives the value:

$$\int_{[0,x]} \sin = 1 - \cos(x)$$

□

We can leverage this result and the symmetries of the sine function to calculate the integral over arbitrary intervals:

Exercise 27.4. Prove that sin is integrable on the interval $[\pi/2, \pi]$ and for any $x \in [\pi/2, \pi]$

$$\int_{[\frac{\pi}{2},x]} \sin = -\cos(x)$$

Hint: proceed either (1) directly, using the fact that sin *is decreasing on this interval or (2) using the above, and the symmetry* $\sin(\pi/2 + x) = \sin(\pi/2 - x)$.

Use this and subdivision to show for any $x \in [0, \pi]$,

$$\int_{[0,x]} \sin = 1 - \cos x$$

Corollary 27.2.

$$\int_{[0,\pi/2]} \sin = 1 \qquad \textit{and} \qquad \int_{[0,\pi]} \sin = 2$$

Exercise 27.5. Use the fact that sine is an odd function and integrable on $[0, \pi]$ to show sin is integrable on $[-\pi, 0]$ and for any $x \in [-\pi, 0]$

$$\int_{[x,0]} \sin = \cos(x) - 1$$

Again by subdivision we can conclude that sin is integrable on $[-\pi, \pi]$.

Proposition 27.4. *Let* $a, b \in [-\pi, \pi]$. *Then* sin *is integrable on* $[a, b]$ *and*

$$\int_{[a,b]} \sin = \cos(a) - \cos(b)$$

Proof. We proceed by cases depending on the location of a, b. If both are positive and lie in $[0, \pi/2]$ we evaluate using Exercise 27.4

$$\begin{aligned}\int_{[a,b]} \sin &= \int_{[0,b]} \sin - \int_{[0,a]} \sin \\ &= (1 - \cos b) - (1 - \cos a) \\ &= \cos a - \cos b\end{aligned}$$

A similar calculation applies if $a, b < 0$. If $a < 0$ and $b > 0$ we evaluate as

$$\begin{aligned}\int_{[a,b]} \sin &= \int_{[a,0]} \sin + \int_{[0,b]} \sin \\ &= (\cos a - 1) + (1 - \cos b) \\ &= \cos a - \cos b\end{aligned}$$

□

Corollary 27.3.

$$\int_{[-\pi,\pi]} \sin = 0$$

Since sin is 2π periodic this is enough to conclude that sin is in fact integrable on *any* interval

Theorem 27.2 (Integrating sine)**.** *Let $a < b$. Then* sin *is integrable on* $[a, b]$ *and*

$$\int_{[a,b]} \sin = \cos(a) - \cos(b)$$

Exercise 27.6. Prove this.

This work has immediate payoff for integrating cosine as well, since we know it to be just a shifted version of the sine:

Theorem 27.3 (Integrating cosine)**.** *Let $a < b$. Then* cos *is integrable on* $[a, b]$ *and*

$$\int_{[a,b]} \cos = \sin(b) - \sin(a)$$

Exercise 27.7. Prove Theorem 27.3 using that $\sin(x + \pi/2) = \cos(x)$ and $\cos(x + \pi/2) = -\sin(x)$ (**?@exr-trig-shift**).

28. Properties

28.1. Integrability

With our minimal notion of integration, we now confront the question of what functions are actually integrable. We know it can't be *all* (as we've seen the Characteristic function of the rationals is one non-example), but it is a very large class of functions - containing all the reasonable functions we wil need! We prove the most important result first, that all continuous functions are integrable, and then continue to show how the similar method generalizes to related cases.

Theorem 28.1 (Continuous functions are Integrable). *Every continuous function on a closed interval is Darboux integrable.*

Proof. Let f be continuous on the interval $[a, b]$ and choose $\epsilon > 0$. We will prove integrability by finding a partition P such that $U(f, P) - L(f, P) < \epsilon$.

As f is continuous it is bounded (by the extreme value theorem), so the upper and lower sums are defined for all partitions. It is also *uniformly continuous* (as $[a, b]$ is a closed interval), so we can find a δ such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{b - a}$$

Now, choose a partition P of $[a, b]$ where the width of each interval is less than δ. Comparing upper and lower sums on this interval,

$$U(f, P) - L(f, P) = \sum_{P_i \in P} M_i|P_i| - \sum_{P_i \in P} m_i|P_i| = \sum_{P_i \in P} [M_i - m_i]|P_i|$$

Since $|P_i| < \delta$, we know that for any $x, y \in P_i$ the values $f(x), f(y)$ differ by less than $\epsilon/(b - a)$. Thus the difference of between the infimum and supremum over this interval must be less than or equal to this bound:

$$M_i - m_i \leq \frac{\epsilon}{b - a}$$

Using this to bound our sum, we see

$$U(f,P) - L(f,P) = \sum_{P_i \in P} [M_i - m_i]|P_i| \leq \frac{\epsilon}{b-a} \sum_{P_i \in P} |P_i|$$
$$= \frac{\epsilon}{b-a}(b-a) = \epsilon$$

Thus, f is integrable! □

But the Darboux integral allows us to integrate even more things than the continuous functions. For example, it is quite straightforward to prove that *all monotone functions are integrable* (even those with many discontinuities!)

Theorem 28.2 (Monotone functions are Integrable). *Every monotone bounded function on a closed interval is integrable.*

Proof. Without loss of generality let f be monotone *increasing* and bounded on the interval $[a,b]$ and choose $\epsilon > 0$. We will prove integrability by finding a partition P such that $U(f,P) - L(f,P) < \epsilon$.

Let $B = f(b) - f(a)$ be the difference between values of f at the endpoints. If $B = 0$ then f is constant, and we already know constant functions are integrable so we are done.

Otherwise, let P be an arbitrary evenly spaced partition of widths $\Delta = \epsilon/B$, we consider the difference $U(f,P) - L(f,P)$:

$$U(f,P) - L(f,P) = \sum_{P_i \in P} M_i|P_i| - \sum_{P_i \in P} m_i|P_i|$$
$$= \sum_{P_i \in P} [M_i - m_i]|P_i| = \Delta \sum_{P_i \in P} [M_i - m_i]$$

Since f is increasing, its supremum on each interval occurs on the right, and its infimum on the left. That is, if $P_i = [t_{i-1}, t_i]$ we have

$$m_i = f(t_{i-1}) \qquad M_i = f(t_i)$$

Plugging this into the above gives a telescoping sum!

$$U(f,P) - L(f,P) = \Delta \sum_{1 \leq i \leq n} [f(t_i) - f(t_{i-1})] = \Delta[f(t_n) - f(t_0)]$$

But $t_0 = a$ and $t_n = b$ are the endpoints of our partition, and so this equals

$$= \Delta[f(b) - f(a)] = \frac{\epsilon}{f(b) - f(a)}[f(b) - f(a)] = \epsilon$$

□

And, inductively its straightforward to show (via subdivision) that if the domain of a function can be partitioned into finitely many intervals on which it is integrable, than its integrable on the entire thing. Thus, for example piecewise continuous functions are Darboux Integrable. The precise statement and theorem is below.s

Definition 28.1 (Piecewise Integrable Functions)**.** A function f defined on a domain I is *piecewise integrable* if I is the disjoint of a finite sequence of intervals $I = I_1 \cup I_2 \cup \ldots \cup I_n$, and f restricted to each interval is integrable.

Proposition 28.1 (Piecewise Integrable $\implies$ Integrable)**.** *If f is piecewise integrable, then it is integrable.*

Proof. We begin with the case that f is piecewise integrable on two subintervals, $[a, c]$ and $[c, b]$ of the interval $[a, b]$. Then the subdivision axiom immediately implies that f is in fact integrable on the entire interval.

Now, assume for induction that all functions that are piecewise integrable on intervals with $\leq n$ subdivisions are actually integrable, and let f be a piecewise integrable function on a union of $n + 1$ intervals

$$[a, b] = I_1 \cup I_2 \cup \cdots \cup I_n \cup I_{n+1}$$

Set J equal to the union of the first n, so that $[a, b] = J \cup I_{n+1}$. Then when restricted to J, the function f is piecewise integrable on n intervals, so its integrable by assumption. And so, f is integrable on both J and I_{n+1}, so its piecewise integrable with *two intervals*, and hence integrable as claimed. □

Because all continuous functions and all monotone functions are integrable, we have the following useful corollary covering most functions usually seen in a calculus course.

Corollary 28.1. *All piecewise continuous and piecewise monotone functions with finitely many pieces are integrable.*

28.2. ⋆ Linearity

We can continue using the definition of the Darboux integral to discover some familiar properties: proving that *if* a given function is integrable, so is any constant multiple, and *if* two functions are integrable, so is their sum. While straightforward, these proofs are a bit tedious as we must again dig into the definitions of the upper and lower sums. Happily, this section is optional to us as we will quickly have much easier proofs of these facts available to us for *continuous functions*, once we've proven the fundamental theorem of calculus.

Theorem 28.3 (Integrability of Constant Multiples). *Let f be an integrable function a closed interval I, and $c \in \mathbb{R}$. Then the function cf is also integrable on I, and furthermore*

$$\int_I cf = c\int_I f$$

We separate into cases depending on the sign of c. Below we complete $c \geq 0$, and leave $c < 0$ as an exercise.

$c = 0$. When $c = 0$ the function cf is identically the zero function. Thus by Proposition 26.4

$$\int_I cf = \int_I 0 = 0|I| = 0$$

This is equal to $c\int_I f = 0\int_I f = 0$, so we've proven $c\int_I f = \int_I cf$ as required. □

$c > 0$. For $c > 0$, note that on any interval J we have

$$\inf_{x\in J}\{cf(x)\} = c\inf_{x\in J}\{f(x)\} \qquad \sup_{x\in J}\{cf(x)\} = c\sup_{x\in J}\{f(x)\}$$

Thus for any partition P,

$$L(cf,P) = \sum_i \inf_{x\in P_i}\{cf(x)\}|P_i| = c\sum_i \inf_{x\in P_i} f(x)|P_i| = cL(f,P)$$

$$U(cf,P) = \sum_i \sup_{x\in P_i}\{cf(x)\}|P_i| = c\sum_i \sup_{x\in P_i} f(x)|P_i| = cU(f,P)$$

Let P_n be any sequence of shrinking partitions: since f is integrable we know $\lim U(f,P_n) = \lim L(f,P_n) = \int_I f$. Computing with the limit laws

$$\lim L(cf,P_n) = \lim cL(f,P_n) = c\lim L(f,P_n) = c\int_I f$$

$$\lim U(cf,P_n) = \lim cU(f,P_n) = c\lim U(f,P_n) = c\int_I f$$

Thus the upper and lower sums are equal in the limit, so cf is integrable and its integral is equal to their common value $c\int_I f$. □

Exercise 28.1. Prove the $c = -1$ case: if f is integrable on I then so is $-f$ and $\int_I(-f) = -\int_I f$. *Hint: what does multiplying by -1 do to $m = \inf$ and $M = \sup$ on each partition? What does it do to the sums $U(f,P)$ and $L(f,P)$?*

Combine this with the $c > 0$ case above to prove the analogous result for any negative constant multiple.

Theorem 28.4 (Integrability of Sums). *Let f, g be integrable functions on a closed interval I. Then their sum $f+g$ is also integrable on I. Furthermore, its integral is the sum of the integrals of f and g:*

$$\int_I (f+g) = \int_I f + \int_I g$$

Proof. The key inequality bounding sums of functions on an arbitrary interval J is

$$\inf_{x\in J}\{f(x)\} + \inf_{x\in J}\{g(x)\} \leq \inf_{x\in J}\{f(x)+g(x)\} \leq \sup_{x\in J}\{f(x)+g(x)\} \leq \sup_{x\in J}\{f(x)\} + \sup_{x\in J}\{g(x)\}$$

Given an arbitrary partition P, summing over the subintervals P_i yields

$$L(f,P) + L(g,P) \leq L(f+g,P) \leq U(f+g,P) \leq U(f,P) + U(g,P)$$

By assumption f and g are both integrable, so we may select a sequence P_n of shrinking partitions such that

$$\lim L(f,P_n) = \lim U(f,P_n) = \int_I f \qquad \lim L(g,P_n) = \lim U(g,P_n) = \int_I g$$

Taking the limit of the above inequalities along this sequence of partitions yields

$$\int_I f + \int_I g \leq \lim L(f+g,P) \leq \lim U(f+g,P) \leq \int_I f + \int_I g$$

Thus by the squeeze theorem these limits are equal; so $f+g$ is integrable, and its integral equals their common value $\int_I f + \int_I g$. □

Each of these theorems does two things: it proves something about the *space of integrable functions* and also about *how the integral behaves* on this space. Below we rephrase the conclusion of these theorems in the terminology of linear algebra - a result so important it deserves the moniker of "Theorem" itself.

Theorem 28.5 (Linearity of the Riemann/Darboux Integral). *For each interval $[a,b] \subset \mathbb{R}$, the set $\mathcal{I}([a,b])$ of Riemann integrable functions forms a* Vector Subspace *of the set of all functions $[a,b] \to \mathbb{R}$. On this subspace, the Riemann integral defines a linear map*

$$\int_{[a,b]} : \mathcal{I}([a,b]) \to \mathbb{R}$$

28.3. ⋆ Dominated Convergence for Integrals

Being able to move limits in and out of sums proved to be an incredibly useful skill in our work with functions defined as sums. Similarly, being able to move limits in and out of integrals proves to be a very useful property for reasoning with calculus. So we pause here to prove a version of the Dominated Convergence Theorem for our minimal integral, the Darboux Integral. This theorem is often called the Arzela Bounded Convergence Theorem, and was first proved by Arzela in 1885.

Theorem 28.6 (Dominated Convergence for the Darboux Integral)**.** *Let $\{f_n\}$ be a sequence of Darboux integrable functions on a closed interval I, and assume that the functions f_n converge pointwise to a Riemann integrable function f. Then if there exists some M where $|f_n(x)| < M$ for all $x \in I$, the order of integration and limit may be interchanged:*

$$\lim \int_I f_n = \int_I f$$

We prove this result in steps; doing a special case $f = 0$ first, and then using the special case to argue the general case. Our proof follows the beautifuly short paper A Concise, Elementary Proof of Arzela's Bounded Convergence Theorem

Proposition 28.2. *Let f_n be a sequence of integrable functions on $[0,1]$, with $f_n(x) \in [0,1]$ and $f_n(x) \to 0$ for all $x \in [0,1]$. Then the limit of their integrals also tends to zero:*

$$\lim_n \int_{[0,1]} f_n = 0$$

Proof. We prove the contrapositive, and show that if $\lim \int_{[0,1]} f_n \neq 0$ the functions must not actually tend to zero, at least at some point. Since Let $[0,1]_n = \int_{[0,1]} f_n$ be the sequence of integrals. Since this is not tending to 0, we can pick a bound 2ϵ (written this way for convenience), and for every N find some $n > N$ where $|\mathcal{I}_n| > 2\epsilon$. Potentially passing to this subsequence (and negating), we can assume without loss of generality we can simply assume $\mathcal{I}_n > 2\epsilon$ for all n.

For each n, we can define a sequence U_n of open subintervals of I of length at least ϵ, where f_n is always greater than ϵ on U_n. To see this, recall our upper and lower estimates: since the integral is *greater* than 2ϵ, we can find a partition P on which the lower sum is *at least* 2ϵ (if we could not, then all lower sums would be $< 2\epsilon$, and so the supremum over lower sums would be $\leq 2\epsilon$. But this is the value of our integral!) Call a rectangle *short* if its height is less than ϵ, and *tall* if its height is $\geq \epsilon$. The collection of all short rectangles have a total area less than ϵ (as their bases together are a subset of the unit interval, of length 1). Since the total area of all the rectangles is $\geq 2\epsilon$, there must *also* be tall rectangles (to make up for the extra at least ϵ of area required). And, since $f_n(x) \leq 1$ for all x, this ϵ of area requires at least ϵ of base. So take U_n to be these (open) bases, then $f_n(x) > \epsilon$ for all $x \in U_n$ and U_n's total length is at least ϵ as required.

Recall our overall goal is to find an x where $f_n(x) \not\to 0$. If there is an x which lies in infinitely many of these sets U_n we are done, because then for these infinitely many values of n we know $f_n(x) > \epsilon$, so the whole sequence cant possibly be converging to zero! So, we've reduced our question about integration to a question about intervals. But we will make one more reduction - for each N build the set $V_N = \cup_{k \geq N} U_k$; the union of all the U's past a given point. Since each of the sets U_n contains intervals of total length at least ϵ, the set V_N also has total length at least ϵ, but now the V_N sets are *nested*: each V_N contains V_{N+1}. So, seeking a point in infinitely many of the Us is the same as seeking a point in the intersection of the Vs, and all we need to do is show the intersection of all the V_N is nonempty. This is a quite intuitive statement, and indeed a straighforward proof:

Theorem: the intersection of a nested collection of intervals, each of total length $> \epsilon$ has total length $\geq \epsilon$. (though the proof goes a bit outside of the techniques we are focusing on here so we omit it, and refer the interested reader to the argument in A Concise, Elementary Proof of Arzela's Bounded Convergence Theorem).

Having found a point in the intersection of V_N gives us a point in infinitely many U_n, which means a point x where $f_n(x) \not\to 0$, finishing the proof of the contrapositive. □

Exercise 28.2. Extend the above proof to apply to sequences of functions $f_n : I \to [0, 1]$ where I is any interval (not just unit length).

Now onto the general case.

Proof. Let f_n be a sequence of Darboux integrable functions on an interval I, which converge pointwise to f. Consider the functions $f_n(x) - f(x)$: since f_n and f are bounded by M in absolute value, their difference is bounded by $2M$, so set

$$g_n(x) = \frac{1}{2M}(f_n(x) - f(x))$$

Then the g_n are a sequence of functions with range $[0, 1]$, which are integrable and converge pointwise to zero on I. Thus we can apply the special case above to conclude

$$\lim_{n\to\infty} \int_I \frac{1}{2M}(f_n - f) = 0$$

From here, we can apply the linearity of the integral for each fixed n to conclude

$$\lim_{n\to\infty} \left(\frac{1}{2M} \int_I f_n - \frac{1}{2M} \int_I f \right) = 0$$

Since the second integral (of f) is constant, we can apply limit laws to conclude the limit of $\int_I f_n$ exists, and distribute the limits:

$$\frac{1}{2M} \lim_n \int_I f_n - \frac{1}{2M} \int_I f = 0$$

Multiplying by $2M$ to clear this constant factor and adding $\int_I f$ to both sides yields

$$\lim_n \int_I f_n = \int_I f$$

as desired. $\square$

Just as we used dominated convergence for series to prove the *continuity* of power series, we can use dominated convergence for integrals to prove the *integrability* of power series. However, we will not pause to do so, as thanks to the fundamental theorem of calculus, we will soon have the technology for a much quicker proof.

Part VII.

Calculus

- In Chapter 29 we prove the Fundamental Theorem of Calculus directly from the axioms of integration
- In **?@sec-calc-antidifferentiation** we use our knowledge of differentiation and the fundamental theorem to prove familiar results about integrals, including u-substitution and integration by parts.
- In Chapter 30 we use calculus to help learn more about elementary functions, finding integral representations for the logarithm and inverse trigonometric functions, and an infinite product for $\sin(x)$.
- In **?@sec-calc-ode** we use calculus to study linear first order differential equations, and prove an existence/uniqueness theorem.
- In Chapter 31 we put everything we've learned to work to find formulae to calculuate π.

29. The Fundamental Theorem

This chapter brings together the two calculus tools of differentiation and integration, proving the fundamental theorem of calculus. We then use this theorem to derive familiar techniques for computing integrals.

The fundamental theorem of calculus is a beautiful result for many different reasons. One of course, is that it forges a deep connection between the theory of areas and the theory of derivatives - something missed by the ancients and left undiscovered until the modern advent of the calculus. But second, it shows how incredibly constraining our simple axioms are: we will not prove the fundamental theorem of calculus for any particular definition of the integral (Riemann's, Lebesgue's, Darboux's, etc) but rather showed that *if continuous functions are integrable* then your theory of integration has no choice whatsoever on how to integrate them!

Theorem 29.1 (The Fundamental Theorem of Calculus)**.** *Let f be a continuous function and assume that f is integrable on $[a, b]$. Denote its area function by*

$$F(x) = \int_{[a,x]} f$$

Then F is differentiable, and for all points $x \in (a, b)$,

$$F' = f$$

Proof. Because f is continuous on a closed interval, it is bounded (by the Extreme Value theorem), and so the area function F is continuous (Theorem 25.1).

Choose an arbitrary $c \in (a, b)$. We wish to show that $F'(c) = f(c)$: that is, we need

$$\lim_{x \to c} \frac{F(x) - F(c)}{x - c} = f(c)$$

In terms of ϵs and δs, this means for arbitrary ϵ we need to find a δ such that if x is within δ of c, this difference quotient is within ϵ of $f(c)$.

It will be convenient to separate this argument into two cases, depending on if $x < c$ or $c < x$ (both arguments are analogous, all that changes is whether the interval in question is $[c, x]$ or $[x, c]$). Below we proceed under the assumption that $c < x$. In this case, looking at the numerator, we see by subidvision (Axiom III) that

$$\begin{aligned}
F(x) &= \int_{[a,x]} f \\
&= \int_{[a,c]} f + \int_{[c,x]} f \\
&= F(c) + \int_{[c,x]} f
\end{aligned}$$

$$\implies F(x) - F(c) = \int_{[c,x]} f$$

Thus the real quantity of interest is this integral over $[c, x]$. Choose $\epsilon > 0$. Since f is continuous, there is some $\delta > 0$ where $|x - c| < \delta$ implies $|f(x) - f(c)| < \epsilon$. Equivalently, for all $x \in [c - \delta, c + \delta]$ we have

$$f(c) - \epsilon < f(x) < f(c) + \epsilon$$

By subdivison (Axiom III), we know that f is integrable on $[c, x]$, and so by comparison (Axiom II) and the area of rectangles (Axiom I) we have

$$(f(c) - \epsilon)(x - c) \leq \int_{[c,x]} f \leq (f(c) + \epsilon)(x - c)$$

Dividing through by $x - c$

$$f(c) - \epsilon \leq \frac{\int_{[c,x]} f}{x - c} \leq f(c) + \epsilon$$

and subtracting $f(c)$

$$-\epsilon \leq \frac{\int_{[c,x]} f}{x - c} - f(c) \leq \epsilon$$

We arrive at the inequality

$$\left| \frac{\int_{[c,x]} f}{x - c} \right| < \epsilon$$

But the numerator here is none other than $F(x) - F(c)$! So, we've done it: for all $x > c$ with $|x - c| < \delta$, we have the difference quotient within ϵ of $f(c)$. This implies the limit exists, and that

$$F'(c) = f(c)$$

$\square$

Exercise 29.1. Write out the case for $x < c$ following the same logic as above.

This tells us that the area function of f is one of its antiderivatives! The theory of area is the *inverse* of the theory of rates of change. But which antiderivative? The mean value theorem assures us that the collection of all possible antiderivatives are easy to understand - any two differ by a constant (Corollary 22.2). So to uniquely specify an antiderivative its enough to give its value at one point. And we can do this! But we first need a small lemma.

One technical condition that will be useful to us later is to think about what happens when the interval is of zero size: intuitively the 'net area' over a point should be zero, but can we prove that from the axioms? Indeed we can!

Proposition 29.1 (Integrating over a Degenerate Interval)**.** *If $\{c\}$ is the degenerate closed interval containing a single point, and f is a function which is integrable on any interval containing a, then*

$$\int_{\{a\}} f = 0$$

Proof. Let f be integrable on the interval $[u, v]$ and $a \in [u, v]$ be a point. Without loss of generality we can in fact take a to be one of the endpoints of the interval, by subdivision: if $a \in (u, v)$ then Axiom III implies that f is integrable on $[u, a]$ and on $[a, v]$ as well.

Thus, we assume f is integrable on $[a, v]$, and further subdivide this interval as

$$[a, v] = [a, a] \cup [a, v] = \{a\} \cup [a, v]$$

By subdivision, we see that f is integrable on $\{a\}$ and that

$$\int_{[a,v]} f = \int_{\{a\}} f + \int_{[a,v]} f$$

Subtracting the common integral over $[a, v]$ from both sides yields the result,

$$\int_{\{a\}} f = 0$$

□

Now that we know integration over a point is zero, we know $F(a) = \int_{[a,a]} f = 0$, which determines the *precise* antiderivative that appears.

Corollary 29.1. *Let f be a continuous function which is integrable on $[a, b]$. Then the function $F(x) = \int_{[a,x]} f$ is uniquely determined as the antiderivative of F such that $F(a) = 0$.*

This connection of integration with antidifferentiation and the classification of antiderivatives has a useful corollary for computation, which is often called the *second fundamental theorem*

Theorem 29.2 (FTC Part II)**.** *Let f be continuous and integrable on $[a,b]$ and let F be* any *antiderivative of f. Then*

$$\int_{[a,b]} f = F(b) - F(a)$$

Proof. Denote the area function for f as $A(x) = \int_{[a,x]} f$. Then the quantity we want to compute is $A(b)$.

Now, let F be any antiderivative of f. The first part of the fundamental theorem assures us that A is an antiderivative of f, and so Corollary 22.2 implies there is some constant C such that $A(x) - F(x) = C$, or $F(x) = A(x) + C$. Now computing,

$$\begin{aligned} F(b) - F(a) &= (A(b) + C) - (A(a) + C) \\ &= A(b) - A(a) + (C - C) \\ &= A(b) - A(a) \\ &= A(b) \end{aligned}$$

Where the last equality comes from the fact that $A(a) = \int_{\{a\}} f = 0$. $\square$

We are going to have a lot of endpoint-subtraction going on, so its nice to have a notation for it.

Definition 29.1. Let $[a,b]$ be an interval and f a function. We write

$$f\Big|_{[a,b]} = f(b) - f(a)$$

as a shorthand for evaluation at the endpoints.

Remark 29.1. It is often convenient when doing calculations to introduce a slight generalization of the integral, which depends on an *oriented interval.* A natural notation for this is already in use in calculus, using the top and bottom of the integral sign for the locations of the 'ending' and 'starting' bound respectively:

$$\int_a^b f = \begin{cases} \int_{[a,b]} f & a \leq b \\ -\int_{[b,a]} f & a \geq b \end{cases}$$

Show that using this notation, we have a clean *generalized subdivision rule*: for **all points a, b, c irrespective of their orderings,

$$\int_a^b f = \int_a^c f + \int_c^b f$$

This notation helps shorten the computations in the proof of the fundamental theorem (at the expense of adding one new thing to remember).

29.1. Linearity

We already have *general proofs* that the integral is linear, over any axiomatically integrable functions (in the chapter *Properties*, of the previous part). But now with the fundamental theorem in hand we can provide much simpler proofs, at least when restricted to continuous functions. We give these arguments here.

Theorem 29.3. *Let f be a continuous function $[a, b]$, and $k \in \mathbb{R}$. Then*

$$\int_{[a,b]} kf = k \int_{[a,b]} f$$

Proof. Since f is continuous on $[a, b]$ it is integrable. Set $F(x) = \int_{[a,x]} f$; and note by the fundamental theorem F we have $F' = f$. Since kf is a constant multiple of a continuous function it is also continuous, and hence integrable. Using the fundamental theorem, we can compute its integral by finding an antiderivative. But this is easy! Since the derivative is linear,

$$(kF(x))' = k\,(F(x))' = kf(x)$$

So kF is an antideriavtive of kf. Thus we can use it to evaluate our integral,

$$\int_{[a,b]} kf = kF(b) - kF(a)$$

Factoring out the k, this is $k(F(b) - F(a))$, and (by our definition of F!) $F(b) - F(a) = \int_{[a,b]} f$. Putting this all togehter gives the claimed identity,

$$\int_{[a,b]} kf = k \int_{[a,b]} f$$

□

Theorem 29.4. *Let f, g be continuous on $[a, b]$. Then*

$$\int_{[a,b]} (f + g) = \int_{[a,b]} f + \int_{[a,b]} g$$

Proof. Since f, g are both continuous so is $f+g$, and hence all three are integrable. Let F, G be antiderivatives of f, g respectively, and note by the linearity of the derivative that

$$(F+G)' = F' + G' = f + g$$

Thus $F+G$ is an antiderivative of $f+g$, so we can use it to evaluate the integral:

$$\int_{[a,b]}(f+g) = (F+G)\Big|_{[a,b]} = (F(b)+G(b)) - (F(a)+G(a))$$

Regrouping the right hand side as $(F(b)-F(a)) + (G(b)-G(a))$ we recognize each as the result of applying the fundamental theorem to calculate the integrals of f, g respectively. Thus

$$\int_{[a,b]}(f+g) = \int_{[a,b]} f + \int_{[a,b]} g$$

□

29.2. Integration Techniques

We can use the fundamental theorem to justify the main integration techniques learned in calculus courses - substitution and integration by parts - as simply antiderivatives of the chain rule and product rule.
:::{#thm-integral-substitution} ## Substitution Let g be a continuously differentiable function on $[a,b]$ and f be continuous on the range of g, with F an antiderivative of f. Then

$$\int_{[a,b]} f(g(x))g'(x) = F \circ g\Big|_{[a,b]}$$

::: :::{.proof} Note g is continuous as it is differentiable. As compositions and products of continuous functions are continuous, $f(g(x))g'(x)$ is a continuous function, and hence integrable. Thus by the fundamental theorem of calculus we can evaluate its integral by finding an antiderivative. The chain rule readily confirms $F \circ g$ is such a function as

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

Thus $\int_{[a,b]} f(g(x))g'(x) = F(g(b)) - F(g(a))$. :::

This justifies the familiar use of *u-substitution* in Calculus

Example 29.1. To integrate $(2x+1)^5$ on the interval $[a,b]$, note that we may write

$$(2x+1)^5 = f(g(x))$$

for $f(x) = x^5$ and $g(x) = 2x + 1$. Then $2(2x+1)^5$ is the derivative of $\frac{1}{6}(2x+1)^6$, so

$$\int_{[a,b]} 2(2x+1)^5 = \frac{1}{6}(2x+1)^6\Big|_{[a,b]}$$

By the linearity of the integral $\int_{[a,b]} 2(2x+1)^5 = 2\int_{[a,b]}(2x+1)^5$ and solving for this yields

$$\int_{[a,b]}(2x+1)^5 = \frac{(2x+1)^6}{12}\Big|_{[a,b]} = \frac{(2b+1)^6 - (2a+1)^5}{12}$$

Theorem 29.5 (Integration by Parts)**.** *Let f be continuous and g continuously differentiable on $[a,b]$. Then*

$$\int_{[a,b]} f(x)g(x) = F(x)g(x)\Big|_{[a,b]} - \int_{[a,b]} F(x)g'(x)$$

where F is an antiderivative of f.

Proof. Since f is continuous we know it is integrable, so let $F(x) = \int_{[a,x]} f$. Then F is differentiable (by the fundamental theorem) and so is g (by assumption), so the product $F(x)g(x)$ is a differentiable function. Taking the derivative with the product rule yields

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

where in the last equality we used that $F' = f$ from the fundamental theorem. Thus $F(x)g(x)$ is an *antiderivative* of the sum on the right hand side, so integrating gives

$$\int_{[a,b]} (f(x)g(x) + F(x)g'(x)) = F(x)g(x)\Big|_{[a,b]}$$

Distributing the integral over addition ☐

Corollary 29.2 (Iterated Integration by Parts)**.** *Applying twice,*

$$\begin{aligned}\int_{[a,b]} fg &= Fg\Big|_{[a,b]} - \int_{[a,b]} Fg' \\ &= Fg\Big|_{[a,b]} - \left(\mathscr{F}g'\Big|_{[a,b]} - \int_{[a,b]} \mathscr{F}g''\right) \\ &= F_g - \mathscr{F}g' + \int_{[a,b]} \mathscr{F}g''\end{aligned}$$

Where F is an antiderivative of f, and $\mathscr{F}$ is an antiderivative of F. Continuing in this fashion, we can replace our integral with one containing n derivatives of g, at the cost of having to take n antiderivatives of f.

30. Elementary Functions

> We derive the integrals of familiar functions, and along the way finally discover a formula for the logarithm.

Using the fundamental theorem of calculus we can effortless find the integrals of many important functions, simply because we know their derivatives! This is much *much* quicker than working directly from the definition (as one can appreciate by looking at the optional chapter *examples* in the last part).

30.1. Polynomials and Power Series

We begin with perhaps the first example one sees in a Calculus I course

Proposition 30.1. *The function x is integrable on any interval $[a, b] \subset \mathbb{R}$, and*

$$\int_{[a,b]} x = \frac{b^2}{2} - \frac{a^2}{2}$$

Proof. The function x is continuous, thus its integrable. We know by the power rule for differentiation that $(x^2)' = 2x$. Thus by linearity of the derivative, if $F(x) = x^2/2$, we have $F' = x$, and we can use this antiderivative to evaluate the integral via the fundamental theorem:

$$\int_{[a,b]} x = F(b) - F(a) = \frac{b^2}{2} - \frac{a^2}{2}$$

□

This technique of inverting the power rule works for all $n \neq -1$ to give a general formula

Theorem 30.1. *For any $n > 0$ the function x^n is integrable on $[a, b]$ for any $[a, b] \subset \mathbb{R}$. For $x < -1$, x^n is not defined at 0 but is integrable on any interval $[a, b]$ not containing zero. In both cases, there is a uniform formula for the integral:*

$$\int_{[a,b]} x^n = \frac{1}{n+1} x^{n+1}\Big|_{[a,b]}$$

Proof. The monomial x^n is continuous, hence integrable, and has as an antiderivative $x^{n+1}n+1$. Thus we can use this to evaluate the integral

$$\int_{[a,b]} x^n = \left.\frac{x^{n+1}}{n+1}\right|_{[a,b]}$$

□

Using linearity of the integral, this gives an immediate calculation of the antiderivative of any polynomial:

Theorem 30.2. *If $p(x) = \sum_{k=1}^{N} a_k x^k$ is any polynomial, then p is integrable on any interval $[a, b]$ and*

$$\int_{[a,b]} p(x) = \left.\sum_{k=1}^{N} \frac{a_k}{k+1} x^{k+1}\right|_{[a,b]}$$

Proof. Polynomials are continuous, hence integrable. We know an antiderivative of each term, so we compute one at a time via linearity of the integral on continuous functions:

$$\int_{[a,b]} \left(a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n\right) = \int_{[a,b]} a_0 + \int_{[a,b]} a_1 x + \int_{[a,b]} a_2 x^2 + \cdots + \int_{[a,b]} a_n x^n$$

$$= a_0 \int_{[a,b]} 1 + a_1 \int_{[a,b]} x + a_2 \int_{[a,b]} x^2 + \cdots + a_n \int_{[a,b]} x^n$$

$$= a_0 x\Big|_{[a,b]} + a_1 \frac{x^2}{2}\Big|_{[a,b]} + a_2 \frac{x^3}{3}\Big|_{[a,b]} + \cdots + a_n \frac{x^{n+1}}{n+1}\Big|_{[a,b]}$$

□

30.1.1. Power Series

As we remember well from differentiation, things that are true for *finite* sums don't always carry over easily to the limit. Indeed, the proof of differentiability of polynomials was identical to their integration above, a direct corollary of linearity. But we cannot use linearity to conclude things about limits, so for power series we instead needed to refine our tools of dominated convergence. A similar strategy goes through without fail here for integration: one can use dominated convergence for integrals (Theorem 28.6) to prove power series can be integrated term by term within their radius of converence.

But the fundamental theorem also makes a much easier technique available to us, given that we know the differentiation case! We follow that line of reasoning here.

Proposition 30.2. *Let $\sum a_n x^n$ be a power series with radius of convergence r. Then the series $\sum \frac{a_n}{n+1} x^{n+1}$ also has radius of convergence r.*

Proof. Like for the differentiable case, we prove this here under the assumption that the Ratio test succeeds in computing the radius of convergence for the original series. (As an exercise, show everything still works even if the ratio test fails, by doing the fully general argument with limsup of the root test). So for any $x \in (-R, R)$

$$\lim \left| \frac{a_{n+1}}{a_n} \right| |x| < 1$$

We now turn to compute the ratio test for our new series $\sum \frac{a_n}{n+1} x^{n+1}$: the ratio in question is

$$\frac{\frac{a_{n+1}}{n+2} x^{n+2}}{\frac{a_n}{n+1} x^{n+1}} = \left(\frac{a_{n+1}}{a_n}\right)\left(\frac{n+1}{n+2}\right) x$$

Since $(n+1)/(n+2) \to 1$ we can compute the overall limit using the limit theorems and see we end up with *the exact same limit as for the original series*! Thus integrating term by term does not change the radius of convergence at all. □

Having confirmed that $\sum \frac{a_n}{n+1} x^n$ converges when the original series does, we can provide a direct proof of the term-by-term integrability of power series, avoiding the use of dominated convergence:

Theorem 30.3 (Integrating Power Series)**.** *Let $f(x) = \sum a_n x^n$ be a power series with radius of convergence r. Then f is integrable on $(-r, r)$ and for any $0 < x < r$*

$$\int_{[0,x]} f = \sum_{n \geq 0} \frac{a_n}{n+1} x^{n+1}$$

Proof. The function $\sum a_n x^n$ is continuous on $(-r, r)$ and thus integrable by Theorem 28.1. Define $F(x) = \sum \frac{a_n}{n+1} x^n$. This converges on $(-r, r)$ by Proposition 30.2, and defines a differentiable function on this interval; whose derivative can be calculated term-by-term, giving

$$F'(x) = \left(\sum_{n \geq 0} \frac{a_n}{n+1} x^{n+1}\right)' = \sum_{n \geq 0} a_n x^n = f(x)$$

Thus F is an antiderivative of f, and by the fundamental theorem of calculus

$$\int_{[0,x]} f = F(x) - F(0)$$

Since F has no constant term, $F(0) = 0$ and so $\int_{[0,x]} f = \sum_{n \geq 0} \frac{a_n}{n+1} x^{n+1}$ as claimed. □

30.2. Exponentials and Trigonometric Functions

We put a good amount of work into defining exponential and trigonometric functions from their functional equations, proving they are continuous and eventually finding their differentiation laws. Now we reap some of the benefits, and find their integrals as effortless as we did the polynomials.

Theorem 30.4 (The Natural Exponential)**.** *THe natural exponential* $\exp(x)$ *is integrable and*

$$\int_{[a,b]} \exp = \exp(b) - \exp(a)$$

Proof. The function exp is continuous, hence integrable, and its its own derivative. Thus, its also its own antiderivative, and

$$\int_{[a,b]} \exp = \exp\Big|_{[a,b]} = \exp(b) - \exp(a)$$

□

Theorem 30.5. *Let* $E(x)$ *be any exponential function. Then* E *is integrable, and*

$$\int_{[a,b]} E = \frac{E(b) - E(a)}{E'(0)}$$

Proof. We proved that every exponential is continuous and differentiable, wtih $E'(x)$ a multiple of $E(x)$: specifically, $E' = E'(0)E$. Dividing by the constant value $E'(0)$ and using linearity of the derivative, we see

$$\left(\frac{E(x)}{E'(0)}\right)' = \frac{1}{E'(0)}\left(E(x)\right)' = \frac{1}{E'(0)}E'(0)E(x) = E(x)$$

Thus we've found an antiderivative! The fundametnal theorem fo calculus then quickly finishes the job:

$$\int_{[a,b]} E = \frac{E(x)}{E'(0)}\Big|_{[a,b]} = \frac{E(b) - E(a)}{E'(0)}$$

□

We succeed equally quickly for the basic trigonometric functions:

Theorem 30.6 (Integrating Sine and Cosine)**.**

$$\int_{[a,b]} \cos = \sin(b) - \sin(a) \qquad \int_{[a,b]} \sin = \cos(a) - \cos(b)$$

Proof. We know $\sin' = \cos$ and $\cos' = -\sin$ from previous homework, Thus, cos has sin as an antiderivative, and

$$\int_{[a,b]} \cos = \sin\Big|_{[a,b]} = \sin(b) - \sin(a)$$

Similarly, sin has $-\cos$ as an antiderivative, so

$$\int_{[a,b]} \sin = -\cos\Big|_{[a,b]} = -(\cos(b) - \cos(a)) = \cos(a) - \cos(b)$$

□

These two formulae, together with the trigonometric identities are enough to fully determine all trigonometric integrals.

30.3. Logarithms

In this section we finally develop a formula for the logarithm: we proved it existed some time ago as we had already proven the inverses of exponentials were logarithms, and we proved that exponentials exist. But this did not give us any way to *compute* a logarithm. This is in stark contrast to the exponential, where right from the beginning we had *some* means of computing it (via the cumbersome task of taking limits of rational powers), and now we have a nice extremely efficient power series. We will remedy all of this in this subsection, by studying the simple function $1/x$.

We already know there is something rather unique about this function, because it is the only case where the power rule fails us, and we can't simply use our knowledge of differentiation to invert. To start, we prove a rather simple lemma that is key to unlocking the logarithm properties:

Lemma 30.1. *If $[a, b]$ is an interval of positive numbers and $k > 0$ then the integral of $1/x$ on the intervals $[a, b]$ and $[ka, kb]$ are the same.*

Proof. Intuitively this is plausible: when we multiply by k the length of our interval increases by a factor of k but the height of our function $1/x$ *decreases* by a factor of k at each point, leaving the area fixed.

To prove it, we (surprise!) invoke the fundamental theorem of calculus. Let F be an antiderivative of $1/x$, so $F'(x) = 1/x$. Then look at $F(kx)$: taking its derivative, we see by the chain rule

$$F(kx)' = F'(kx)(kx)' = kF'(kx)$$

And using that $F' = 1/x$,

$$= k\frac{1}{kx} = \frac{1}{x}$$

Thus both $F(x)$ **and** $F(kx)$ are antiderivatives of $1/x$! This means we can use either to evaluate our integral: so, using $F(kx)$,

$$\int_{[a,b]} \frac{1}{x} = F(kx)\Big|_{[a,b]} = F(kb) - F(ka)$$

But this quantity is exactly an antiderivative of $1/x$ evalluated at kb and ka, so

$$F(kb) - F(ka) = F(x)\Big|_{[ka,kb]} = \int_{[ka,kb]} \frac{1}{x}$$

Stringing these equalities together yields the result. □

We can immediately use this to show the integral of $1/x$ has the logarithm property.

Theorem 30.7 (The Logarithm as an Integral)**.** *The function $f(x) = 1/x$ is integrable on $(0, \infty)$, and its integral*

$$L(x) = \int_{[1,x]} \frac{1}{x}$$

satisfies the law of logarithms $L(xy) = L(x) + L(y)$\$ for $x, y > 1$.

Proof. The function $1/x$ is continuous on $(0, \infty)$ so it is integrable on any closed subinterval $[a, b]$. Let $x, y > 1$ and consider the interval $[1, xy]$. We can decompose this into two intervals $[1, x] \cup [x, xy]$ and so by subdivision

$$L(xy) = \int_{[1,xy]} \frac{1}{t} = \int_{[1,x]} \frac{1}{t} + \int_{[x,xy]} \frac{1}{t}$$

The first of these terms is by definition $L(x)$, and the second can be calculated via our lemma:

$$\int_{[x,xy]} \frac{1}{t} = \int_{[1,y]} \frac{1}{t} = L(y)$$

Thus, $L(xy) = L(x) + L(y)$, as claimed! □

Exercise 30.1. Confirm this also works for arbitrary $x, y \in \mathbb{R}$ if we interpret our integral as an *oriented integral* (Definition 25.4).

Using the fundamental theorem we can easily calculate the derivative of this logarithm at 1:

Corollary 30.1 (The Natural Logarithm)**.** *The integral of* $1/x$ *is the* natural logarithm

$$\log(x) \int_{[1,x]} \frac{1}{x}$$

Proof. By the fundamental theorem $\log(x)' = \frac{1}{x}$ so evaluating at 1,

$$\log(x)'\Big|_{x=1} = \frac{1}{1} = 1$$

□

As is, this is not a very useful formula for the logarithm, as we can't use antidifferentiation to compute it: the function we care about is *defined as the antiderivative*! But this does have dramatic implications: we can use this to derive a formula for the logarithm, via power series.

Theorem 30.8 (Logarithm Power Series)**.** *The function* $\log(1+x)$ *has a power series representation for* $x \in (-1,1)$

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots$$

Proof. The geometric series converges on $(-1,1)$ to

$$\frac{1}{1-x} = \sum_{n\geq 0} x^n$$

Making the substitution $x \mapsto -x$ we can rewrite this as

$$\frac{1}{1+x} = \sum_{n\geq 0}(-x)^n = \sum_{n\geq 0}(-1)^n x^n$$

This power series converges on $(-1,1)$ still, and so is integrable term-by-term along this entire interval:

$$\int_{[0,x]} \frac{1}{1+x} = \int_{[0,x]} \sum_{n\geq 0}(-1)^n x^n = \sum_{n\geq 0}\int_{[0,x]}(-1)^n x^n = \sum_{n\geq 0}(-1)^n \frac{x^{n+1}}{n+1}$$

Since the integral of $1/x$ is $\log(x)$, its easy to see the antiderivative of $1/(1+x)$ is $\log(1+x)$, by differentiation using the chain rule:

$$(\log(1+x))' = \frac{1}{1+x}(1+x)' = \frac{1}{1+x}$$

Thus our power series is indeed a logarithm!

$$\log(1+x) = \sum_{n\geq 0}(-1)^n \frac{x^{n+1}}{n+1} = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

□

Example 30.1 (Integrating the Tangent)**.** The function $\tan(x) = \sin(x)/\cos(x)$ is continuous on $(-\pi/2, \pi/2)$ and hence integrable on any $[a, b] \subset (0, \pi/2)$. To find the value of the integral, we notice that

$$\tan(x) = \frac{\sin(x)}{\cos(x)} = \frac{1}{\cos(x)}\sin(x) = \frac{-1}{\cos(x)}(\cos(x))'$$

Since $1/x$ is the derivative of the natural logarithm function, we see this is the result of a chain rule!

$$\log(\cos(x))' = \frac{1}{\cos(x)}(\cos(x))'$$

, and so

$$-\log(\cos(x))' = \tan(x)$$

We have found an antiderivative for tangent, so the fundamental theorem yields its integral:

$$\int_{[a,b]} \tan(x) = -\log(\cos(x))\Big|_{[a,b]} = -\log(\cos(b)) + \log(\cos(a))$$

Using the rules of logarithms we can simplify this a bit to

$$\int_{[a,b]} \tan(x) = \log\left(frac\cos(a)\cos(b)\right)$$

30.4. Inverse Trigonometric Functions

Understanding the inverse trigonometric functions will prove exceedingly useful to us in our end goal of calculating π: we have defined π as a particular *input* to the trigonometic functions (the first positive input at which sine is zero, for example), and so we don't have a way to compute π by plugging something into a function: we've had to resort to methods like Newton's approximation scheme, which requires *a lot* of calculation since we are working with a power series!

Our lives would be much easier if we had functions that yielded π as an *output* (of a simple value): we could then simply derive a means of computing this function! The natural such functions would be the inverse trigonometric functions, and so we take a moment to study these here.

30.4.1. ⋆ The ArcSine

Our first fundamental problem of course is we have no idea how to get a formula for the inverse trigonometric functions! To get one, we will use the fact that we understand differentiation quite well, and then apply the fundamental theorem.

Proposition 30.3. *The derivative of the inverse sine function is*

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$$

Proof. Let $f(x) = \arcsin(x)$. Then where defined, $f(\sin(\theta)) = \theta$ by definition, and we may differentiate via the chain rule: on the left side

$$\frac{d}{d\theta} f(\sin(\theta)) = f'(\sin(\theta))\cos(\theta)$$

and on the right $\frac{d}{d\theta}\theta = 1$. Equating these and solving for f' yields

$$f'(\sin(\theta)) = \frac{1}{\cos(\theta)}$$

The only remaining problem is that we want to know f' as a function of x and we only know its value implicitly, as a function of $\sin(\theta)$. But setting $x = \sin\theta$ we can express $\cos\theta = \sqrt{1-x^2}$ via the pythagorean identity $\sin^2\theta + \cos^2\theta = 1$. Thus

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

□

Before integration this would have been a mere curiosity. But, armed wtih the fundamental theorem this is an extremely powerful fact: indeed, it directly gives us a representation as an integral:

Corollary 30.2. *The inverse sine function is defined on the interval* $[0,1]$ *by the integral*

$$\arcsin(x) = \int_{[0,x]} \frac{1}{\sqrt{1-x^2}}$$

Proof. Since $(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$, the inverse sine is an antiderivative of $\frac{1}{\sqrt{1-x^2}}$, and also $\sin(0) = 0$ implies $\arcsin(0) = 0$, so it is zero at $x = 0$. Thus, it is exactly the area function

$$\arcsin(x) = \int_{[0,x]} \frac{1}{\sqrt{1-t^2}}$$

□

Exercise 30.2. Carry out the analogous reasoning to derive an integral expression for the inverse cosine function.

30.4.2. The ArcTangent

Proposition 30.4.

$$(\arctan x)' = \frac{1}{1+x^2}$$

Proof. We again proceed by differentiating the identity $\arctan(\tan\theta) = \theta$. This yields $\arctan'(\tan\theta)\frac{1}{\cos^2\theta} = 1$ and multiplying through by $\cos^2$ we can solve for the derivative of arctangent:

$$\arctan'(\tan\theta) = \cos^2\theta$$

The only problem is again we have the derivative as a function implicitly of of $\tan\theta$, and we need it in terms of just an abstract variable x. Setting $x = \tan\theta$ we see that $x^2 = \tan^2\theta$ and (using the pythagorean identity) $x^2 + 1 = \tan^2\theta + 1 = \frac{1}{\cos^2\theta}$. Thus

$$\cos^2\theta = \frac{1}{1+x^2}$$

and putting these two together, we reach what we are after

$$\arctan'(x) = \frac{1}{1+x^2}$$

□

Proposition 30.5. *The inverse function* $\arctan(x)$ *to the tangent* $\tan(x) = \sin(x)/\cos(x)$ *admits an integral representation*

$$\arctan(x) = \int_{[0,x]} \frac{1}{1+t^2}$$

Proof. This follows as $\arctan'(x) = 1/(1+x^2)$, so both arctan and this integral have the same derivative. As antiderivatives of the same function this means that they differ by a constant. Finally, this constant is equal to zero as $\arctan(0) = 0$ and $\int_{[0,0]} \frac{1}{1+x^2} = 0$ as it is an integral over a degenerate interval. □

This integral expression is quite nice - the arctangent like the logarithm is shown to be the *integral* of a rather simple rational function. But like arcsine, an integral expression is rather difficult to use for computing actual values: we'd need to actually compute (or estimate) some Riemann sums. So it's helpful to look for other expressions as well, and here arctan has a particularly nice power series.

Recall the geometric series

$$\frac{1}{1-x} = \sum_{n\geq 0} x^n$$

We can substitute $-x^2$ for the variable here to get a series for $1/(1+x^2)$:

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

This power series has radius of convergence 1 (inherited from the original geometric series) and converges at neither endpoint. We know from the above that this function is the derivative of the arctangent, so we should integrate it!

$$\arctan(x) = \int_{[0,x]} \frac{1}{1+t^2}\,dt = \int_{[0,x]} \sum_{n\geq 0}(-1)^n t^{2n}\,dt$$

Inside its radius of convergence we can exchange the order of the sum and the integral:

$$\begin{aligned}\int_{[0,x]}\left(\sum_{n\geq 0}(-1)^n t^{2n}\right)dt &= \sum_{n\geq 0}\int_{[0,x]}(-1)^n t^{2n}\,dt\\ &= \sum_{n\geq 0}(-1)^n\int_{[0,x]} t^{2n}dt\\ &= \sum_{n\geq 0}(-1)^n\frac{x^{2n+1}}{2n+1}\end{aligned}$$

Theorem 30.9. *For $x \in (-1,1)$,*

$$\arctan(x) = \sum_{n\geq 0}(-1)^n\frac{x^{2n+1}}{2n+1}$$

$$= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \cdots$$

31. π

31.1. Area of a Circle

We have defined π as the first zero of the sine function - a definition, but have finally developed enough tools to relate it to the area of a circle. This provides a relationship between the modern, rigorous theory of trigonometric functions and the ancient quest of Archimedes to measure the area of the circle.

Indeed, since we have defined area rigorously with integration, we can now make sense of *the area of the circle* as long as we can express the unit circle as a function. While this is not directly possible, we can take the implicit equation $x^2 + y^2 = 1$ and solve for y giving *two* functions (one for the top half and one for the bottom). Then we can measure the area of the circle as twice the top half, or

$$\text{Area} = 2\int_{[-1,1]} \sqrt{1-x^2}$$

Now we compute this integral with our newfound integration techniques (substitution), and show it equals the half-period of our trigonometric functions in natural units.

Theorem 31.1.

$$2\int_{[-1,1]} \sqrt{1-x^2} = \pi$$

Proof. By subsitution, we see that the following two integrals are equal

$$\int_{[0,1]} \sqrt{1-x^2} = \int_I \sqrt{1-(\sin(t))^2}(\sin(t))'$$

Where $I = [a, b]$ is the interval such that $[\sin(a), \sin(b)] = [0, 1]$. Since $\sin(0) = 0$ and $\sin(\pi/2) = 1$ we see $I = [0, \pi/2]$. Now we focus on simplifying the integrand:

By the Pythagorean identity, $1 - \sin^2(t) = \cos^2(t)$, thus by Example 2.3,

$$\sqrt{1-\sin^2(t)} = \sqrt{\cos^2(t)} = |\cos(t)|$$

and by definition we recall $(\sin t)' = \cos t$. Thus

$$\begin{aligned}\int_{[0,\pi/2]} &= \int_{[0,\pi/2]} |\cos(t)|\cos(t)\\ &= \int_{[0,\pi/2]} \cos^2(t)\end{aligned}$$

Where we can drop the absolute value as cos is nonnegative on $[0,\pi/2]$ (its first zero is at half the period, so π). We can simplify this using the "half angle formula" $\cos^2(x) = (1+\cos(2x))/2$

$$\int_{[0,\pi/2]} \cos^2(t) = \int_{[0,\pi/2]} \frac{1+\cos(2t)}{2}$$

Using the linearity of the integral, this reduces to

$$\begin{aligned}\int_{[0,\pi/2]} \cos^2(t) &= \frac{1}{2}\int_{[0,\pi/2]} 1 + \frac{1}{2}\int_{[0,\pi/2]} \cos(2t)\\ &= \frac{\pi}{4} + \frac{1}{2}\int_{[0,\pi/2]} \cos(2t)\end{aligned}$$

The first of these integrals could be immediately evaluated as the integral of a constant, but the second requires us to do another substitution. If $u = 2t$ then

$$\int_{[0,\pi/2]} \cos(2t) = \frac{1}{2}\int_{[0,\pi]} \cos u$$

We recall again that by definition $\cos u = (\sin u)'$, so by the first fundamental theorem

$$\int_{[0,\pi]} \cos u = \int_{[0,\pi]} (\sin u)' = \sin u\Big|_{[0,\pi/]}$$

But, sin is equal to 0 both at 0 and π! So after all this work, this integral evaluates to zero. Thus

$$\begin{aligned}\int_{[0,1]} \sqrt{1-x^2} &= \int_{[0,\pi/2]} \cos^2 t\\ &= \frac{\pi}{4} + \frac{1}{2}\int_{[0,\pi]} \cos(2t)\\ &= \frac{\pi}{4} + 0\end{aligned}$$

Now, we are ready to assemble the pieces. Because x^2 is an even function so is $\sqrt{1-x^2}$, and so its integral over $[-1,1]$ is twice its integral over $[0,1]$. Thus

$$\text{Area} = 2\int_{[-1,1]} \sqrt{1-x^2} = 4\int_{[0,1]} \sqrt{1-x^2} = 4\frac{\pi}{4} = \pi$$

$\square$

This single result ties together so many branches of analysis, and proves a worthy capstone calculation for the course. However after all this work we shouldn't let ourselves be satisfied too quickly! Now that we have related the area of a circle to *trigonometry*, we can hope to use other techniques from analysis to accurately calculate its value.

31.2. Calculating π's Value

31.2.1. From the Area Integral

Having proven that π is the area of the circle, we may attempt to estimate its value by estimating the integral of $\sqrt{1-x^2}$:

$$\frac{\pi}{4} = \int_{[0,1]} \sqrt{1-x^2}$$

Using the evenly spaced partition P_n with n bars of width $\Delta = 1/n$ and the fact that $\sqrt{1-x^2}$ is monotone decreasing on $[0,1]$, we can evaluate this integral as a limit of upper sums:

$$\int_{[0,1]} \sqrt{1-x^2} = \lim U(\sqrt{1-x^2}, P_n) = \sum_{i=1}^{n} \sqrt{1-(i\Delta)}\Delta$$

Simplifying gives an explicit limit of sums to compute:

Example 31.1. The following limit of infinite sums converges to π.

$$\pi = \lim 4\sum_{i=1}^{n} \sqrt{1-\frac{i^2}{n^2}}\frac{1}{n} = \lim 4\sum_{i=1}^{n} \frac{\sqrt{n^2-i^2}}{n^2}$$

This series is difficult to compute because it involves *square roots*: irrational quantities that we will also have to approximate in order to get a good approximate value for π. It also converges *slowly*, so there's many square roots to approximate! Using a computer to help we find

$$4\sum_{i=1}^{10}\frac{\sqrt{100-i^2}}{100}\approx 2.904\ldots$$

$$4\sum_{i=1}^{100}\frac{\sqrt{10000-i^2}}{10000}\approx 3.120\ldots$$

$$4\sum_{i=1}^{1000}\frac{\sqrt{1000000-i^2}}{1000000}\approx 3.139\ldots$$

31.2.2. From Inverse Trigonometry

One may use the inverse trigonometric functions to get integral representations of π. Perhaps the most natural thought is to use that $\sin(\pi/2)=1$, so $\arcsin(1)=\pi/2$ or

$$\frac{\pi}{2}=\arcsin(1)=\int_{[0,1]}\frac{1}{\sqrt{1-x^2}}$$

This integral is *improper* as the integrand becomes unbounded in a neighborhood of $x=1$: thus it must be calculated as a limit over intervals $[0,t]$ with $t\to 1$ which is rather difficult in practice: certainly more involved than the calculation from the area integral above.

Remark 31.1. If we were not bothered by the square roots for our computation-focused goals, one could easily replace the problematic integral above with something avoiding its problems. For instance, since $\sin(\pi/4)=1/\sqrt{2}$, we have

$$\frac{\pi}{4}=\int_{[0,1/\sqrt{2}]}\frac{1}{\sqrt{1-x^2}}$$

But this is much worse in terms of square roots: if you write out a Riemann sum here it'll be a sum of nested roots, and *still* more complicated than the estimate from the area integral.

The same trouble plagues the cosine function, but things get much nicer with the tangent. We know that sin and cos are equal when evaluated at $\pi/4$, which means their ratio is $1=\tan\pi/4$. Inverting this,

Corollary 31.1.

$$\frac{\pi}{4} = \arctan(1) = \int_{[0,1]} \frac{1}{1+x^2}, dx$$

This function is integrable (its continuous), so we can compute its value as the limit of any shrinking sequence of Riemann sums. Below is an explicit example, given for evenly spaced partitions sampled at their right endpoints.

Example 31.2. The following infinite series converges to π:

$$\begin{aligned}\pi &= \lim_n 4 \sum_{i=1}^{n} \frac{1}{1+(i\Delta)^2}\Delta \\ &= \lim_n 4 \sum_{i=1}^{n} \frac{n}{n^2+i^2}\end{aligned}$$

This sequence of sums is much better to work with: each term is a rational number, so it can be computed exactly, giving a sequence of better and better rational approximations to π.

$$4 \sum_{i=1}^{10} \frac{10}{100+i^2} \approx 3.0395\ldots$$

$$4 \sum_{i=1}^{100} \frac{100}{10000+i^2} \approx 3.13155\ldots$$

$$4 \sum_{i=1}^{1000} \frac{1000}{1000000+i^2} \approx 3.140592\ldots$$

$$4 \sum_{i=1}^{1000000} \frac{1000000}{(1000000)^2+i^2} \approx 3.14159165359\ldots$$

This is great - these sums are trivial to do on a computer (I did these in a simple python for loop) and get us an accurate value for π. But we shouldn't be satisfied just yet! First of all, these sums take a while to converge - we need a thousand terms to get the first two digits after the decimal, and a million to get the first five!

31.2.3. From Series

Power series are much easier to deal with than the limits arising from integrals: to get a better approximation of a power series you *keep the terms you have, and just add more* whereas to compute better approximate Riemann sums you need to start all over from scratch! Thus its certainly advantageous from a computational perspective to look for series converging to π.

A particularly nice example is given by the arctangent, whose series we computed in Theorem 30.9 to be $\sum_{n\geq 0}(-1)^n\frac{x^{2n+1}}{2n+1}$ on the interval $(-1,1)$. Since $\tan(\pi/4)=1$, we can calculate π as $\pi/4=\arctan(1)$, which lies right at the boundary of the interval of convergence. Luckily, this proves not to be an issue

Proposition 31.1.

$$\frac{\pi}{4}=\arctan(1)=\sum_{n\geq 0}\frac{(-1)^n}{2n+1}$$

Proof. The arctangent function is continuous on $\mathbb{R}$, so

$$\arctan(1)=\arctan(\lim_{x\to 1^-}x)=\lim_{x\to 1^-}\arctan(x)$$

For $x\in(-1,1)$ the arctangent can be expressed as a power series, so

$$\lim_{x\to 1^-}\arctan(x)=\lim_{x\to 1^-}\sum_{n\geq 0}(-1)^n\frac{x^{2n+1}}{2n+1}$$

This series converges at $x=\pm 1$ by the alternating series test. Hence, by Abel's theorem (**?@thm-pseries-continuous-endpoints**) it defines a continuous function on $[-1,1]$ and so the limit can be pulled inside:

$$\begin{aligned}
\lim_{x\to 1^-}\sum_{n\geq 0}(-1)^n\frac{x^{2n+1}}{2n+1} &= \sum_{n\geq 0}(-1)^n\frac{(\lim_{x\to 1^-}x)^{2n+1}}{2n+1}\\
&= \sum_{n\geq 0}(-1)^n\frac{1^{2n+1}}{2n+1}\\
&= \sum_{n\geq 0}(-1)^n\frac{1}{2n+1}
\end{aligned}$$

Putting this all together yields the claim:

$$\frac{\pi}{4} = \arctan(1) = \sum_{n\geq 0} \frac{(-1)^n}{2n+1}$$

$\square$

This formula is exceedingly beautiful, and worthy of writing out without summation notation to take in

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots$$

However, way out here at the endpoint the series converges *very slowly*. Using a computer to do a little experimenting:

$$4\sum_{n=0}^{10} \frac{(-1)^n}{2n+1} = 3.2323\ldots$$

$$4\sum_{n=0}^{100} \frac{(-1)^n}{2n+1} = 3.1549\ldots$$

$$4\sum_{n=0}^{1{,}000} \frac{(-1)^n}{2n+1} = 3.1425\ldots$$

Like the Riemann sum approach, we needed a thousand terms to get the first two decimals right. This problem *only* occurs as we are evaluating a series at the very boundary of its interval of convergence: we know via comparison that power series converge exponentially quickly *within their radius of convergence*, so to get better behavior we should seek a point *inside* $(-1, 1)$ at which the arctan will give us information aboout π. How do we find such a value? Here's one clever possibility: we actually realize $\pi/4$ as the *sum* of two different arctangent values:

Proposition 31.2.

$$\frac{\pi}{4} = \arctan\left(\frac{1}{2}\right) + \arctan\left(\frac{1}{3}\right)$$

Proof. Let $\theta = \arctan(1/2)$ and $\psi = \arctan(1/3)$. Now use the tangent addition law $\tan(\theta+\psi) = \frac{\tan\theta + \tan\psi}{1-\tan\theta\tan\psi}$ to compute $\theta + \psi$:

$$\tan(\theta+\psi) = \frac{\frac{1}{2}+\frac{1}{3}}{1-\frac{1}{2}\frac{1}{3}} = \frac{\frac{5}{6}}{1-\frac{1}{6}} = 1$$

Thus, $\tan(\theta+\psi) = 1$ so $\theta + \psi = \pi/4$, as claimed. $\square$

Now, both $1/2$ and $1/3$ lie well within the radius of convergence of the arctangent, so we can add the two together to get a formula for π. Since series converge absolutely within their radii of convergence, we can re-arrange terms as we please, even combining the two into a single sum:

Theorem 31.2.

$$\frac{\pi}{4} = \sum_{k\geq 0} \frac{(-1)^k}{(2k+1)2^{2k+1}} + \sum_{n\geq 0} \frac{(-1)^k}{(2k+1)3^{2k+1}}$$

$$= \sum_{k\geq 0} \frac{(-1)^k}{2k+1} \left(\frac{1}{2^{2k+1}} + \frac{1}{3^{2k+1}} \right)$$

This series converges very quickly, as the exponents 2^{2k+1} and 3^{2k+1} in the denominators grow rapidly. Indeed, summing up to $N = TWO$ already gives the first two decimal digits!

$$\left(\frac{1}{2} + \frac{1}{3} \right) - \frac{1}{3} \left(\frac{1}{8} + \frac{1}{27} \right) + \frac{1}{5} \left(\frac{1}{32} + \frac{1}{243} \right) = 3.14558$$

Using up until $N = 10$ terms in this series gives the approximation

$$\pi \approx 3.14159257960635$$

Which is correct to 7 decimal digits. To get 15 significant digits using 22 terms in this series is enough!

This is truly a marvelous machine we have built - conjuring directly from the lowly geometric series an efficient formula for π.

Example 31.3. Want to be even more clever? In 1796 John Machin showed the following identity:

$$\frac{\pi}{4} = 4 \arctan(1/5) - \arctan(1/239)$$

Note: If you wish to prove this, probably the easiest way is to notice that $(5+i)^4(239-i) = -114244(1+i)$ *and use the polar form of complex numbers to get the result. See here: https://people.math.sc.edu/howard/Classes/555c/trig.pdf*

This allows you to compute π to five or six decimals without much trouble. Just using the first five terms in the series gives $\pi \approx 3.14159268240440$ so we are already good to seven decimals. Using nine terms in the series gives you 15 significant digits