GEM 360. Foundations of Analysis

Franz Lemmermeyer, Spring 2001

March 9, 2002

Abstract

The mathematicians who have inventend (Newton, Leibniz) and developed calculus (the Bernoullis, Euler, etc.) used geometric insight freely in their proofs, and even Gauss considered the fact that e.g. a polynomial of odd degree has at least one real zero as being obvious. In the 19th century, however, people began to ask why it is that $x^2 = 2$ has a solution in $\mathbb{R}$ but not in $\mathbb{Q}$. The result on which this fact is based is called the 'completeness property' of the real numbers, and is essential for nearly every theorem in calculus.

In most textbooks on analysis (including Ross's), however, the completeness of the reals is taken on faith, and the situation is not really improved by calling it an axiom (why don't we call the fundamental theorem of calculus an axiom?). There are other omissions I'm not really happy with: the square root of real numbers, for example, is freely used but never defined, and I challenge all of you to read Chapter 1 of Ross and then tell me where to find the proof that $\sqrt{2}$ is a real number but $\sqrt{-1}$ is not. And while we're at it, let me repeat Dedekind's lament that most students (actually he was talking about mathematicians) have never seen a proof of $\sqrt{2} \cdot \sqrt{3} = \sqrt{6}$.

In short: I don't consider Chapter 1 of Ross as a 'Foundation of Analysis' because it doesn't contain a foundation of the real numbers. To quote the author (p. vi):

> Even after studying this book (or writing it) it will not be easy to handle questions such as "What is a number?"

And even though the construction of the reals (and the proof of the completeness property) is quite abstract and takes much time, I will at least try to explain how mathematicians create the universe of real numbers and analysis out of the natural numbers.

Contents

Chapter 1

From $\mathbb{N}$ to $\mathbb{Q}$

1.1 The Natural Numbers $\mathbb{N}$

In every deductive theory there are certain statements you must take for granted: you can't prove theorems by assuming nothing. What we are taking for granted here are elementary notions of sets and the basic properties of natural numbers as encoded by the following statements called the Peano axioms: Let $\mathbb{N}$ be a set together with a 'successor' function s such that

N1 $1 \in \mathbb{N}$;

N2 if $x \in \mathbb{N}$, then $s(x) \in \mathbb{N}$;

N3 there is no $x \in \mathbb{N}$ with $s(x) = 1$;

N4 if $s(x) = s(y)$, then $x = y$;

N5 if S is a subset of $\mathbb{N}$ containing 1, and if $s(n) \in S$ whenever $n \in S$, then $S = \mathbb{N}$.

Remark 1. Ross is not really exact in his statement of the Peano axioms as he already assumes the existence of an addition on $\mathbb{N}$ and then simply replaces $s(n)$ by $n + 1$.

Remark 2. Axiom [N2] states that s is a map $\mathbb{N} \longrightarrow \mathbb{N}$, that is: each element of $\mathbb{N}$ gets mapped to another element of $\mathbb{N}$.

Axiom [N4] states that the map $s : \mathbb{N} \longrightarrow \mathbb{N}$ is injective. A map $f : A \longrightarrow B$ is called injective if $f(a) = f(a')$ for $a, a' \in A$ implies that $a = a'$, in other words: if different elements get mapped to different images.

Remark 3. Axiom [N5] is called the Principle of Induction. Assume you want to prove a statement $P(n)$ (like $n^2 + n$ is even) for all $n \in \mathbb{N}$; let S denote the set of natural numbers $z \in \mathbb{N}$ for which $P(n)$ is true. If you can show that $P(1)$ holds (i.e. that $1 \in S$) and that $P(s(n))$ holds whenever $P(n)$ does (i.e. that

$s(n) \in S$ whenever $n \in S$) then this axiom allows you to conclude that $P(n)$ holds for every natural number.

Naively speaking, these axioms describe the basic properties of natural numbers; logicians can prove that if a set $\mathbb{N}$ with such a successor function s satisfying [N1]–[N5] exists, then it is essentially unique (this means that the Peano axioms *characterize* the natural numbers), but we won't need this.

What we want to do here is to show how the arithmetic of the natural numbers can be defined from the Peano axioms. We start by giving the natural numbers their usual names: we put $2 := s(1)$, $3 = s(2)$, $4 = s(3)$, etc.; in particular $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$.

Exercise. Which of these Peano axioms are satisfied by the ring $\mathbb{Z}$ of integers and successor function $z \longmapsto z + 1$?

Exercise. Consider the set $N = \{1\}$ with successor function $s : N \longrightarrow N : 1 \longmapsto 1$. Show that this system satisfies all Peano axioms except one – which one?

Exercise. Consider the set $N = \{1, 2\}$ with successor function $s : N \longrightarrow N$ mapping $1 \longmapsto 2$ and $2 \longmapsto 2$. Show that this system satisfies all Peano axioms except one – which one?

Exercise. Consider the set $N = \mathbb{N} \cup (\mathbb{N} + \frac{1}{2})$, where $\mathbb{N} = \{1, 2, 3, \ldots\}$ and $\mathbb{N} + \frac{1}{2} = \{\frac{1}{2} + 1, \frac{1}{2} + 2, \ldots\}$. Define a successor function $s : N \longrightarrow N$ by mapping $n \longmapsto n + 1$ and $\frac{1}{2} + n \longmapsto \frac{1}{2} + (n + 1)$ for all $n \in \mathbb{N}$. Show that this system satisfies all Peano axioms except one – which one?

Addition

Next we define an operation $+$ on $\mathbb{N}$ that we call addition. We have to say what $m + n$ should mean. How can we do that in terms of our axioms? Well, we can certainly define $m + 1$ by putting

$$m + 1 := s(m). \tag{1.1}$$

Once we know what $m + 1$ is, we can put $m + 2 = (m + 1) + 1$. In order to find a simple rule that does it in general we observe that $m + 2 = m + s(1)$, and that our definition reads $m + 2 = s(m + 1)$. This suggests that we define addition 'inductively': assume that we already know what $m + n$ means; we then define

$$m + s(n) := s(m + n), \tag{1.2}$$

that is, $m + (n + 1) := (m + n) + 1$.

Here's our first theorem (just kidding):

Theorem 1.1. *We have* $1 + 1 = 2$.

Proof. By (1.1), we have $1 + 1 = s(1)$. On the other hand, we have defined $2 = s(1)$. $\square$

Remark. The way we have done it, the equation $2 = 1+1$ follows immediately from the definitions. There are other axiom systems (for example, it is possible to *define* natural numbers using the axioms of set theory), and in some of them the statement $1 + 1 = 2$ is a theorem whose proof requires some work. In their book 'Principia Mathematica', Whitehead & Russels prove this equation after about 500 pages. This means that they must be doing something we don't, and they do. Not only are we using naive set theory (we haven't defined notions like sets, subsets, elements etc.), we haven't even defined what a proof is, or by which methods we are allowed to draw conclusions. It is possible to make all this precise, but you shouldn't be surprised to learn that this gets quite complicated (as a matter of fact, it doesn't 'get', it *is* complicated right from the start). In any case, this is part of what is called 'mathematical logic', and some people like it. You can even get famous by doing it: Gödel was a logician, and Hofstaedter's book *Gödel, Escher, Bach* was a bestseller in the early 1980s. One of Gödel's surprising (if not shocking) results was the proof that there are statements about natural numbers that are true but cannot be proved, that is to say: cannot be derived from the Peano axioms in a finite number of steps. In fact, for each finite set of axioms including those of Peano, there are true statements about natural numbers that can't be proved. That hasn't kept number theorists from proving theorems, however: life goes on.

Lemma 1.2. *Addition $m + n$ is now defined for all $m, n \in \mathbb{N}$.*

Proof. The following proof is fairly typical for much that follows. Make sure you understand everything.

Let $m \in \mathbb{N}$ be any natural number. Let S be the set of all $n \in \mathbb{N}$ for which $m+n$ is defined. We want to show that $m+n$ is defined for all $n \in \mathbb{N}$, i.e., that $S = \mathbb{N}$. We shall accomplish this by using axiom N5 (induction).

First, we have $1 \in S$ since, by (1.1), $m + 1$ is defined as $s(m)$.

Next, if $n \in S$, then $m + n$ is defined, and since $m + s(n) = s(m + n)$ by (1.2), so is $m + s(n)$. In other words: if $n \in S$, then $s(n) \in S$.

By the Induction axiom [N5], we conclude that $S = \mathbb{N}$, hence addition $m+n$ is defined for all $n \in \mathbb{N}$ (and also for all $m \in \mathbb{N}$ since m was arbitrary). $\square$

Now we can prove all the laws that are satisfied by the addition of natural numbers; we shall be content with a few examples, however.

Proposition 1.3 (Associativity of Addition). *For all $x, y, z \in \mathbb{N}$, we have $x + (y + x) = (x + y) + z$.*

Proof. Let $x, y \in \mathbb{N}$ be arbitrary and put

$$S = \{z \in \mathbb{N} : \ x + (y + x) = (x + y) + z\}.$$

Again, S is the set of natural numbers $z \in \mathbb{N}$ for which the claim is true, and our task is to show that $S = \mathbb{N}$.

Now $1 \in S$ because

$$\begin{array}{rcll} x+(y+1) & = & x+s(y) & \text{by (1.1)} \\ & = & s(x+y) & \text{by (1.2)} \\ & = & (x+y)+1 & \text{by (1.1)} \end{array}$$

Next assume that $z \in S$. Then we want to show that $s(z) \in S$, and to this end we have to prove that $x+(y+s(z)) = (x+y)+s(z)$. Here we go:

$$\begin{array}{rcll} x+(y+s(z)) & = & x+s(y+z) & \text{by (1.2)} \\ & = & s(x+(y+z)) & \text{by (1.2)} \\ & = & s((x+y)+z) & \text{since } z \in S \\ & = & (x+y)+s(z) & \text{by (1.2)} \end{array}$$

By the induction principle, this proves that $S = \mathbb{N}$ and we are done. □

Note that there's not the smallest possibility of a gap in our proof; we explained each and every step, and the result holds without any shred of doubt.

Lemma 1.4. *For all $x \in \mathbb{N}$, we have $x+1 = 1+x$.*

Proof. Let S denote the set of all $x \in \mathbb{N}$ for which $x+1 = 1+x$. Then $1 \in S$ since $1+1 = 1+1$. Now assume that $x \in S$. Then

$$\begin{array}{rcll} s(x)+1 & = & (x+1)+1 & \text{by (1.1)} \\ & = & (1+x)+1 & \text{since } x \in S \\ & = & 1+(x+1) & \text{by Prop. 1.3} \\ & = & 1+s(x) & \text{by (1.1)} \end{array}$$

Thus $S = \mathbb{N}$ by the induction principle. □

Now we can prove

Proposition 1.5 (Commutativity of Addition). *For all $x, y \in \mathbb{N}$ we have $x+y = y+x$.*

Proof. You know the game by now: for an arbitrary $x \in \mathbb{N}$, let S denote the set of all $y \in \mathbb{N}$ such that $x+y = y+x$. By Lemma 1.4, we have $1 \in S$.

Now assume that $y \in S$. Then

$$\begin{array}{rcll} x+s(y) & = & s(x+y) & \text{by (1.2)} \\ & = & s(y+x) & \text{since } y \in S \\ & = & y+(x+1) & \text{by (1.1)} \\ & = & y+(1+x) & \text{by Lemma 1.4} \\ & = & (y+1)+x & \text{by Prop. 1.3} \\ & = & s(y)+x & \text{by (1.1)} \end{array}$$

Thus $S = \mathbb{N}$, and we are done. □

OK, now it's your turn: the proofs of the following results are left as exercises.

Proposition 1.6 (Cancellation Law). *If $x + z = y + z$ for some $x, y, z \in \mathbb{N}$, then $x = y$.*

Hint. This is easy for $z = 1$ (use Peano); now use induction.

Proposition 1.7. *For $x, y \in \mathbb{N}$, we have $x + y \neq x$.*

Hint. Use induction on x.

Theorem 1.8 (Trichotomy Law for Addition). *For any $x, y \in \mathbb{N}$, exactly one of the following three cases holds:*

(i) $x = y$;

(ii) $x = y + z$ *for some* $z \in \mathbb{N}$;

(iii) $y = x + z$ *for some* $z \in \mathbb{N}$.

Proof. We first show that no two of these statements can hold simultaneously.

Assume that (i) and (ii) are both true. Then $x = x + z$, contradicting Prop. 1.7.

The claim that (i) and (iii) [or (ii) and (iii)] cannot hold simultaneously is left as an exercise.

Now we have to prove that, given $x, y \in \mathbb{N}$, at least one of these claims is true. We consider an arbitrary $y \in \mathbb{N}$ and do induction on x, that is, we put

$$S = \{x \in \mathbb{N} : (i) \text{ or } (ii) \text{ or } (iii) \text{ is true}\}.$$

We claim that $1 \in S$. If $1 = y$, this is clear, so assume that $1 \neq y$. In this case, y must be the successor of some natural number, say $y = s(z)$. But then $y = s(z) = z + 1 = 1 + z$ by the definition of addition and commutativity, and $y = 1 + z$ shows that (iii) is satisfied.

Now we claim that $x \in S$ implies $s(x) \in S$, so assume that $x \in S$. Then we are in exactly one of three cases:

a) $x = y$; then $s(x) = s(y) = y + 1$, so (ii) holds;

b) $x = y + z$ for some $z \in \mathbb{N}$; then $s(x) = s(y + z) = y + s(z)$, so again (ii) is true.

c) $y = x + z$ for some $z \in \mathbb{N}$. If $z = 1$, then $y = s(x)$, and (i) holds. If $z \neq 1$, then $z = s(v)$ for some $v \in \mathbb{N}$, hence

$$y = x + z = x + s(v) = x + (v + 1) = (x + 1) + v = s(x) + v,$$

where we have used various properties of addition, so (iii) holds.

Thus if $x \in S$, then $s(x) \in S$, hence $S = \mathbb{N}$ by induction, and we are done. $\square$

Finally, let us introduce the well known (or so I hope) sigma notation for sums: for integers $x_1, \ldots, x_n, \ldots \in \mathbb{N}$ we define $\sum_{k=1}^{n} x_k$ inductively by

$$\sum_{k=1}^{1} x_k = x_1 \tag{1.3}$$

and

$$\sum_{k=1}^{s(n)} = \Big(\sum_{k=1}^{n} x_k\Big) + x_{n+1}. \tag{1.4}$$

Generalizing the properties proved above to finite sums is then again left as an exercise in induction; let us just mention the examples

$$\begin{aligned}
\sum_{k=n+1}^{n+m} x_k &= \sum_{k=1}^{m} x_{n+k} \\
\sum_{k=1}^{n} x_k + \sum_{k=1}^{m} x_{n+k} &= \sum_{k=1}^{n+m} x_k \\
\sum_{k=1}^{n} x_k + \sum_{k=1}^{n} y_k &= \sum_{k=1}^{n} (x_k + y_k).
\end{aligned}$$

Let us see how to prove the first claim. For $m = 1$, the claim is

$$\sum_{k=n+1}^{n+1} x_k = \sum_{k=1}^{1} x_{n+k},$$

and by our definition (1.3), both sides are equal to x_{n+1}.

Now assume that the claim is true for some $m \in \mathbb{N}$; then we claim that it also holds for $s(m)$. In fact,

$$\begin{aligned}
\sum_{k=n+1}^{n+s(m)} x_k &= \Big(\sum_{k=n+1}^{n+m} + x_k\Big) + x_{n+m+1} && \text{by (1.4)} \\
&= \Big(\sum_{k=1}^{m} x_{n+k}\Big) + x_{n+m+1} && \text{by assumption} \\
&= \sum_{k=1}^{n+s(m)} x_{n+k} && \text{by (1.4),}
\end{aligned}$$

and the claim follows. Induction does the rest.

Exercise. Prove the other two formulas.

We also should mention the following generalization of associativity: the sum $\sum_{k=1}^{n} x_k$ is by definition equal to $((((x_1 + x_2) + x_3) + x_4) + \ldots) + x_n$; "of course" this sum does not depend on how we place the brackets, but we still have to prove this. A concise formulation of this property is the following: if $x_1, \ldots, x_n$ is a finite set of natural numbers, and if $y_1, \ldots, y_n$ is a permutation of the x_k, then $\sum_{k=1}^{n} x_k = \sum_{k=1}^{n} y_k$.

Multiplication

We are now going to define the multiplication of natural numbers. For the definition of $x \cdot z$ we use induction. First we put

$$x \cdot 1 = x \tag{1.5}$$

(of course). Now assume that we have defined $x \cdot y$; then we put

$$x \cdot s(y) = x \cdot y + x \tag{1.6}$$

(in other words: we put $x \cdot (y+1) := x \cdot y + x$). It should be obvious by now that the induction principle guarantees that xy is defined for any $x, y \in \mathbb{N}$. In general, we omit the multiplication sign $\cdot$ and write xy instead of $x \cdot y$' We shall also write $xy + z$ instead of $(xy) + z$ and agree that we always evaluate expressions by multiplying first and then adding the products.

Next we prove the basic properties of multiplication:

Proposition 1.9 (Left Distributive Law). *For $x, y, z \in \mathbb{N}$ we have $x(y+z) = xy + xz$.*

Proof. Take $z, y \in \mathbb{N}$ and do induction on z. We find

$$\begin{array}{rcll} x(y+1) & = & x \cdot s(y) & \text{by (1.1)} \\ & = & xy + x & \text{by (1.6)} \\ & = & xy + x \cdot 1 & \text{by (1.5).} \end{array}$$

Next we assume that the left distributive law holds for z and prove that it also holds for $s(z)$:

$$\begin{array}{rcll} x(y+s(z)) & = & x \cdot (s(y+z)) & \text{by (1.2)} \\ & = & x(y+z) + x & \text{by (1.6)} \\ & = & (xy + xz) + x & \text{by assumption} \\ & = & xy + (xz + x) & \text{by Prop. 1.3} \\ & = & xy + x \cdot s(z) & \text{by (1.6)} \end{array}$$

This proves the claim by induction. □

Where there's a left distributive law, there's a

Proposition 1.10 (Right Distributive Law). *We have $(x+y)z = xz + yz$ for all $x, y, z \in \mathbb{N}$.*

This proof is left as an exercise. Note that right distributivity would follow immediately from left distributivity if we already knew that multiplication was commutative. Fact is, however: we don't. But it comes right next on our (that is your) agenda: we start out with commutativity for multiplication by 1:

Lemma 1.11. *For all $x \in \mathbb{N}$, we have $1 \cdot x = x$.*

and then do induction:

Proposition 1.12 (Commutativity of Multiplication). *For all $x, y \in \mathbb{N}$, we have $xy = yx$.*

Yet another exercise:

Proposition 1.13 (Associativity of Multiplication). *For $x, y, z \in \mathbb{N}$, we have $x(yz) = (xy)z$.*

And another one:

Proposition 1.14 (Cancellation Law of Multiplication). *If $xz = yz$ for x, y, $z \in \mathbb{N}$, then $x = y$.*

Exercise. Prove that $a\Big(\sum_{k=1}^{n} x_n\Big) = \sum_{k=1}^{n}(ax_n)$ for $a, x_1, \ldots, x_n \in \mathbb{N}$.

Exercise. Define $\prod_{k=1}^{n} x_k$ for $x_1, \ldots, x_k \in \mathbb{N}$.

Exercise. Prove that $\prod_{k=1}^{m} x_k \prod_{k=m+1}^{n} x_k = \prod_{k=1}^{n} x_k$.

Exercise. Prove that $\prod_{k=1}^{n} x_k \prod_{k=1}^{n} y_k = \prod_{k=1}^{n}(x_k y_k)$.

Now that we know how to multiply, we can go forth and define exponentiation a^n for $a, n \in \mathbb{N}$: we put $a^1 = a$, and if a^n is already defined, then $a^{s(n)} = a^n \cdot a$. Armed with this definition, we perform the same procedure as last time:

- a^n is defined for all $a, n \in \mathbb{N}$ (induction on n)
- $a^{m+n} = a^m a^n$ for $a, m, n \in \mathbb{N}$ (induction on n)
- $a^{mn} = (a^m)^n$ for $a, m, n \in \mathbb{N}$ (induction on n)
- $a^n b^n = (ab)^n$ for $a, b, n \in \mathbb{N}$.

This is recommended as an exercise unless it is *perfectly clear* for you how to proceed.

There is one last set of properties of the naturals that we have not yet touched upon: those based on the relation $<$.

$\mathbb{N}$ as a well-ordered set

We start by defining the relevant concept. For $x, y \in \mathbb{N}$ we write

- $x < y$ if there is an $n \in \mathbb{N}$ such that $x + n = y$;
- $x > y$ if $y < x$;
- $x \leq y$ if $x < y$ or $x = y$;
- $x \geq y$ if $y \leq x$.

We say that a set R is simply ordered if we have a relation $<$ such that the following conditions are satisfied for all $x, y, z \in R$:

O1 Trichotomy: We either have $x < y$ or $x = y$ or $x > y$.

O2 Transitivity: if $x < y$ and $y < z$ then $x < z$.

The proofs of the following claims are now straight forward:

Proposition 1.15. *The set $\mathbb{N}$ of natural numbers is simply ordered.*

Proof. That condition O1 holds follows immediately from the trichotomy law of addition (Thm. 1.8).

Now assume that $x < y$ and $y < z$. By definition, there exist $m, n \in \mathbb{N}$ such that $y = x + m$ and $z = y + n$. But then $z = y + b = x + m + n$, hence $x < z$. This proves O2. □

Proposition 1.16. *For $x, y, z \in \mathbb{N}$, $<$ and $\leq$ have the following properties:*

- *$x \leq x$;*
- *if $x \leq y$ and $y \leq x$ then $x = y$;*
- *either $x \leq y$ or $y \leq x$;*
- *if $x < y$, then $x + z < y + z$ for $z \in \mathbb{N}$.*
- *if $x < y$ then $xz < yz$.*

We shall also need a couple of almost trivial results:

Proposition 1.17. *We have*

1. *There is no $x \in \mathbb{N}$ with $x < 1$;*
2. *$x < s(y)$ if and only if $x \leq y$, where $x, y \in \mathbb{N}$;*
3. *$1 \leq x$ for all $x \in \mathbb{N}$;*
4. *$s(y) \leq x$ if and only if $y < x$, where $x, y \in \mathbb{N}$;*

For a subset $S \in \mathbb{N}$, we say that S has a smallest element if there is an $s \in S$ such that $s \leq t$ for all $t \in S$. The following result is basic (we say that $\mathbb{N}$ is well-ordered):

Theorem 1.18. *Every nonempty subset $S \in \mathbb{N}$ has a smallest element.*

Proof. The idea is quite simple: since S is non-empty, there is an $s_1 \in S$. If $s_1 \leq s$ for all $s \in S$, then s_1 is a smallest element, and we are done; if not, then there is an $s_2 \in S$ such that $s_2 < s_1$. If s_2 is a smallest element, the proof is complete, if not then pick an $s_3 \in S$ with $s_3 < s_2$. After finitely many steps, this process must terminate.

While the last claim may be intuitively clear (and almost obvious), it is quite hard to prove because we are going 'from the top to the bottom', whereas our principal method of proof (induction) goes in the inverse direction. If someone can come up with a simpler proof, I'm all ears.

Let $S \subseteq \mathbb{N}$ be non-empty, and assume for contradiction that S does not contain a smallest element. Define

$$R = \{x \in \mathbb{N} : x < y \text{ for all } y \in S\}.$$

Then $R \subseteq \overline{S}$, where $\overline{S} = \mathbb{N} \setminus S$. In fact, if $x \in R$ were also an element of S, then $x < y$ does not hold for all $y \in S$ because it fails for $y = x$.

We now prove by induction that $R = \mathbb{N}$. This gives the desired contradiction, since $R = \mathbb{N}$ and $R \subseteq \overline{S}$ imply that $\mathbb{N} \subseteq \overline{S}$, hence $S = \varnothing$.

We start by showing that $1 \in R$. We have $1 \leq y$ for all $y \in S$ by Prop. 1.17.3. Now if $1 \in S$, then 1 would be a smallest element in S; since by assumption there is no such beast, we have $1 \notin S$, and thus $1 < y$ for all $y \in S$.

Next we have to show that if $x \in R$, then $s(x) \in R$. So assume that $x \in R$, and let $y \in S$. Then $x < y$, so $s(x) \leq y$ by Prop. 1.17.4. If $s(x) \in S$, then $s(x)$ would be a smallest element in S contrary to our assumption. Hence $s(x) \notin S$, and so $s(x) < y$ for all $y \in S$, thus $s(x) \in R$. □

Let us also prove a simple result that will evolve into the Archimedean property of the reals:

Proposition 1.19. *If $x < y$ are natural numbers, then there exists an $n \in \mathbb{N}$ such that $nx > y$.*

Proof. Put $n = y + 1$. □

Finally we take the opportunity to explain the difference between arbitrarily long and infinitely long sequences:

- there exist arbitrarily long decreasing sequences of natural numbers;
- there do not exist infinitely long decreasing sequences of natural numbers.

The first statement means that we can find sequences of decreasing natural numbers that are as long as we wish: in fact, if we want a sequence of length N, we simply pick N, $N-1$, $N-2$, ..., 2, 1. On the other hand, there clearly are no infinite sequences of this type: if the sequence starts with a number N, then the sequence cannot have more than N elements, and in particular it is finite.

1.2 The Integers $\mathbb{Z}$

The first step in constructing the real numbers from $\mathbb{N}$ is of course the construction of $\mathbb{Z}$, that is, of the (ring of) integers. Clearly mathematicians don't like to mumble something about debts or negative temperatures when defining negative integers. So how do they do it?

Here's how. We can represent every natural number n as a difference of two natural numbers in many ways, e.g. $2 = 3 - 1 = 4 - 2 = 5 - 3 = \ldots$. Thus we can represent 2 by the pairs $(3, 1)$, $(4, 2)$, $(5, 3)$ etc. of natural numbers. If we already knew negative numbers, then of course $-2 = 1 - 3 = 2 - 4 = \ldots$ would be represented by the pairs $(1, 3)$, $(2, 4)$, $(3, 5)$ etc. The idea is now to turn everything around and create negative integers using pairs (m, n) of natural numbers.

We define an equivalence relation on the set

$$W = \{(m, n) : m, n \in \mathbb{N}\}$$

of these pairs by putting $(m, n) = (m', n')$ if $m + n' = n + m'$. This is indeed an equivalence relation:

- Symmetry: $(m, n) \sim (m, n)$;
- Reflexitivity: $(m, n) \sim (m', n') \Longrightarrow (n', m') \sim (m, n)$;
- Transitivity: $(n, m) \sim (n', m')$ and $(n', m') \sim (n'', m'') \Longrightarrow (n, m) \sim (n'', m'')$.

Of course there is something to prove here. DO IT!

Note that $(3, 1) \sim (4, 2) \sim (5, 3) \sim \ldots$.

Let $[m, n] = \{(x, y) : (x, y) \sim (m, n)\}$ denote the equivalence class of (m, n), and let $\mathbb{Z} = \{[m, n] : m, n \in \mathbb{N}\}$ denote the set of all equivalence classes.

We can make $\mathbb{N}$ into a subset of $\mathbb{Z}$ by identifying a natural number n with the equivalence class $[n + 1, 1]$ (in our example, we would identify 2 with $[3, 1]$). Moreover, we shall simply write $-n$ for the class $[1, 1 + n]$ and introduce the symbol $0 = [1, 1]$.

Remark. At this point I regret that I have not defined $\mathbb{N}$ as $\mathbb{N} = \{0, 1, 2, \ldots\}$ (simply replace 1 in the Peano axioms by 0): if I had, then we could identify natural numbers n with the integers $[n, 0]$ instead of $[n + 1, 1]$; this seems more natural, and it would make several formulas below considerably simpler.

Remark. There are many definitions of what mathematics is (or should be), each of them emphasizing different aspects. The definition that describes what we are doing here is

> Mathematics is the art of identifying different things.

This 'identification' can be given a precise mathematical formulation by introducing the map $\iota : \mathbb{N} \longrightarrow \mathbb{Z}$ that identifies $\mathbb{N}$ with a subset of $\mathbb{Z}$: we put

$\iota(n) = [n+1, 1]$.

$$\begin{array}{ccccc}
\mathbb{N} & \longrightarrow & \mathbb{Z} & & \\
\vdots & \longmapsto & \vdots & & \\
3 & \longmapsto & [4,1] & \simeq & 3 \\
2 & \longmapsto & [3,1] & \simeq & 2 \\
1 & \longmapsto & [2,1] & \simeq & 1 \\
 & & [1,1] & \simeq & 0 \\
 & & [1,2] & \simeq & -1 \\
 & & [1,3] & \simeq & -2 \\
 & & \vdots & &
\end{array}$$

Now we only are allowed to 'identify' $\mathbb{N}$ with its image $\iota(\mathbb{N}) \subseteq \mathbb{Z}$ if ι does not map two natural numbers to the same integer, that is, if ι is injective. Let's check this:

Assume that $\iota(m) = \iota(n)$, i.e., that $[m+1, 1] = [n+1, 1]$. By definition of these equivalence classes this means that $(m+1, 1) \sim (n+1, 1)$, that is, $m+1+1 = n+1+1$. The cancellation law then implies $m = n$, hence ι is injective.

We remark that

$$\mathbb{Z} = -\mathbb{N} \cup \{0\} \cup \mathbb{N} = \{\ldots -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

To this end, we have to prove that for every $(m, n) \in W$ there is an $x \in \mathbb{N}$ such that $(m, n) \sim (x, 1)$ or $(m, n) \sim (1, x)$, because this implies $[m, n] = x \in \mathbb{N}$ (if $(m, n) \sim (x, 1)$ and $x > 1$) or $[m, n] = 0$ (if $(m, n) \sim (x, 1)$ and $x = 1$) or $[m, n] = -x$ (if $(m, n) \sim (1, x)$ and $x > 1$). This follows from the Trichotomy Law.

We also mention that $\mathbb{Z}$ is countable. Recall that a set A is said to be countable if there exists a bijection between A and $\mathbb{N}$. The fact that $\iota : \mathbb{N} \hookrightarrow \mathbb{Z}$ is not a bijection does not imply that $\mathbb{Z}$ is not countable: it simply means that ι is not bijective, so we have to look for another map.

This isn't hard: the sequence 0, 1, -1, 2, -2, etc. contains every integer once, hence $0 \longleftrightarrow 1$, $1 \longleftrightarrow 2$, $-1 \longleftrightarrow 3$, $2 \longleftrightarrow 4$, $-2 \longleftrightarrow 5$ etc. gives us the desired bijection $\mathbb{Z} \longleftrightarrow \mathbb{N}$. We could write down a formula if we wanted to, but this would only complicate things.

Addition

We now show that we can define addition, multiplication and an ordering $<$ on $\mathbb{Z}$ in such a way that the properties of $\mathbb{N}$ proved in the last section continue to hold.

We start by defining addition $\oplus$ on $\mathbb{Z}$. We have to say what $[r, s] \oplus [t, u]$ is supposed to be. Clearly we would like to have $[r, s] = r - s$, $[t, u] = t - u$, so the sum should be $r - s + t - u = r + t - (s + u) = [r + t, s + u]$. With this idea in mind we now define

$$[r, s] \oplus [t, u] = [r + t, s + u], \tag{1.7}$$

where the addition inside the brackets is the addition in $\mathbb{N}$.

There is no need to prove that we have defined addition on all elements of $\mathbb{Z}$, because this is obvious here; the point is that we did not define $\oplus$ inductively, we simply wrote down a formula.

Nevertheless there's a lot of work to do. First we have to prove that this addition is well defined (this is something that comes up whenever we define something on equivalence classes). What this means is: assume that $[r, s] = [r', s']$ and $[t, u] = [t', u']$. On the one hand, we have

$$[r, s] \oplus [t, u] = [r + t, s + u].$$

If we replace the left hand side by $[r', s'] \oplus [t', u']$, then we clearly get

$$[r', s'] \oplus [t', u'] = [r' + t', s' + u'].$$

But if our addition is to make any sense, then the right hand sides should be equal because, after all, the left hand sides are. Thus we want to show that

$$[r' + t', s' + u'] = [r + t, s + t]. \tag{1.8}$$

What do we know? We know that $[r, s] = [r', s']$, which by definition means $(r, s) \sim (r', s')$, that is, $r + s' = s + r'$. Similarly, $[t, u] = [t', u']$ implies $t + u' = u + t'$. Adding these equations and using commutativity and associativity for natural numbers we get $r+t+s'+u' = s+u+r'+t'$, which in turn is equivalent to (1.8).

Next we have to show that the two additions agree on $\mathbb{N}$; after all, we are using the very same symbols for natural numbers 1, 2, ... and their images 1, 2, ... under ι in $\mathbb{Z}$. This can only work if, for natural numbers m, n, the sum $m + n$ is the same whether evaluated in $\mathbb{N}$ or in $\mathbb{Z}$. In other words: we want to be sure that

$$\iota(m + n) = \iota(m) \oplus \iota(n).$$

This is a straight forward computation:

$$\begin{array}{rcll} \iota(m) \oplus \iota(n) & = & [m + 1, 1] \oplus [n + 1, 1] & \text{by definition of } \iota \\ & = & [m + n + 2, 2] & \text{by (1.7)} \\ & = & [m + n + 1, 1] & (m + n + 2, 2) \sim (m + n + 1, 1) \\ & = & \iota(m + n) & \text{by definition of } \iota \end{array}$$

Now that there is no need to distinguish between the two types of addition anymore, we shall often write $+$ instead of $\oplus$ for the addition on $\mathbb{Z}$.

Of course we have prove that associativity and commutativity also holds for our addition in $\mathbb{Z}$. So why is $(x + y) + z = x + (y + z)$ for all $x, y, z \in \mathbb{Z}$? Write $x = [r, s]$, $y = [t, u]$ and $z = [v, w]$ with $r, s, t, u, v, w \in \mathbb{N}$. Then $(x + y) + z = [r + t, s + u] + [v, w] = [(r + t) + v, (s + u) + w]$, and similarly $x + (y + z) = [r + (t + v), s + (u + w)]$. Because addition in $\mathbb{N}$ is associative, so is addition in $\mathbb{Z}$ (again, observe that there's no need for invoking induction here).

Exercise. Prove that addition on $\mathbb{Z}$ is commutative.

As we know, we also can subtract integers in $\mathbb{Z}$. So how do we define $x - y$? Well we write $x = [r, s]$ and $y = [t, u]$; we cannot put $x - y = [r - t, s - u]$ because $r - t$ and $s - u$ might not be natural numbers; but if they were, we would have $[r - t, s - u] = [r + u, s + t]$, and nothing prevents us from defining

$$[r, s] \ominus [t, u] = [r, s] \oplus [u, t] = [r + u, s + t]. \tag{1.9}$$

Again, we have to show that it is well defined. Once this is done, it is easy to prove that $\mathbb{Z}$ is a group with respect to addition, and that $0 = [1, 1]$ is the neutral element.

What does that mean? A group is a set G of elements together with a composition, that is, a map $+ : G \times G \longrightarrow G$ that maps a pair of elements $(g, g') \in G \times G$ to another element $g + g' \in G$; we also demand that this composition satisfy the following rules:

G1 there is a neutral element $0 \in G$ such that $g + 0 = g$ for all $g \in G$;

G2 for every $g \in G$ there is an element $g' \in G$ such that $g + g' = 0$ (we shall write $g' = -g$);

G3 the composition is associative: we have $(g + g') + g'' = g + (g' + g'')$ for all $g, g', g'' \in G$.

If the group also satisfies the condition

G4 $g + g' = g' + g$ for all $g, g' \in G$;

then we say that G is commutative (abelian).

The set $\mathbb{N}$ of natural numbers is not a group with respect to $+$: there is a composition $+ : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$, but there is no neutral element (let alone inverses).

The set $\mathbb{Z}$ of integers, on the other hand, is a group with respect to $+$ (CHECK IT!). In fact, $\mathbb{Z}$ is not only a group, it also carries the structure of a ring. But to see this, we have to define multiplication in $\mathbb{Z}$ first.

Multiplication

We want to define $[r, s] \odot [t, u]$ in such a way that multiplication of natural numbers $\mathbb{N} \subset \mathbb{Z}$ agrees with the one we have defined before, namely that $\iota(m) \odot \iota(n) = \iota(mn)$. In other words, our multiplication should give

$$[m + 1, 1] \odot [n + 1, 1] = [mn + 1, 1].$$

More generally, if, for a minute, we think of $[r, s]$ as the 'integer' $r - s$, then we want $[r, s] \odot [t, u] \simeq (r - s)(t - u) = rt + su - ru - st \simeq [rt + su, ru + st]$, and this suggests the definition

$$[r, s] \odot [t, u] = [rt + su, ru + st]. \tag{1.10}$$

Exercise. Show that the operation $[r,s] * [t,u] = [rt, su]$ is not well defined.

Once more we have to show that the multiplication (1.10) is well defined and that it agrees with multiplication in $\mathbb{N}$ (actually we have defined it in such a way that it must; what we have to check here is that $\iota(m) \odot \iota(n) = \iota(mn)$). Then one generalizes distributivity, commutativity, associativity and the cancellation law to integers in $\mathbb{Z}$.

Let us just note in passing that

$$\begin{array}{rcll} (-1)\cdot(-1) & = & [1,2]\odot[1,2] & \text{by our identification} \\ & = & [5,4] & \text{by (1.10)} \\ & = & [2,1] & \text{since } (5,4)\sim(2,1) \\ & = & +1 & \text{since } \iota(1)=[2,1] \end{array}$$

In fact, for $m, n \in \mathbb{Z}$ it is now easy to prove that

$$\begin{array}{rcl} (-m)\cdot \quad n & = & -mn, \\ m\cdot(-n) & = & -mn, \\ (-m)\cdot(-n) & = & mn. \end{array}$$

Thus the rules for multiplying signs come out naturally from our definition of multiplication on $\mathbb{Z}$, which in turn was forced upon us as the simplest way of extending multiplication on $\mathbb{N}$.

Now we are ready to state that $\mathbb{Z}$ is a ring. A ring R is a set on which two kinds of compositions are defined; they are usually denoted by $+$ and $\cdot$. Of course, these compositions are to satisfy certain conditions; first of all, $r+s$ and $r\cdot s$ should be elements of R whenever r and s are. Moreover, we demand

R1 R is an abelian group with respect to $+$;

R2 (associativity): $r(st) = (rs)t$ for $r, s, t \in R$;

R3 (distributivity): we have $r(s+t) = rs + rt$ and $(r+s)t = rt + st$ for $r, s, t \in R$.

R4 R contains a unit element e satisfying $er = re = r$ for all $r \in R$;

If R also satisfies $rs = sr$ for all $r, s \in R$, then we say that R is commutative. Finally, a ring R is called a domain if $xy = 0$ implies $x = 0$ or $y = 0$.

In any ring we have $0x = 0$: in fact,

$$\begin{array}{rcll} 0x & = & (0+0)x & \text{since 0 neutral element of } + \\ & = & 0x + 0x & \text{by distributivity,} \end{array}$$

so substracting $0x$ from both sides gives $0 = 0x$.

Theorem 1.20. *The integers $\mathbb{Z}$ form a commutative domain with respect to addition and multiplication.*

Let us prove that $\mathbb{Z}$ is indeed a domain b

Assume that $xy = 0$ for $x, y \in \mathbb{Z}$. Write $x = [r, s]$ and $y = [t, u]$. Then $[1, 1] = 0 = xy = [r, s] \odot [t, u] = [rt+su, ru+st]$ by assumption, hence $rt+su+1 = ru + st + 1$, and by the cancellation law, $rt + su = ru + st$.

Now assume that $x \neq 0$; then $r + m = s$ or $r = s + m$ for some $m \in \mathbb{N}$ by Theorem 1.8. In the first case, $r+m = s$ for some $m \in \mathbb{N}$. Then $rt+(r+m)u = rt + su = ru + st = ru + (r + m)t$, hence $mu = mt$ and so $u = t$, that is, $y = 0$. The case $r > s$ is treated similarly.

$\mathbb{Z}$ as an ordered domain

Last not least we have to extend the relation $<$ to $\mathbb{Z}$. We put

$$[r, s] < [t, u] \quad \text{if} \quad r + u < t + s. \tag{1.11}$$

This is well defined and agrees with the ordering on $\mathbb{N}$.

For showing that the relation is well defined, we have to assume that $(r, s) \sim (r', s')$ and $(t, u) \sim (t', u')$, and then show that $r+u < t+s$ implies $r'+u' < t'+s'$. DO IT.

For showing that the order just defined agrees with the one we know from $\mathbb{N}$ we have to prove that $n < m$ if and only if $\iota(n) < \iota(m)$.

Proposition 1.21. *The set $\mathbb{Z}$ is simply ordered with respect to $<$.*

An ordered domain is a domain R together with an order $<$ such that

OD1 R is simply ordered with respect to $<$.

OD2 If $x < y$, then $x + z < y + z$ for $x, y, z \in R$.

OD3 If $x < y$ and $0 < z$, then $xz < yz$.

Proposition 1.22. *$\mathbb{Z}$ is an ordered domain with respect to $<$.*

Proof. Write $x = [r, s]$ and $y = [t, u]$. If $z \in \mathbb{N}$, then we may put $z = [v, 1]$. Now $x < y$ means $r + u < t + s$, and $xz < yz$ is equivalent to $(r + u)v < (t + s)v$. Now look back at Prop. 1.16.

The rest is left as an exercise. □

Proposition 1.23. *In any ordered ring R, the following assertions are true:*

1. *If $x < 0$, then $-x > 0$.*
2. *If $x < y$ and $z < 0$, then $xz > yz$.*
3. *We have $x^2 \geq 0$ for all $x \in R$, with equality if and only if $x = 0$.*

Proof. 1. If $x < 0$, then $x + (-x) < 0 + (-x)$ by OD2, and so $0 < -x$.

2. We have $0 < -z$, hence $-xz = x \cdot (-x) < y \cdot (-z) = -yz$, hence $xz > yz$.

3. If $x \geq 0$, then multiplying through by $x \geq 0$ gives $x^2 \geq= 0$; if $x \leq 0$, then multiplying through by $x \leq 0$ gives $x^2 \geq 0$.

□

We now introduce absolute values in any ordered domain by putting

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

Here are a few simple properties of

Lemma 1.24. *In any ordered domain R, The absolute value $|\cdot|$ has the following properties.*

1. $|x| \geq 0$*: this is clear if $x \geq 0$; if $x < 0$, multiply through by $-1 < 0$.*

2. $|xy| = |x| \cdot |y|$*: Just consider the four possible cases: 1) if $x > 0$, $y > 0$, then $xy > 0$, so the claim is $xy = xy$, which obviously holds; 2) if $x > 0$ and $y < 0$, then $xy < 0$, hence the claim is $-xy = x \cdot (-y)$. The other two cases are treated similarly.*

3. *If $s \geq 0$ and $-s \leq r \leq s$, then $|r| \leq s$. In fact, it is sufficient to prove that $r \leq s$ and $-r \leq s$. The first one is true by assumption, the second one follows from multiplying $-s \leq r$ through by -1.*

Proposition 1.25 (Triangle Inequality). *For all x, y in an ordered domain, we have $|x + y| \leq |x| + |y|$.*

Proof. By adding $-|x| \leq x \leq |x|$ and $-|y| \leq y \leq |y|$ we obtain $-(|x| + |y|) \leq x + y \leq |x| + |y|$. Now apply Lemma 1.24.3 to $r = x + y$ and $s = |x| + |y|$. □

1.3 The Rationals $\mathbb{Q}$

Just as mathematicians avoid thinking of debts when introducing negative integers, they don't think of splitting pies when introducing rationals. Instead, they play the game of constructing them using – pairs of integers, of course.

In this case, we want to identify a fraction $\frac{a}{b}$ with the equivalence class $[a, b]$, where $a, b \in \mathbb{Z}$ are integers. Since we want $\frac{a}{b} = \frac{c}{d}$ whenever $ad = bc$, we start by defining an equivalence relation on the set

$$V = \{(r, s) : r \in \mathbb{Z}, s \in \mathbb{N}\}$$

of pairs of an integer and a natural number, and of course we put

$$(r, s) \sim (t, u) \iff ru = st.$$

It is easily seen that this is an equivalence relation (DO IT!), and we now put

$$[r, s] = \{(x, y) \in V : (x, y) \sim (r, s)\}.$$

Let $\mathbb{Q}$ denote the set of all equivalence classes $[r, s]$ with $(r, s) \in V$; a class $[r, s]$ is called a rational number, and we often write $\frac{r}{s}$ instead of $[r, s]$.

Remark. This should be no problem at all for computer scientists: if they were to write a program that does arithmetic with rational numbers, they would realize $\frac{r}{s}$ as a pair $[r, s]$ of integers, identify $[r, s] = [nr, ns]$ for $n \in \mathbb{N}$, and define sums and products of such pairs in exactly the same way as we are doing here.

First we identify an integer $z \in \mathbb{Z}$ with the rational number $[z, 1]$, that is, we consider the map $\iota : \mathbb{Z} \longrightarrow \mathbb{Q}$ defined by $\iota(z) = [z, 1]$ and first prove that ι is injective (that is, no two different integers correspond to the same rational number). In fact, assume that $x, y \in \mathbb{Z}$ are such that $\iota(x) = \iota(y)$. Then $[x, 1] = [y, 1]$, i.e., $(x, 1) \sim (y, 1)$, and by definition of equivalence in V this means $x \cdot 1 = y \cdot 1$, hence $x = y$.

We want to have $\frac{r}{s} + \frac{t}{u} = \frac{ru+st}{su}$, so we are led to define

$$[r, s] \oplus [t, u] = [ru + st, su]. \tag{1.12}$$

This is well defined (!), and agrees with addition on $\mathbb{Z}$ under the identification ι: in fact,

$$\begin{aligned} \iota(x) \oplus \iota(y) &= [x, 1] \oplus [y, 1] \\ &= [x \cdot 1 + 1 \cdot y, 1 \cdot 1] = [x + y, 1] \\ &= \iota(x + y). \end{aligned}$$

Thus it does not matter whether we add in $\mathbb{Z}$ and then identify the result with a rational number, or first view the integers as elements of $\mathbb{Q}$ and add there.

Next we define multiplication of rationals by

$$[r, s] \odot [t, u] = [rt, su]. \tag{1.13}$$

Here we were motivated by $\frac{r}{s} \cdot \frac{t}{u} = \frac{rt}{su}$. Again, this is well defined and agrees with multiplication of integers on the subset $\mathbb{Z} \subset \mathbb{Q}$: we have $\iota(x) \odot \iota(y) = \iota(xy)$ because

$$\begin{aligned} \iota(x) \odot \iota(y) &= [x, 1] \odot [y, 1] && \text{by definition of } \iota \\ &= [xy, 1] && \text{by definition (1.13)} \\ &= \iota(xy) && \text{by definition of } \iota \end{aligned}$$

Remark. The map $\iota : \mathbb{Z} \longrightarrow \mathbb{Q}$ from the ring of integers to the ring of rational numbers satisfies

$$\begin{aligned} \iota(x) \oplus \iota(y) &= \iota(x + y), \\ \iota(x) \odot \iota(y) &= \iota(xy), \end{aligned}$$

Maps $R \longrightarrow S$ between rings with these properties (we say that they 'respect the ring structure') are called ring homomorphisms if they map the unit element of R to the unit element of S. In particular, our 'identification map' ι is a ring homomorphism.

Using these definitions, we can prove associativity, commutativity, distributivity, thereby verifying that $\mathbb{Q}$ is a ring. In fact, $\mathbb{Q}$ is even a field!

A field F is a commutative ring in which we can divide by nonzero elements: thus F is a field if F satsifies the ring axioms, and if in addition

F1 For every $r \in F \setminus \{0\}$ there is an $s \in F$ such that $rs = 1$.

This is a strong axiom: together with some other ring axioms it implies that fields are domains, in other words: if $xy = 0$ for field elements $x, y \in F$, then $x = 0$ or $y = 0$. In fact, if $y \neq 0$, then there is some $z \in F$ such that $yz = 1$ by F1, hence $0 = 0z = (xy)z = x(yz) = x \cdot 1 = x$.

One of the most important properties of fields F is

Proposition 1.26. *If F is a field and if $xy = 0$ for $x, y \in F$, then $x = 0$ or $y = 0$.*

Proof. In fact, assume that $xy = 0$ and $y \neq 0$. Since the nonzero elements of F form a group, y has an inverse, that is, there is a $z \in F$ such that $yz = 1$. But now $0 = xy$ implies $0 = 0z = (xy)z = x(yz) = x \cdot 1 = x$; here we have used associativity of multiplication. □

Theorem 1.27. *The set $\mathbb{Q}$ of rational numbers forms a field with respect to addition and multiplication.*

Starting from $\mathbb{N}$, we have step by step constructed larger domains: we introduced $\mathbb{Z}$ in order to be able to subtract, and we went over to $\mathbb{Q}$ in order to be able to divide by nonzero elements. Now in $\mathbb{Q}$ we can add, subtract, multiply and divide by nonzero elements, so why should we want to extend $\mathbb{Q}$ again? One reason is that there 'seem to exist' numbers that are not rational:

Proposition 1.28. *There is no $x \in \mathbb{Q}$ with $x^2 = 2$.*

Proof. Assume that $x^2 = 2$ for $x = \frac{m}{n}$. Assume that m and n are not both even (if they are, we can cancel factors of 2 until one of them is odd). Multiplying through by n^2 gives $2n^2 = m^2$. Since the left hand side is even so is the right hand side, and this implies that m is even, say $m = 2p$. Then $2n^2 = 4p^2$, hence $n^2 = 2p^2$. This in turn shows that n is even: contradiction, since we assumed that m or n is odd. □

There is another proof that uses less number theory and is much more powerful:

Theorem 1.29. *If $n \in \mathbb{N}$ is not the square of an integer, then it is not the square of a rational number.*

Proof. In fact, if n is not a square of an integer, then it lies between two squares, that is, we can find an integer a such that $a^2 < n < (a+1)^2$. Assume that $\sqrt{n} = \frac{p}{q}$ with $q > 0$ minimal. Then $p^2 = nq^2$, hence $p(p - aq) = p^2 - apq = nq^2 - apq = q(nq - ap)$, so

$$\frac{p}{q} = \frac{nq - ap}{p - aq}.$$

But $a < \frac{p}{q} < a+1$ implies $0 < p - aq < q$: this contradicts the minimality of the denominator q. $\square$

Observe the difference between the statement '2 is not a square in $\mathbb{Q}$' and '$\sqrt{2}$ is an irrational number': the second statement only makes sense if we already know what $\sqrt{2}$ means.

Countability of $\mathbb{Q}$

Although there are 'many more' rational numbers than integers, in some way their cardinality is the same:

Theorem 1.30. *The set $\mathbb{Q}$ is countable.*

Proof. We first show that the positive rational numbers are countable. To this end we consider the lattice of numbers

$$\begin{array}{ccccc} \frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots \\ \frac{2}{1} & \frac{2}{2} & \frac{2}{3} & \frac{2}{4} & \cdots \\ \frac{3}{1} & \frac{3}{2} & \frac{3}{3} & \frac{3}{4} & \cdots \\ \frac{4}{1} & \frac{4}{2} & \frac{4}{3} & \frac{4}{4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Clearly, every positive rational number is contained in here. Now we form a sequence by starting at the left top like this:

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

that is, we enumerate the fractions $\frac{m}{n}$ according to growing $m+n$. Now here we have counted several rational numbers twice (as a matter of fact, infinitely often), because $\frac{1}{1} = \frac{2}{2} = \frac{3}{3} = \ldots$, $\frac{1}{2} = \frac{2}{4} = \frac{3}{6} = \ldots$ etc. But if we omit every rational number that has appeared before then we get a sequence that contains every positive rational number exactly once:

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

Now we have to take care of 0 and negatives. But this is easy: put 0 at the beginning, and after each positive rational number squeeze in its negative:

$$0, \frac{1}{1}, -\frac{1}{1}, \frac{1}{2}, -\frac{1}{2}, \frac{2}{1}, -\frac{2}{1}, \ldots$$

This sequence contains each rational number exactly once, and our claim is proved. $\square$

Part of the above arguments can be formalized easily. For example, if you look at our proof carefully, you should be able to extract a proof of

Proposition 1.31. *If A is a subset of a countable set B, then A is countable.*

Hilbert's Hotel.

The intuitive problems with infinite countable sets become apparent in Hilbert's Hotel. This is a hotel with infinitely many rooms: one room for each natural number $1, 2, 3, \ldots$. Assume that all rooms are taken, and that a new guest arrives. There's nothing you can do in a hotel with finitely many rooms, but the manager of Hilbert's hotel simply tells the guest in room n to go to room $s(n) = n + 1$ and gives the room with #1 to the new guest. Similarly, if m guests arrive, he sends the guest in room n to room $n + m$ and gives the m guests the rooms 1, 2, $\ldots$, m.

It is hard to imagine how many space there is in Hilbert's hotel even if there are no vacancies: assume that Hilbert's bus arrives, bringing infinitely but countably many new guests. The trick above doesn't work, but the manager saves the day by sending the guest in room n to room $2n$, and then giving the room $2n - 1$ to the new guest with number n.

Proposition 1.32. *If A and B are countable sets, then so is $A \cup B$.*

Proof. Exercise. (Hint: think of the elements of A as guests at Hilbert's Hotel, and of the elements of B as newly arrived people in Hilbert's Bus). □

$\mathbb{Q}$ as an ordered set

We define an order relation $<$ on $\mathbb{Q}$ by putting

$$[r, s] < [t, u] \iff ru < st$$

(recall that $s, u \in \mathbb{N}$). This is well defined: if $[r, s] = [r', s']$ and $[t, u] = [t', u']$, then $rs' = r's$ and $tu' = t'u$. Now

$$\begin{array}{rcll} [r, s] < [t, u] & \iff & ru < st & \text{by definition} \\ & \iff & rus'u' < sts'u' & \text{since } s'u' > 0 \\ & \iff & r'suu' < ss't'u & \text{since } rs' = r's \text{ and } tu' = t'u \\ & \iff & r'u' < s't' & \text{since } su > 0 \\ & \iff & [r', s'] < [t', u'] & \text{by definition} \end{array}$$

Now we have

Theorem 1.33. *$\mathbb{Q}$ is an ordered domain (even field).*

Proof. Since exactly one of the relations $ru < st$, $ru = st$ or $ru > st$ is true by the trichotomy law for integers, exactly one of $x < y$, $x = y$ or $x > y$ is true, where $x = [r, s]$ and $y = [t, u]$.

Next assume that $x < y$ and $y < z$, where $z = [v, w]$. Then $ru < st$ and $tw < uv$, hence $ruw < stw$ and $stw < suv$ since $w, s > 0$; transitivity for the integers gives $ruw < suv$, and since $u > 0$, this is equivalent to $rw < sv$, i.e., $x < z$.

This shows that $\mathbb{Q}$ is simply ordered. The rest of the proof that $\mathbb{Q}$ is an ordered domain is left as an exercise. □

Thus everything proved for general ordered domains holds for the rationals; in particular, $x^2 \geq 0$ for all $x \in \mathbb{Q}$, and $|x+y| \leq |x| + |y|$ for $x, y \in \mathbb{Q}$.

Now let us collect a few simple results that will turn out to be useful.

Lemma 1.34. *We have $|x| < |y|$ if and only if $n|x| < n|y|$ for some $n \in \mathbb{N}$.*

Proof. Exercise. □

Proposition 1.35. *If $x, y \in \mathbb{Z}$ satisfy $|x - y| < 1$, then $x = y$.*

Proof. If $x \neq y$, then $x = y + z$ for some nonzero $z \in \mathbb{Z}$. This means $|z| \in \mathbb{N}$, hence $|z| \geq 1$, and we see that $|x - y| = |z| \geq 1$. □

Proposition 1.36. *Let $x, y \in \mathbb{Q}$ and assume that for every rational $\varepsilon > 0$ we have $|x - y| < \varepsilon$; then $x = y$.*

Proof. Assume that this is false, i.e. that $x - y \neq 0$. Then $\varepsilon = |x - y|$ is a positive rational number, so by assumption we have $|x - y| < \varepsilon$. This implies $\varepsilon < \varepsilon$, which is a contradiction. □

Proposition 1.37. *Let $0 < x < y$ be rational numbers. Then there is an $n \in \mathbb{N}$ such that $nx > y$.*

Proof. Write $x = \frac{r}{s}$ and $y = \frac{s}{t}$ with $r, s, t, u \in \mathbb{N}$ (here we have used that $x, y > 0$). Then $x < y$ is equivalent to $ru < st$; by Prop. 1.19 there is an $n \in N$ such that $n(ru) > st$. But the last inequality is equivalent to $nx > y$. □

Chapter 2

Cauchy Sequences in $\mathbb{Q}$

In this chapter, ε will always denote a *rational* number > 0.

2.1 Sequences in $\mathbb{Q}$

A sequence (s_n) in $\mathbb{Q}$ is a sequence $s_1, s_2, s_3, \ldots$ of elements $s_i \in \mathbb{Q}$. In mathematical terms, we could define it as a function $s : \mathbb{N} \longrightarrow \mathbb{Q}$ assigning a rational number s_n to each natural number n.

If s is a rational number such that the elements s_n get as close as you want to s, then we say that s_n tends to s, or that s is the limit of the sequence s. Here's the correct definition; it does away with all the dynamics of 'tending' or 'getting closer', and replaces it by something static.

Definition.[1] A sequence (s_n) in $\mathbb{Q}$ is said to converge to a rational number $s \in \mathbb{Q}$ if the following condition is satisfied:

For all $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that $|s_n - s| < \varepsilon$ for all $n > N$.

If this is true, then we write $s = \lim_{n\to\infty} s_n$ or simply $s = \lim s_n$.

Remark. What the definition says is that s_n converges to s if we can make the difference $|s_n - s|$ as small as we want if we only choose n large enough.

Remark. A definition equivalent to the above was given already in 1665 by John Wallis, a contemporary of Pièrre Fermat.

As for proofs, you should convince yourself of the following facts:

1. If you want to use the fact that s_n converges, then you pick an $\varepsilon > 0$; the fact that s_n converges provides you with an N such that $|s_n - s| < \varepsilon$ for $n > N$.

2. If you have to prove that s_n converges to s, you are given an $\varepsilon > 0$, and you have to find an $N \in \mathbb{N}$ such that $|s_n - s| < \varepsilon$ for all $n > N$.

[1] Ross, Def. 7.1.

Go through the proofs below to see what this remark means.

Example 1. Consider the sequence $s_n = \frac{1}{n}$. Then the elements of the sequence get closer and closer to 0. As a matter of fact, they also get closer and closer to -1; but while the difference between $\frac{1}{n}$ and 0 can be made as small as you want by choosing n sufficiently large, the difference between $\frac{1}{n}$ and -1 is always larger than 1.

In fact, we claim that $\lim \frac{1}{n} = 0$. For a proof, we have to show that for every $\varepsilon > 0$ we can find a natural number N with the property that $s_n = |s_n - 0| < \varepsilon$ for all $n > N$. Since $s_n = \frac{1}{n}$, we have to check that $\frac{1}{n} < \varepsilon$, i.e., that $n > 1/\varepsilon$. If we choose N as any natural number larger than $1/\varepsilon$, then $n > 1/\varepsilon$ for all $n > N$.

Example 2. A sequence (s_n) of integers converges if and only if there is an $N \in \mathbb{N}$ and an $s \in \mathbb{Z}$ such that $s_n = s$ for all $n > N$. In this case, $\lim s_n = s$.

This shows that the notion of limits is not very important for sequences of integers. Let's now prove our claim.

Assume first that s_n converges. We pick $\varepsilon = \frac{1}{2}$; by definition of convegence there is an $s \in \mathbb{Q}$ and an $N \in \mathbb{N}$ such that $|s_n - s| < \varepsilon$ for all $n > N$. Let $m > N$ be another integer; then $|s_n - s_m| = |s_n - s + s - s_m| \leq |s_n - s| + |s_m - s| < 2\varepsilon = 1$. But $s_n, s_m \in \mathbb{Z}$, and if two integers differ by less than 1, then they are equal. This proves that $s_n = s_m =: t$ for some $t \in \mathbb{Z}$ and all $m, n > N$. Since we clearly have $\lim s_n = t$, the uniqueness of limits proves that s must be an integer.

Now assume that there is an $N \in \mathbb{N}$ and an $s \in \mathbb{Z}$ such that $s_n = s$ for all $n > N$. Then we claim that (s_n) converges (to s). In fact, given any $\varepsilon > 0$, we have $|s_n - s| = 0 < \varepsilon$ for all $n > N$.

Here are some simple properties of the limit.

Proposition 2.1. [2] *Limits are unique: if* $\lim s_n = s$ *and* $\lim s_n = t$*, then* $s = t$.

Proof. Assume that $s \neq t$ and put $\varepsilon = |\frac{s-t}{2}|$. Since $\lim s_n = s$, there is an $N_1 \in N$ such that $|s_n - s| < \varepsilon$ for all $n > N_1$. Similarly, there is an $N_2 \in \mathbb{N}$ such that $|s_n - t| < \varepsilon$ for all $n > N_2$. Let $N = \max\{N_1, N_2\}$. Then $|s_n - s| < \varepsilon$ and $|s_n - t| < \varepsilon$ for all $n > N$. But now

$$\begin{aligned} 2\varepsilon &= |s - t| = |s - s_n + s_n - t| \\ &\leq |s_n - s| + |s_n - t| < \varepsilon + \varepsilon = 2\varepsilon, \end{aligned}$$

hence $2 < 2$, which is a contradiction. □

Proposition 2.2. *If* (s_n) *converges, then so does the sequence* (a_n) *defined by* $a_1 = s_1$ *and* $a_{n+1} = s_{n+1} - s_n$ *for* $n \geq 2$*, and we have* $\lim a_n = 0$.

Proof. Given $\varepsilon > 0$ we have to find an $N \in \mathbb{N}$ such that $|a_n| = |a_n - 0| < \varepsilon$ for all $n > N$. Our only chance of getting such an N is using what we know

[2] Ross, p. 27.

about (s_n). So how are a_n and s_n related? Well, $|a_{n+1}| = |s_{n+1} - s_n|$. We have to make this small, but all we know is that $|s_n - s|$ is small for large n, where $s = \lim s_n$. Here is where you learn to appreciate the power of the triangle inqequality:

$$|s_{n+1} - s_n| = |s_{n+1} - s + s - s_n| \leq |s_{n+1} - s| + |s_n - s|.$$

But the last two terms can be made small, so we have won!

Here's the formal proof: We know that there is an $N \in \mathbb{N}$ such that $|s_n - s| < \varepsilon/2$ for all $n > N$. Then

$$|a_{n+1} - 0| = |s_{n+1} - s_n| \leq |s_{n+1} - s| + |s_n - s| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $n > N$, hence $\lim a_n = 0$. $\square$

Beginners often feel that the converse of this proposition should also be true, i.e., that if the s_n differ less and less from each other, then the sequence (s_n) should converge. But: Beware of the harmonic series!

In fact, consider the sequence $s_1 = 1$, $s_2 = 1 + \frac{1}{2}$, ..., $s_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$, ... Here we have $a_n = s_{n+1} - s_n = \frac{1}{n+1}$, so clearly $\lim a_n = 0$. Nevertheless, (s_n) does not converge. In fact,

$$\begin{aligned}
\frac{1}{3} + \frac{1}{4} &> \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \\
\frac{1}{5} + \ldots + \frac{1}{8} &> 4 \cdot \frac{1}{8} = \frac{1}{2}, \\
\frac{1}{9} + \ldots + \frac{1}{16} &> 8 \cdot \frac{1}{16} = \frac{1}{2}, \\
&\ldots \\
\frac{1}{2^n + 1} + \ldots + \frac{1}{2^{n+1}} &> 2^n \frac{1}{2^{n+1}} = \frac{1}{2}, \\
&\ldots
\end{aligned}$$

Thus $s_4 > 1 + \frac{1}{2} + \frac{1}{2} = 2$, $s_8 > 1 + 3 \cdot \frac{1}{2} = \frac{5}{2}$, $s_{16} > 1 + 4 \cdot \frac{1}{2} = 3$, ..., and in general $s_{2^n} > 1 + n \cdot \frac{1}{2} = \frac{n+2}{2}$. Thus the sequence s_n is not bounded, and we shall see below (Prop. 4) that such sequences never converge.

Nevertheless the result Prop. 4 is very useful for guessing limits: see the example at the beginning of Section 2.2.

Theorem 2.3. [3] *Assume that the sequences s_n and t_n converge. Then the sequences $s_n \pm t_n$ and $s_n t_n$ converge, and we have* $\lim(s_n \pm t_n) = \lim s_n \pm \lim t_n$ *and* $\lim(s_n t_n) = (\lim s_n)(\lim t_n)$.

Remark. This shows that the set C of converging sequences is a ring: if s_n and t_n converge, then so do $s_n \pm t_n$ and $s_n t_n$. The neutral element with respect to addition is the sequence defined by $s_n = 0$, the unit element is the sequence 1,

[3] Ross, 9.3, 9.4.

1, 1, ... defined by $s_n = 1$. Converging sequences do not form a field, however: we cannot divide by sequences s_n containing a 0.

The map lim is a map from C to the ring (even field) $\mathbb{Q}$, and Theorem 2.3 says that lim is a ring homomorphism. A map $f : R \longrightarrow S$ between rings is a ring homomorphism if it respects the structure, that is, if $f(r+r') = f(r)+f(r')$ and $f(rr') = f(r)f(r')$ for all $r, r' \in R$, and if it maps the identity of R to the identity of S. The unit element of C is the sequence $1, 1, 1, 1, \ldots$, and lim maps it to $1 \in \mathbb{Q}$.

Proposition 2.4. [4] *Convergent sequences (s_n) are bounded: there exists an $M \in \mathbb{N}$ such that $|s_n| < M$ for all $n \in \mathbb{N}$.*

Proof. The idea is simple: assume that $\lim s_n = s$; then for $\varepsilon = 1$, there is an N such that $|s_n - s| < \varepsilon = 1$. Put $M_0 = \max\{s_1, \ldots, s_N\}$ (this is ok since the set is finite). Then $|s_n| \leq M_0$ for all $n \leq N$ and $|s_n| = |s_n - s + s| \leq |s_n - s| + |s| < |s| + 1$ for all $n > N$. Now put $M = \max\{M_0, |s| + 1\}$. □

Proof of Thm. 2.3. There's a lot of work to do; we'll start with verifying that $(s_n + t_n)$ converges if (s_n) and (t_n) do.

So assume that $\lim s_n = s$ and $\lim t_n = t$, and assume we are given some $\varepsilon > 0$. We have to find an $N \in \mathbb{N}$ such that $|s_n + t_n - (s + t)| < \varepsilon$ for all $n > N$. We know that we can make $|s_n - s|$ and $|t_n - t|$ small, so we start by using the triangle inequality

$$|s_n + t_n - (s + t)| = |s_n - s + t_n - t| \leq |s_n - s| + |t_n - t|.$$

Since s_n converges to s, there is an $N_1 \in \mathbb{N}$ such that $|s_n - s| < \varepsilon/2$ for all $n > N_1$. Since t_n converges to t, there is an $N_2 \in \mathbb{N}$ such that $|t_n - t| < \varepsilon/2$ for all $n > N_2$. Put $N = \max\{N_1, N_2\}$. Then these inequalities hold for any $n > N$, and we find

$$|s_n + t_n - (s + t)| \leq |s_n - s| + |t_n - t| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which completes the proof.

The proof for $(s_n - t_n)$ is basically the same, so let's consider the product $s_n t_n$. Here we have to make $|s_n t_n - st|$ small. How is this related to $|s_n - s|$ and $|t_n - t|$? Here's how:

$$|s_n t_n - st| = |s_n t_n - s_n t + s_n t - st| \leq |s_n| \cdot |t_n - t| + |t| \cdot |s_n - s|.$$

This looks very good because we can make $|s_n - s|$ and $|t_n - t|$ small. There is a problem, however: $|t_n - t|$ is multiplied by $|s_n|$, so even if we can make $|t_n - t|$ small, $|s_n|$ may be so large that it kills all efforts to make the product small.

Fortunately, however, $|s_n|$ can't be too large: after all, (s_n) converges, hence is bounded, say by $|s_n| \leq M$ for some constant M. Now choose N_2 so large that $|t_n - t| < \varepsilon/2M$ for all $n > N_2$, and N_1 so large that $|s_n - s| < \varepsilon/2(|t| + 1)$

[4] Ross Thm. 9.1.

(we don't want to divide by 0 if (t_n) happens to converge to 0) for all $n > N_1$. Then put $N = \max\{N_1, N_2\}$ again and observe that

$$|s_n t_n - st| \leq |s_n| \cdot |t_n - t| + |t| \cdot |s_n - s| < M \cdot \frac{\varepsilon}{2M} + |t| \cdot \frac{\varepsilon}{2(|t|+1|)} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

where we have used that $\frac{|t|}{|t|+1} < 1$. This shows that $\lim s_n t_n = st$ and completes the proof. □

Example. In order to show that $\lim \frac{1}{n^2} = 0$, we would have to work with N's and ε's if we started with the definition. No one in his right mind would do so, however (excluding textbook authors). What we do is observe that $\frac{1}{n^2} = \frac{1}{n} \cdot \frac{1}{n}$; since we have already shown that $\lim \frac{1}{n} = 0$, we get $\lim \frac{1}{n^2} = 0$.

Exercise. Let $\mathcal{B}$ denote the set of bounded sequences. Show that $\mathcal{B}$ is a ring containing the ring C of converging sequences as a subring.

Exercise. Show that the set K of constant sequences in $\mathbb{Q}$ forms a subring of C.

We have already remarked that we cannot divide by sequences that contain a 0 term; if $t_n \neq 0$ for all $n \in \mathbb{N}$, then we can introduce the 'quotient' (s_n/t_n). What we would like to have is $\lim \frac{s_n}{t_n} = \frac{\lim s_n}{\lim t_n}$, but this cannot possibly hold if $\lim t_n = 0$. If we exclude these bad guys, however, we get

Proposition 2.5. *Assume that s_n and t_n converge; assume moreover that $t_n \neq 0$ for all $n \in \mathbb{N}$ and that $\lim t_n \neq 0$. Then s_n/t_n converges, and we have $\lim \frac{s_n}{t_n} = \frac{\lim s_n}{\lim t_n}$.*

Proof. The proof of the convergence of quotients is similar to that of products, but with an additional twist. In fact, we need to make $\left|\frac{s_n}{t_n} - \frac{s}{t}\right|$ small. We find

$$\left|\frac{s_n}{t_n} - \frac{s}{t}\right| = \left|\frac{s_n t - t_n s}{t_n t}\right| \leq \frac{|s_n t - st| + |ts - t_n s|}{|t_n t|}.$$

This looks good, because we can make $|s_n - s|$ and $|t_n - t|$ small and since s and t are constants, but we get a problem from the t_n in the denominator: if the t_n are very small, then their inverse will be very large, which is bad. But can the t_m be very small without t_n converging to 0? They cannot:

Lemma 2.6. [5] *If (t_n) converges to $t \neq 0$, then there exists an $N \in \mathbb{N}$ such that for all $n > N$ we have*

$$\begin{cases} t_n > \frac{1}{2}t & \text{if } t > 0, \\ t_n < \frac{1}{2}t & \text{if } t < 0. \end{cases}$$

In particular, $|t_n| > \frac{1}{2}|t|$ for all $n > N$.

[5] Ross, Example 6, §8

Assuming this for a moment, we can prove Prop. 2.5 as follows: let an $\varepsilon > 0$ be given; choose $N_1 \in \mathbb{N}$ such that $|s_n - s| < |t|\varepsilon/4$ for all $n > N_1$; choose $N_2 \in \mathbb{N}$ such that $|t_n - t| < t^2\varepsilon/4|s|$ for all $n > N_2$; and choose $N_3 \in \mathbb{N}$ such that $|t_n| > \frac{1}{2}|t|$ for all $n > N_3$. Then

$$\left|\frac{s_n}{t_n} - \frac{s}{t}\right| \;\leq\; \frac{|t||s_n - s| + |s||t - t_n|}{|t_n t|} \;\leq\; \frac{t^2/2}{t^2\varepsilon/4 + t^2\varepsilon/4} \;=\; \varepsilon$$

whenever $n > N := \max\{N_1, N_2, N_3\}$. □

Proof of Lemma 5. Put $\varepsilon = |t|/2$. Then there exists an $N \in \mathbb{N}$ such that $|t_n - t| < \varepsilon$ for all $n > N$. But then $|t| = |t - t_n + t_n| \leq |t_n - t| + |t_n|$, hence $|t_n| > |t| - |t_n - t| > |t|/2$.

If $t > 0$, then $|t_n| > t/2$ and $|t_n - t| < t/2$ imply $t_n > t/2$. If $t < 0$, on the other hand, the same argument shows that $t_n < t/2$. □

In some proofs, subsequences of sequences will turn up. These are just what the name says: if $a_1,\ a_2,\ a_3,\ \ldots$ is a sequence of rational numbers, then any infinite sequence you get by omitting terms from (a_n) is called a subsequence of (a_n). In mathematical terms: (b_n) is called a subsequence of (a_n) if there are integers

$$n_1 < n_2 < \ldots < n_k < n_{k+1} < \ldots$$

such that $b_k = a_{n_k}$ for all $k \in \mathbb{N}$. Note that $n_k \geq k$ (induction: $n_1 \geq 1$; if $n_k \geq k$, then $n_{k+1} > n_k \geq k$, hence $n_{k+1} \geq k + 1$).

Subsequences are well behaved:

Proposition 2.7. [6] *If (a_{n_k}) is a subsequence of a converging sequence (a_n), then (a_{n_k}) converges, and we have*

$$\lim_{k\to\infty} a_{n_k} = \lim_{n\to\infty} a_n.$$

Proof. Let $a = \lim a_n$; we have to make $|a_{n_k} - a|$ small. We know that $|a_n - a|$ is small, but that's enough because the a_{n_k}'s are among the a_n somewhere.

In fact, let $\varepsilon > 0$; then there is an $N \in \mathbb{N}$ such that $|a_k - a| < \varepsilon$ for all $k > N$. But then we have $|a_{n_k} - a| < \varepsilon$ for all $k > N$ because $n_k \geq k$, and that's it. □

Proposition 2.8. *Let (s_n) be a sequence converging to s. If $a \leq s_n \leq b$ for almost all n (that is, for all n larger than some integer N), then $a \leq s \leq b$.*

Proof. It is sufficient to prove that $a \leq s_n$ for almost all n implies $a \leq s$. Assume that this is false; then $a > s$, and since $s_n \geq a$ for all $n > N$, we get $s_n - s = s_n - a + a - s \geq a - s$. If we put $\varepsilon = a - s > 0$, then this shows $s_n - s \geq \varepsilon$, in particular $|s_n - s| > \varepsilon$, for all $n > N$. This contradicts the fact that $|s_n - s|$ can be made arbitrarily small. □

[6] Ross, Thm. 11.2.

Note that strict inequalities in general do not survive the limit: we have $\frac{1}{n} > 0$ for all $n \in \mathbb{N}$, but $\lim \frac{1}{n} = 0$.

Consider the sequence defined by $a_1 = 2$ and $a_{n+1} = \frac{1}{2}a_n + 1/a_n$. The first few terms are $a_1 = 2$, $a_2 = 1.5$, $a_3 = 1.41666\ldots$, $a_4 = 1.41421\ldots$, so it seems the sequence converges. But does it?

Let us start with the observation

$$a_{n+1}^2 = \Big(\frac{1}{2}a_n + \frac{1}{a_n}\Big)^2 = \Big(\frac{1}{2}a_n - \frac{1}{a_n}\Big)^2 + 2;$$

we deduce that $a_{n+1}^2 - 2 \geq 0$; in fact, since the a_n are rational numbers (induction!), we must have $a_{n+1}^2 > 2$ for all $n \geq 1$. This in turn implies that

$$a_{n+1}^2 = \frac{1}{4}a_n^2 + 1 + \frac{1}{a_n^2} < \frac{1}{4}a_n^2 + \frac{3}{2},$$

hence

$$a_{n+1}^2 - 2 < \frac{1}{4}(a_n^2 - 2).$$

What this means is that the difference between a_n^2 and 2 shrinks by at least a factor of 4 every time we increase n by 1. In other words: we have $\lim(a_n^2 - 2) = 0$, or $\lim a_n^2 = 2$.

Now if a_n would converge, say $\lim a_n = a$, then the above discussion would show that $a^2 = 2$. But there is no rational a whose square equals 2, so a_n does not converge. By now you should have the impression that this is due to the fact that $\mathbb{Q}$ contains not enough elements to provide limits for sequences that seem to converge. This observation is at the basis for our construction of the reals: we simply 'adjoin to $\mathbb{Q}$ the limits of sequences that seem to converge'. We shall make this precise in the next section by introducing Cauchy sequences.

Some Algebra

[This is for those of you who are familiar with algebra. Don't be alarmed if you are not.]

We have seen in Theorem 2.3 that the set C of converging sequences in $\mathbb{Q}$ forms a ring with respect to addition and multiplication, and that $\lim : C \longrightarrow \mathbb{Q}$ is a ring homomorphism.

Now consider the subset $N \subset C$ of sequences converging to 0; clearly, the sum and product of sequences converging to 0 again converge to 0, hence N is not only a subset but a subring of C. Even better: N is an ideal!

Recall that a subring S of R is called an ideal if not only products of elements of S are again in S, but if products of elements of S with elements of R are elements of S again, that is: if S is closed under multiplication by elements of R.

In our case, the product of a converging sequence and a sequence converging to 0 is a sequence converging to 0 (PROOF!), that is, N is closed with respect to multiplication by elements in C.

In algebra you learn that ideals are exactly the kernels of ring homomorphisms. The only ring homomorphism around here was the map $\lim : C \longrightarrow \mathbb{Q}$; the kernel of a ring homomorphism is the set of elements that get mapped to 0, so the kernel of lim is indeed N. Since lim is clearly surjective (given $q \in \mathbb{Q}$, the constant sequence $q, q, q, \ldots$ converges to q), a standard theorem in algebra shows

Theorem 2.9. *We have $C/N \simeq \mathbb{Q}$, where the isomorphism is induced by* lim.

Series

Out of any sequence a_n we can form a new sequence $s_n = \sum_{k=1}^{n} a_n$. If (s_n) converges to $s = \lim s_n$, then we write

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

and call s an infinite series.

The most important series in much of calculus and complex analysis is the geometric series; in addition, it does us the favor of converging:

Theorem 2.10. *The series $1 + q + q^2 + q^3 + \ldots$ is called the geometric series. It converges for any $q \in \mathbb{Q}$ with $|q| < 1$, and we have $\sum_{k=1}^{\infty} q^k = \frac{1}{1-q}$.*

Proof. We have to show that the sequence $s_n = \sum_{k=1}^{n} q^k$ of partial sums converges. This is easy if you remember the little trick involved:

$$\begin{array}{rcllllll}
S_n & = & 1+ & q+ & q^2+ & \ldots + & q^n & \\
qS_n & = & & q+ & q^2+ & \ldots + & q^n+ & q^{n+1} \\
\hline
S_n - qS_n & = & 1 & & & & & -q^{n+1}
\end{array}$$

hence $S_n(1-q) = 1 - q^{n=1}$ and finally

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

whenever $q \neq 1$. If $|q| < 1$, then $\lim q^n = 0$, hence $\lim S_n = \frac{1}{1-q}$ as claimed. $\square$

Here's a little application:

Exercise. An $m \times m$-matrix M is called nilpotent if there exists an integer $n \geq 1$ such that $M^n = 0$ (if you've never seen this, show that upper triangular matrices with 0's on the diagonal are all nilpotent). If M is nilpotent, show that $E - M$ has an inverse, where E is the $m \times m$ unit matrix.

2.2 Cauchy Sequences

The problem with our definition of convergence is: how can we prove that a sequence converges if we can't write down it's limit? Well, we can't.

What we can do is to define a (possibly larger) set of sequences that 'behave as if they would converge'. This definition should not involve the limit of the sequence (because, in general, we don't know it).

How do we do that? If a sequence (s_n) converges to s, then the difference between s_n and s becomes as small as we want for large n. This means that the difference between any two terms s_n and s_m becomes small if only n and m are large enough. So here's what we'll do:

A Cauchy sequence in $\mathbb{Q}$ is a sequence (s_n) with the following property:

> For every $\varepsilon > 0$ there is an $N > 0$ such that $|s_n - s_m| < \varepsilon$ for all $m, n > N$.

Remark. This condition is due to Cauchy (Analyse algébrique (1821), p. 125). Actually he proved that sequences of real numbers converge if and only if they satisfy the condition above.

From the way we arrived at the definition we expect that converging sequences are Cauchy. They are:

Proposition 2.11. [7] *Any converging sequence is a Cauchy sequence.*

Proof. Let (s_n) be a sequence converging to s. This means that we can make $|s_n - s|$ small. However, we need to make $|s_n - s_m|$ small. How do we do that? You should be able to guess the answer by now:

$$|s_n - s_m| = |s_n - s + s - s_m| \leq |s_n - s| + |s_m - s|, \tag{2.1}$$

and these terms can be made small.

OK, here's the formal proof: given $\varepsilon > 0$, we pick an integer N such that $|s_n - s| < \varepsilon/2$ for all $n > N$. By (2.1), this implies that $|s_n - s_m| < \varepsilon$ for all $m, n > N$, and this proves that (s_n) is Cauchy. □

As a matter of fact, Cauchy sequences share most of the properties of converging series; as an Exercise, formulate the analog of Prop 2.2 and prove it. As an example of how to do it, we will go through the proof of the following

Proposition 2.12. [8] *Cauchy sequences are bounded.*

Proof. Let (a_n) be a Cauchy sequence. Since it suffices to prove that $(|a_n|)$ is bounded, we may assume without loss of generality that $a_n \geq 0$.

Since (a_n) is Cauchy, there is an $N \in \mathbb{N}$ such that $|a_m - a_n| < 1$ for all $m, n > N$. Assume that (a_n) is not bounded; then for $m > N$ there exists an $n > m$ such that $a_n > a_m + 1$: otherwise, $\max\{a_1, \ldots, a_m, a_m + 1\}$ would be an upper bound. But now $|a_m - a_n| = a_m - a_n > 1$: contradiction! □

[7] Ross 10.9.
[8] Ross, Lemma 10.10.

We shall also need the analog of Lemma 5 for Cauchy sequences:

Lemma 2.13. *Let (t_n) be a Cauchy sequence that does not converge to 0. Then there exists an $N \in \mathbb{N}$ and a $\tau > 0$ such that either $t_n > \tau$ for all $n > N$ or $t_n < -\tau$ for all $n > N$. In particular, we have $|t_n| > \tau$ for all $n > N$.*

Proof. We clearly have to exploit the fact that (t_n) does not converge to 0 for getting a $\tau > 0$ that will work. Now (t_n) converges to zero if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n > N : \ |t_n| = |t_n - 0| < \varepsilon.$$

What does this tell us about sequences not converging to 0? Well, if (t_n) does not converge to 0, then

$$\exists \varepsilon > 0 \quad \forall N \in \mathbb{N} \quad \exists n > N : \ |t_n| = |t_n - 0| > \varepsilon.$$

In plain words: there must be an $\varepsilon > 0$ such that no matter how large you choose N, there is always some $n > N$ such that t_n lies outside the ε-strip around 0.

Now put $\tau = \varepsilon/2$ and choose $N \in \mathbb{N}$ such that $|t_m - t_n| < \varepsilon/2$ for all $m, n > N$. Now we pick a particular m by demanding that $|t_m| > \varepsilon$, which we can do by the preceding discussion.

If $t_m > \varepsilon$, then $t_m - t_n < \varepsilon/2$ gives $t_n > t_m - \varepsilon/2 > \varepsilon/2$ for all $n > N$, and our claim follows.

If $t_m < -\varepsilon$, then $t_m - t_n > -\varepsilon/2$ gives $t_n < t_m + \varepsilon/2 < -\varepsilon/2$ for all $n > N$, and again we are home. This concludes the proof. □

We have seen above that converging sequences are Cauchy. Conversely, Cauchy sequences in $\mathbb{Z}$ converge: in fact, let (s_n) be a Cauchy sequence of integers, and pick $\varepsilon = \frac{1}{2}$; then there is an $N \in \mathbb{N}$ such that $|s_m - s_n| < \varepsilon$ for all $m, n > N$. This implies that $s_m = s_n$ for all $m, n > N$, that is, the sequence eventually becomes constant; in particular, it converges.

Unfortunately, this result does not carry over to $\mathbb{Q}$: there are Cauchy sequences in $\mathbb{Q}$ that do not converge. In fact, the sequence (a_n) constructed above is such a beast. Since we already know that it does not converge, all we need to prove is that it is Cauchy. This follows from a slightly more general result which shows that any sequence (a_n) with $|a_{n+1} - a_n|$ decreasing quickly is Cauchy:

Proposition 2.14. [9] *Let (a_n) be a sequence with $|a_{n+1} - a_n| < Cq^n$ for constants $C > 0$ and $0 \leq q < 1$. Then (a_n) is a Cauchy sequence.*

Proof. For $m > n$ we have

$$\begin{aligned} |a_m - a_n| &= |a_m - a_{m-1} + a_{m-1} - \ldots - a_{n+1} + a_{n+1} - a_n| \\ &< |a_m - a_{m-1}| + |a_{m-1} - a_{m-2}| + \ldots + |a_{n+1} - a_n| \\ &\leq C[q^{m-1} + q^{m-2} + \ldots + q^n] \\ &= Cq^n[1 + q + q^2 + \ldots + q^{m-n-1}]. \end{aligned}$$

[9] Ross, Exercise 10.6.

The sum in the brackets is smaller than $\sum_{k=1}^{\infty} q^k = \frac{1}{1-q}$, hence $|a_m - a_n| < q^n \frac{C}{1-q}$. Since we may make this expression as small as we wish by choosing n sufficiently large, the claim follows. $\square$

So (a_n) is a Cauchy sequence, but, as we have already seen, it does not converge. Here's another proof for the last claim: assume that $a = \lim a_n$; then $a \neq 0$ because clearly $a_n \geq 1$ for all $n \geq 1$ (induction). Moreover $a = \lim a_{n+1}$, hence $a = \lim a_{n+1} = \lim(\frac{1}{2}a_n + \frac{1}{a_n}) = \frac{a}{2} + \frac{1}{a}$, hence $a^2 = 2$. But there is no $a \in \mathbb{Q}$ such that $a^2 = 2$.

There are other useful criteria that allow us to conclude that a given sequence is Cauchy; the following is particularly simple and useful:

Theorem 2.15. *Every monotone bounded sequence is Cauchy.*

Here, a bounded sequence is a sequence (a_n) such that $|a_n| \leq M$ for all $n \in \mathbb{N}$ and some constant M; a monotone sequence (a_n) is one that satisfies one of the following conditions:

$$a_1 \leq a_2 \leq a_3 \leq \dots \quad \text{or}$$
$$a_1 \geq a_2 \geq a_3 \geq \dots$$

We call (a_n) monotone increasing or decreasing according as we are in the first or the second case.

Proof of Thm. 2.15. Assume that (a_n) is monotone (say decreasing; otherwise, consider $(-a_n)$) and not Cauchy; we have to show that it is not bounded.

Here's my original attempt: Let $\varepsilon > 0$ and put $s_1 = a_1$; since (a_n) is not Cauchy, there is an $m > 1$ such that $|a_1 - a_m| \geq \varepsilon$; we put $s_2 = a_m$. By repeating this construction I could show that (a_n) is not bounded.

However, not being Cauchy does not guarantee that, for any $\varepsilon > 0$ there is a pair $m, n \in \mathbb{N}$ with $|a_n - a_m| \geq \varepsilon$: if the sequence is bounded and ε is large enough, all terms will differ from each other by something less than ε.

Here's the correct way of exploiting that (a_n) is not Cauchy: we know that (a_n) is Cauchy if

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{N}\ \forall\, n, m > N:\ |a_n - a_m| < \varepsilon.$$

Thus if a_n is not Cauchy, there must be some $\varepsilon > 0$ for which this is false; more exactly:

$$\exists\, \varepsilon > 0\ \forall\, N \in \mathbb{N}\ \exists\, n, m > N:\ |a_n - a_m| \geq \varepsilon.$$

This works in general: finding the negation of a statement involving quantors you simply replace $\exists$ by $\forall$ and vice versa, and you replace the actual claim by its negation.

Now let $\varepsilon > 0$ be a positive number such that for all $N \in \mathbb{N}$ there exist $m, n \in \mathbb{N}$ with $|a_n - a_m| \geq \varepsilon$. Pick any such pair $n_1 < m_1$ (they are necessarily different, and the order doesn't matter) and put $N = m_1$. Then pick a pair $n_2 < m_2$ with $|a_{n_2} - a_{m_2}| \geq \varepsilon$, put $N = m_2$, and repeat this reasoning. This

way we get a sequence of indices $n_1 < m_1 < n_2 < m_2 < n_3 < \ldots$ such that $a_{n_1} > a_{m_1} \geq a_{n_2} > a_{m_2} \geq a_{n_3} > \ldots$; in fact, we have

$$\begin{aligned} a_{n_1} - a_{m_k} &= a_{n_1} - a_{m_1} + a_{m_1} + \ldots - a_{m_{k-1}} + a_{m_{k-1}} - a_{m_k} \\ &\geq a_{n_1} - a_{m_1} + a_{n_2} + \ldots + a_{n_{k-1}} - a_{m_{k-1}} + a_{n_k} - a_{m_k} \\ &> k\varepsilon. \end{aligned}$$

By making k large enough we can make $a_{n_1} - a_{m_k}$ as large as we please (by the Archimedean property of the rationals): this means that (a_n) can't be bounded. $\square$

Ring of Cauchy Sequences

We have seen in the preceding section that the set of converging sequences in $\mathbb{Q}$ form a ring. The same is true for the set $\mathcal{R}$ of Cauchy sequences: if (a_n) and (b_n) are Cauchy sequences, then so are $(a_n \pm b_n)$ and $(a_n b_n)$.

Exercise. Prove that $\mathcal{R}$ is a ring.

Here's a hint for the case $a_n + b_n$. Given an $\varepsilon > 0$, you have to find an integer N such that $|(a_m + b_m) - (a_n + b_n)| < \varepsilon$ for all $m, n > N$. Now use the fact that (a_n) and (b_n) are Cauchy, and don't forget the triangle inequality.

Exercise. Prove that if (a_n) is a Cauchy sequence not converging to 0 and such that $a_n \neq 0$ for all $n \in \mathbb{N}$, then $(1/a_n)$ is also Cauchy.

Exercise. Prove that $\mathcal{R}$ contains the set $\mathcal{N}$ of sequences converging to 0 as a subring.

Actually, much more is true: $\mathcal{N}$ is not only a subring of $\mathcal{R}$, it is an ideal. This means that, for any Cauchy sequence $(a_n) \in \mathcal{R}$ and any sequence $(b_n) \in \mathcal{N}$, the product $(a_n b_n)$ is not only Cauchy, but lies again in $\mathcal{N}$.

Exercise. Prove that $\mathcal{N}$ is an ideal in $\mathcal{R}$.

Remark. Those with a working knowledge about rings and ideals can now immediately define the real numbers $\mathbb{R}$ as the ring $\mathcal{R}/\mathcal{N}$. We shall define $\mathbb{R}$ in the next chapter as a set of equivalence classes, and then define a ring structure on this set: in the algebraic approach, we get this ring structure for free.

Remark. We can define the following sets of sequences of rational numbers:

$$\begin{aligned} \mathcal{N} &= \text{sequences converging to 0} \\ \mathcal{C} &= \text{converging sequences} \\ \mathcal{R} &= \text{Cauchy sequences} \\ \mathcal{B} &= \text{bounded sequences} \\ \mathcal{S} &= \text{sequences} \end{aligned}$$

These are all rings, and in fact

$$\mathcal{N} \subset \mathcal{C} \subset \mathcal{R} \subset \mathcal{B} \subset \mathcal{S},$$

where all inclusions are proper.

Chapter 3

From $\mathbb{Q}$ to $\mathbb{R}$ (and $\mathbb{C}$)

3.1 The Reals $\mathbb{R}$

The construction of the real numbers differs fundamentally from the construction of the integers or the rationals, and in fact some mathematicians (e.g. Kronecker) did not accept the following construction (and thus the existence of real numbers as we know them) for mainly philosophical reasons.

Let $\mathcal{R}$ denote the set of all Cauchy sequences of rational numbers. We say that two sequences (a_n) and (b_n) are equivalent (and write $(a_n) \sim (b_n)$) if $\lim(a_n - b_n) = 0$, that is, if they differ by a null sequence (a sequence converging to 0).

WARNING. This does not imply that $\lim a_n = \lim b_n$, because these limits may not exist!

Let $\mathbb{R} = \mathcal{R}/\sim$ be the set of all equivalence classes of $\mathcal{R}$. Then we can identify $\mathbb{Q}$ with a subset of $\mathbb{R}$: in fact, for $r \in \mathbb{Q}$, the constant sequence $r_n = r$ is certainly Cauchy, so the map $\iota : \mathbb{Q} \longrightarrow \mathbb{R}$ sending r to the real number $[r_n] := [(r_n)]$ allows us to identify r with the real number $\iota(r)$.

The map ι is injective: if $\iota(r) = \iota(s)$, then $(r_n) \sim (s_n)$ for the constant seties $r_n = r$ and $s_n = s$; this implies that $r_n - s_n$ converges to 0, and so $r = s$.

Defining addition and multiplication on the reals is a breeze: we set

$$[r_n] \pm [s_n] = [r_n \pm s_n], \tag{3.1}$$

$$[r_n] \cdot [s_n] = [r_n \cdot s_n]. \tag{3.2}$$

First we have to verify that this is defined at all: but if (r_n) and (s_n) are Cauchy sequences, then so are $(r_n \pm s_n)$ and $(r_n s_n)$, so the expressions on the right hand side of (3.1) and (3.2) make sense.

Next we claim that (3.1) is well defined: assume that $[r_n] = [r'_n]$ and $[s_n] = [s'_n]$; we have to show that $[r_n + s_n] = [r'_n + s'_n]$. But $r_n - r'_n$ and $s_n - s'_n$ are sequences converging to 0, hence so is $r_n + s_n - (r'_n + s'_n)$, and this implies the claim. The proof that multiplication is well defined is similar.

The new addition on real numbers agrees with the corresponding operations on $\mathbb{Q}$ under the embedding $\iota : \mathbb{Q} \hookrightarrow \mathbb{R}$, in other words: that the map ι is a ring homomorphism. In fact, if r and s are rational numbers, and if $\iota(r) = [r_n]$ and $\iota(s) = [s_n]$ are the corresponding real numbers defined by the constant sequences $r_n = r$ and $s_n = s$, then

$$\iota(r) + \iota(s) = [r_n] + [s_n] = [r_n + s_n] = \iota(r + s),$$

and similarly for products. Moreover, $\iota(0)$ is the neutral element of addition, and $\iota(1)$ that of multiplication.

So far we know that $\mathbb{R}$ is a ring containing $\mathbb{Q}$ as a subring. Before we can show (or even claim) that $\mathbb{R}$ is a field we have to define division, which is a bit more complicated: if (r_n) and (s_n) are Cauchy sequences such that $\lim s_n$ exists and is different from 0, then we cannot simply put

$$[r_n]/[s_n] = [r_n/s_n]$$

because although s_n does not converge to 0, some of the s_n might vanish. Here's one possible way around it: we claim

Lemma 3.1. *If (s_n) is a Cauchy sequence not converging to 0, then $s_n \neq 0$ for almost all n, that is, for all n larger than some natural number N.*

Proof. Assume not. Then there are infinitely many indices n_1, n_2, $\ldots$, such that $a_{n_k} = 0$. Since (a_n) is Cauchy, for every $\varepsilon > 0$ there is an N such that $|a_n - a_m| < \varepsilon$ for all $m, n > N$. In particular, $|a_{n_k} - a_m| < \varepsilon$ for all $m, n_k > N$. But $a_{n_k} = 0$, hence we have proved that for every $\varepsilon > 0$ we have $|a_m| < \varepsilon$ for all sufficiently large m: but this implies that (a_m) converges to 0. $\square$

Actually this lemma is a special case of Prop. 2.13; the proof given here is slightly simpler because we have proved less.

Thus if (s_n) is a Cauchy sequence not converging to 0, only finitely many terms actually vanish. The sequence (s'_n) defined by

$$s'_n = \begin{cases} s_n & \text{if } s_n \neq 0 \\ \frac{1}{n} & \text{if } s_n = 0 \end{cases}$$

differs from (s_n) in only finitely many places. In particular, $s_n - s'_n$ is a null sequence, hence $[s_n] = [s'_n]$. Moreover, we know that $(1/s'_n)$ is a Cauchy sequence by Lemma 2.13 (in fact, $s_n = s'_n$ for all $n > N$ for some N, and the sequence $a_n = s_{n+N}$ is a Cauchy sequence satisfying the conditions of Lemma 2.13. Thus $1/a_n$ is a Cauchy sequence, and since adding finitely many terms to a_n does not change this, so is $1/s'_n$). Thus if (r_n) and (s_n) are Cauchy sequences and if (s_n) is not a null sequence, then (r_n/s'_n) is a Cauchy sequence, and we may define division by

$$[r_n]/[s_n] = [r_n/s'_n]. \tag{3.3}$$

Using these definitions, it is now straight forward (if tedious) to verify that the operations on the reals as constructed above satisfy the usual properties, and in particular that the set $\mathbb{R}$ of reals forms a field.

Let us show for example that every nonzero element of $\mathbb{R}$ has an inverse. To this end, consider a real number $s \neq 0$; then $s = (s_n)$ for a Cauchy sequence (s_n) that does not converge to 0. Replacing (s_n) by (s'_n) as above we get a real number $r = [1/s_n]$, and clearly $rs = \iota(1)$, where $\iota(1)$ is the real number represented by the constant sequence $1, 1, 1, \ldots$.

Theorem 3.2. *The real numbers $\mathbb{R}$ form a field containing $\mathbb{Q}$ as a subfield.*

What we would like to do now is to prove that $\mathbb{R}$ does what we want it to do, namely provide a limit for every Cauchy sequence (r_n) in $\mathbb{Q}$. But what does it mean that $\lim r_n = r$? It should mean that we can make $|r_n - r|$ small, but this does not yet make sense because we do not yet have an absolute value on $\mathbb{R}$. This will be taken care of next.

$\mathbb{R}$ as an ordered set

We say that $x > y$ if $x - y > 0$, so it is sufficient to define when $r > 0$ for an $r \in \mathbb{R}$. Assume that $r \neq 0$ and let $r = [r_n]$. By Lemma 2.13, we either have $r_n > 0$ or $r_n < 0$ for sufficiently large n. In the first case, we put $r > 0$, in the second we put $r < 0$. As for rationals, we say $r \geq 0$ if $r = 0$ or $r > 0$, and define $r \leq 0$ accordingly.

We claim that this is well defined. In order to prove this, let (r_n) and (r'_n) be two Cauchy sequences in $\mathbb{Q}$ representing the same real number $r = [r_n] = [r'_n]$. If $r > 0$, then there is a $\tau > 0$ such that $r_n > \tau$ for all $n > N$ for a suitable $N \in \mathbb{N}$. Since $(r_n - r'_n)$ converges to 0, there is an $N_0 \in \mathbb{N}$ such that $|r_n - r'_n| < \frac{1}{2}\tau$ for all $n > N_0$. But then $r'_n = r'_n - r_n + r_n > r_n - \frac{1}{2}\tau > \frac{1}{2}\tau$ for all $n > \max\{N, N_0\}$, so $r'_n > 0$ for all sufficiently large n. Thus it does not matter whether we represent r by (r_n) or (r'_n).

It is now easy to arrest the usual suspects:

Proposition 3.3. *The order $<$ on $\mathbb{R}$ satisfies the usual properties:*

1. *The identification map $\iota : \mathbb{Q} \hookrightarrow \mathbb{R}$ respects the order: for $x, y \in \mathbb{Q}$ we have $\iota(x) < \iota(y)$ if and only if $x < y$.*

2. *(Trichotomy Law) For $r \in \mathbb{R}$, exactly one of the statements $r < 0$, $r = 0$ and $r > 0$ holds.*

3. *If $x, y \in \mathbb{R}$ and $x > 0$, $y > 0$, then $xy > 0$. Similarly, if $x > 0$ and $y < 0$, then $xy < 0$, and if $x, y < 0$, then $xy > 0$.*

4. *We have $x^2 \geq 0$ for all $x \in \mathbb{R}$. In particular, there is no real number whose square equals -1.*

5. *If $x < y$ and $y < z$, then $x < z$.*

6. *If $x \geq y$ and $y \geq x$, then $x = y$.*

Proof. We'll leave some of the proofs as an exercise. Assume that $x, y \in \mathbb{Q}$. Then $\iota(x)$ is the real number that is represented by the Cauchy sequence $x_n = x$ for $n \in \mathbb{N}$. Thus $\iota(x) - \iota(y) = \iota(x - y)$ is represented by the sequence $z_n = x - y$. The claim is that $x - y < 0$ if and only if $[z_n] < 0$. But if $x < y$, then the terms of (z_n) (namely $x - y$) are all negative, so $[z_n] < 0$ by our definition, and the converse is just as obvious. This proves 1.

For 2, observe that $r < 0$ and $r = 0$ cannot hold at the same time by definition. Moreover, we cannot have $r < 0$ and $r > 0$: in fact, put $r = [r_n]$. Then $r < 0$ implies that $r_n < 0$ for all sufficiently large n, whereas $r > 0$ implies $r_n > 0$ for large n. But the r_n are rational, and by the trichotomy law for rationals we cannot have $r_n > 0$ and $r_n < 0$ at the same time.

Now let $r = [r_n]$ be a real number. If r_n converges to 0, then $r = 0$ because (r_n) and the sequence 0, 0, 0, ... differ by a null sequence. If (r_n) does not converge to 0, then Lemma 2.13 shows that either $r_n > 0$ or $r_n < 0$ for all sufficiently large n, hence by definition either $r > 0$ or $r < 0$. This proves the trichotomy law for $\mathbb{R}$.

Now assume that x and y are real numbers with $x > 0$ and $y > 0$. By definition this means that if $x = [x_n]$ and $y = [y_n]$ for Cauchy sequences (x_n) and (y_n) of rational numbers, then there exist $N_1, N_2 \in \mathbb{N}$ with $x_n > 0$ for $n > N_1$ and $y_n > 0$ for $n > N_2$. But then $x_n y_n > 0$ for all $n > \max\{N_1, N_2\}$, hence $xy = [x_n y_n] > 0$.

The proofs of the other claims are left as an exercise. □

Now we introduce an absolute value on $\mathbb{R}$ by defining

$$|r| = \begin{cases} +r & \text{if } r \geq 0, \\ -r & \text{if } r < 0. \end{cases}$$

It is easy to check that $|\cdot|$ has the usual properties; in particular, it satisfies the triangle inequality.

We shall occasionally need to show that certain real numbers are small; assume that $\varepsilon > 0$ is rational and $r = [r_n]$ is a real number. Then $|r| < \varepsilon$ means $-\varepsilon < r < \varepsilon$. and by definition of $<$ this means that there is a natural number N such that $-\varepsilon < r_n < \varepsilon$.

Finally, let's prove the Archimedean property.

Lemma 3.4. *If $x > 0$ is a real number, then there is a rational number r such that $0 < r < x$.*

Proof. Let (x_n) be a sequence of rational numbers with $x = [x_n]$. Lemma 2.13 provides us with a positive rational number $r > 0$ such that $x_n > r$ for all sufficiently large $n \in \mathbb{N}$. By definition of the order, this means $x > r$. □

Corollary 3.5 (Archimedean Property). [1] *If $a, b > 0$ are real numbers, then there is a natural number $n \in \mathbb{N}$ such that $na > b$.*

[1] Ross, 4.6.

Proof. Put $x = \frac{a}{b}$; by the above Lemma there is a rational number $\frac{p}{q}$ such that $0 < \frac{p}{q} < x$. Then $\frac{b}{a} < \frac{q}{p} < q + 1$, so $n = q + 1 \in \mathbb{N}$ does it. $\square$

Corollary 3.6. *For any real $x > 0$, there is an integer $z \geq 0$ such that $z \leq x < z + 1$.*

Proof. Consider the set $M = \{n \in \mathbb{N} : n > x\}$. Then M is a non-empty set of natural numbers, since applying the archimedean property with $a = 1$ and $b = x$ shows that there is some $n \in \mathbb{N}$ with $n > x$. Since $\mathbb{N}$ is well-ordered, M has a minimal element n. Put $z = n - 1$. Then $0 \leq z \in \mathbb{Z}$ and $z + 1 \geq x$ follow immediately, and $z \leq x$ also holds: for if not, then $z = n - 1 > x$, which contradicts the minimality of n unless $n = 1$, in which case $0 = z > x$ contradicts the assumption $x > 0$. $\square$

The integer z in the above corollary is called the largest integer less than or equal to x; we write $z = \lfloor x \rfloor$.

Theorem 3.7. [2] *The rationals $\mathbb{Q}$ are dense in $\mathbb{R}$: for all real numbers $a < b$, there is a rational number q such that $a < q < b$.*

Proof. We have $b - a > 0$, so by the Archimedean property there is an integer n with $n(b - a) > 1$, that is, with $nb > na + 1$. Put $z = \lfloor nb \rfloor$; then $z + 1 > nb \geq z$, hence $z > nb - 1 > na$. Together, this shows that $nb \geq z > na$, hence $b \geq \frac{z}{n} > a$. $\square$

$\mathbb{R}$ as a complete field

Now that we have an absolute value, we can carry over a couple of definitions from the rationals: a sequence (r_n) of real numbers is called

- *bounded* if there exists an $M \in \mathbb{R}$ such that $|r_n| < M$ for all $n \in \mathbb{N}$.
- convergent if there is an $r \in \mathbb{R}$ such that, for all $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that $|r_n - r| < \varepsilon$; we call r the limit of (r_n) and write $r = \lim r_n$.
- a Cauchy sequence if for all $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that $|r_n - r_m| < \varepsilon$ for all $m, n > N$.

Let $\mathcal{N}$ ($\mathcal{C}$, $\mathcal{R}$, $\mathcal{B}$) denote the set of null (converging, Cauchy, bounded sequences. Again, all these sets are rings, in fact subrings of the ring $\mathcal{S}$ of sequences of real numbers. The proofs are *exactly* the same. The main difference is that in the inclusions

$$\mathcal{N} \subset \mathcal{C} \subseteq \mathcal{R} \subset \mathcal{B} \subset \mathcal{S},$$

the inclusion $\mathcal{C} \subseteq \mathcal{R}$ is not proper anymore: a major theorem that we shall prove below is that $\mathcal{C} = \mathcal{R}$ in $\mathbb{R}$. Moreover, converging sequences have exactly one limit (same proof).

Now we are ready to show

[2] Ross, 4.7.

Proposition 3.8. *Every Cauchy sequence in $\mathbb{Q}$ converges to some real number.*

Proof. Let (r_n) be a Cauchy sequence of rational numbers. We claim that $\lim r_n = r$, where $r = [r_n]$. To this end, let an $\varepsilon > 0$ be given (we may assume that ε is rational; if it is not, we replace it by some smaller positive rational number); we have to find an integer N such that $|r_n - r| < \varepsilon$ for all $n > N$. By definition of the order $<$, this holds if and only if the inequality $|r_n - r_m| < \varepsilon$ of rational numbers holds for all sufficiently large m. But since (r_n) is a Cauchy sequence, there is an N such that $|r_n - r_m| < \varepsilon$ for all $m, n > N$. $\square$

We now formulate the result on which the rest of this course is built:

Theorem 3.9. *The field $\mathbb{R}$ of reals is complete, that is: every Cauchy sequence converges.*

At first it might seem that there is nothing to prove here. After all, didn't we introduce $\mathbb{R}$ by 'adding' the limits of every Cauchy sequence to $\mathbb{Q}$? Doesn't that mean that every Cauchy sequence converges?

We have to be very careful here: by Prop. 3.8, only Cauchy sequences consisting of rational numbers converge in $\mathbb{R}$. But maybe there is a Cauchy sequence consisting of real numbers whose limit is not real? Well, there is no such thing, but this must be proved.

Proof. Assume that (a_n) is a Cauchy sequence of real numbers. We have to find a real number a such that $\lim a_n = a$. Real numbers are limits of Cauchy sequences in $\mathbb{Q}$, so we need to construct a Cauchy sequence in $\mathbb{Q}$ with the same limit as (a_n). Here's how we do that:

Every real number a_n has the form $a_n = [a_{n,k}]$, where $(a_{n,k})$ is a Cauchy sequence of rational numbers. By Proposition 3.8, this Cauchy sequence converges to a_n. Let us arrange all these numbers as follows:

$$\begin{array}{cccccc}
a_{1,1} & a_{2,1} & a_{3,1} & \dots & a_{n,1} & \dots \\
a_{1,2} & a_{2,2} & a_{3,2} & \dots & a_{n,2} & \dots \\
a_{1,3} & a_{2,3} & a_{3,3} & \dots & a_{n,3} & \dots \\
\vdots & \vdots & \vdots & & \vdots & \\
a_{1,n} & a_{2,n} & a_{3,n} & \dots & a_{n,n} & \dots \\
\vdots & \vdots & \vdots & & \vdots & \\
a_1 & a_2 & a_3 & \dots & a_n & \dots
\end{array}$$

Now the first idea that came to my mind was making up a new sequence by putting $s_1 = a_{1,1}$, $s_2 = a_{2,2}$, $\dots$, $s_n = a_{n,n}$. If one could show that this sequence of rational numbers is Cauchy, it would have a limit a, and then a proof that $a = \lim a_n$ would be mere child's play. After an hour of futile attempts to prove that (s_n) is Cauchy I realized that it is easy to show that this approach cannot possibly work: all you have to do is to replace $a_{n,n}$ by n; then for all $m \in \mathbb{N}$, $a_{m,\nu}$ would still be a Cauchy sequence converging to a_m, but our sequence $s_n = n$ would certainly not converge.

This shows that you have to make sure that the s_m you pick from the sequence $a_{m,n}$ is very close to the limit a_m. Here's how to do that: since $(a_{m,n})$ converges to a_m, we can find an n such that $|a_{m,n} - a_m| < \frac{1}{m}$ and then put $s_m = a_{m,n}$.

Now we make two claims:

A. the sequence (s_n) is Cauchy;

B. we have $\lim s_n = \lim a_n$.

A) (s_n) is a Cauchy sequence. We have to make $|s_\mu - s_\nu|$ small. All we know about the s_μ is that they are close to a_μ, so we try to invoke the old triangle:

$$\begin{aligned} |s_\mu - s_\nu| &= |s_\mu - a_\mu + a_\mu - a_\nu + a_\nu - s_\nu| \\ &\leq |s_\mu - a_\mu| + |a_\mu - a_\nu| + |a_\nu - s_\nu|. \end{aligned}$$

This looks promising.

In fact, we can now prove A. Let $\varepsilon > 0$ be given. By our choice of s_μ, we have $|s_\mu - a_\mu| < 1/\mu$; for $\mu > 3/\varepsilon$, we therefore have $|s_\mu - a_\mu| < \varepsilon/3$. Similarly, we find $|s_\nu - a_\nu| < \varepsilon/3$ for $\nu > 3/\varepsilon$. Finally, there is an $N_0 \in \mathbb{N}$ such that $|a_\mu - a_\nu| < \varepsilon/3$ whenever $\mu, \nu > N_0$. Thus, for $N = \max\{N_0, \lfloor 3/\varepsilon \rfloor + 1\}$ and all $\mu, \nu > N$, we have

$$|s_\mu - s_\nu| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Since (s_n) is a Cauchy sequence of rational numbers, it converges to a real number a by Prop. 3.8. We now claim that $a = \lim a_n$. The triangle inequality gives $|a_n - a| \leq |a_n - s_n| + |s_n - a|$. We know $|a_n - s_n| < \frac{1}{n}$ by construction, so for $\varepsilon > 0$ we can pick $N_1 > 2/\varepsilon$, and then $|a_n - s_n| < \varepsilon/2$ for all $n > N_1$. Moreover, $|s_n - a| < \varepsilon/2$ for all n larger than some N_2 (since $a = \lim s_n$), hence $|a_n - a| < \varepsilon$ for all $n > \max\{N_1, N_2\}$, and this implies $a = \lim a_n$. This completes the proof of B. $\square$

Supremum and Infimum

Let $S \subset \mathbb{R}$ be a nonempty set of real numbers. We say that $M \in \mathbb{R}$ is an upper bound of S if $s \leq M$ for all $s \in S$. We say that M is a least upper bound (or supremum) if M is an upper bound of S and if there is no upper bound $m < M$ of S. The corresponding concept for lower bounds is the infimum of S (the greatest lower bound). We write $\sup S$ and $\inf S$ to denote the supremum and the infimum of S (in fact, if a supremum (infimum) exists, then it is unique: EXERCISE).

It is extremely important to understand the difference between the supremum and the maximum of a set: for $S = \{1 - \frac{1}{n} : n \in \mathbb{N}\}$, we have $\sup S = 1$, but $\max S$ does not exist.

Corollary 3.10. *Every nonempty set S of real numbers with an upper bound M has a supremum.*

Proof. Put $m_1 = M$ and pick $s_1 \in S$. If $m_1 = s_1$, then m_1 is a supremum of S (in fact, a maximum). If $m_1 > s_1$, then we consider the number $M_2 = \frac{1}{2}(m_1 + s_1)$. If $s \leq M_2$ for all $s \in S$, then M_2 is an upper bound, and we set $m_2 = M_2$ and $s_2 = s_1$; if not, then there is an $s_2 \in S$ with $s_2 > M_2$, and we put $m_2 = m_1$.

Assume that we have constructed upper bounds $m_1 \geq m_2 \geq \ldots \geq m_n$ and elements $s_1 \leq s_2 \leq \ldots \leq s_n$ of S with $m_k - s_k = \frac{1}{2}(m_{k-1} - s_{k-1})$ for all $1 \leq k \leq n$. Put $M_{n+1} = \frac{1}{2}(m_n + s_n)$. Again there are two cases:

- $s \leq M_{n+1}$ for all $s \in S$; then M_{n+1} is an upper bound for S, and we set $m_{n+1} = M_{n+1}$ and $s_{n+1} = s_n$. Observe that $m_{n+1} - s_{n+1} = M_{n+1} - s_n = \frac{1}{2}(m_n - s_n)$.
- $s > M_{n+1}$ for some $s \in S$; put $s_{n+1} = s$ and $m_{n+1} = m_n$. Observe that $m_{n+1} - s_{n+1} = m_n - s_{n+1} < m_n - M_{n+1} = \frac{1}{2}(m_n - s_n)$.

We note for later use that

$$0 \leq m_{n+1} - s_{n+1} \leq \frac{1}{2}(m_n - s_n). \tag{3.4}$$

Clearly (m_n) is a monotone decreasing sequence, and since $m_n \geq s$ for all $s \in S$, we have $m_n \geq s_1$, that is, (m_n) is bounded from below by s_1. But monotone decreasing sequences bounded from below are Cauchy, and since $\mathbb{R}$ is complete, (m_n) converges; put $m = \lim m_n$.

Similarly, (s_n) is a monotone increasing sequence bounded from above by m_1, hence it is Cauchy, so converges in $\mathbb{R}$. Put $s = \lim s_n$; we claim that $m = s$. In fact, applying the limit to the inequalities (3.4), we get $0 \leq m - s \leq \frac{1}{2}(m - s)$. Subtracting $\frac{1}{2}(m - s)$ from $m - s \leq \frac{1}{2}(m - s)$ gives $\frac{1}{2}(m - s) \leq 0$; on the other hand, $\frac{1}{2}(m - s) \geq 0$ since $m - s \geq 0$, hence $m = s$.

Now we claim that m is a least upper bound of S. There are two things to prove:

- m is an upper bound. Assume it isn't. Then there is an $s_0 \in S$ such that $s_0 > m$. Put $\varepsilon = s_0 - m$. Then there is an integer N such that $|m_n - m| < \varepsilon$; moreover, $s_0 \leq m_n$ since by construction, the m_n were upper bounds for S. But now

  $$0 \leq m_n - s_0 = m_n - m + m - s_0 < \varepsilon - \varepsilon = 0,$$

  which is a contradiction.
- There is no upper bound $m' < m$ of S. For assume there is; then $s \leq m'0$ for all $s \in S$, in particular $s_n \leq m'$ for the elements of the sequence (s_n); but then $\lim s_n \leq m' < m$, contradicting the fact that $\lim s_n = m$.

□

Decimal Expansion

Fix a natural number $g \geq 2$. Given any sequence (d_n) of integers $d_n \in \{0, 1, \ldots, g\}$, the sequence of rational numbers

$$s_n = \sum_{k=1}^{n} d_k g^{-k}$$

is a Cauchy sequence. In fact, if $m > n > N$ for some $N \in \mathbb{N}$, then

$$\begin{aligned} |s_m - s_n| &= \sum_{k=n+1}^{m} d_k g^{-k} \leq \sum_{k=n+1}^{m} (g-1) g^{-k} < (g-1) \sum_{k=N+1}^{\infty} g^{-k} \\ &= (g-1) g^{-N} \sum_{k=+1}^{\infty} g^{-k} = (g-1) g^{-N} \frac{1}{g-1} = g^{-N} \end{aligned}$$

Thus if we are given an $\varepsilon > 0$, and if we choose N so large that $g^{-N} < \varepsilon$, then $|s_m - s_n| < \varepsilon$ for all $m, n > N$. This means that all numbers of the form $\sum d_k g^{-k}$ are real numbers, and in fact they are real numbers in the interval $[0, 1]$.

Conversely, let $r \in [0, 1)$ be a real number. We claim that it has a g-adic expansion. Assume for the moment we already have what we are looking for: $r = 0.d_1 d_2 d_3 \ldots$; then $gr = d_1.d_2 d_3 \ldots$, hence $d_1 = \lfloor gr \rfloor$. Turning this observation around we see how we have to proceed.

In fact, put $d_1 = \lfloor gr \rfloor$ and $r_1 = gr - d_1$; then $d_1 \in \{0, 1, \ldots, g-1\}$ and $r_1 \in [0, 1)$. The first claim follows from d_1 being an integer by definition and the inequality $0 \leq gr < g$; the second claim follows from the definition of the floor function. This allows us to define a sequence (d_n) of integers $d_n \in \{0, 1, \ldots, g-1\}$ inductively: assume that we have found $d_1, \ldots, d_n$, and that $r_n \in [0, 1)$; then put $d_{n+1} = \lfloor g r_n \rfloor$ and $r_{n+1} = g r_n - d_n$.

What remains to show is that we do in fact have $r = 0.d_1 d_2 d_3 \ldots$. By construction, $0 \leq gr - d_1 < 1$; using induction, we can prove $0 < g^n r - (d_1 g^{n-1} + d_2 g^{n-2} + \ldots + d_n) < 1$, which is equivalent to $0 < r - 0.d_1 d_2 \ldots d_n < g^{-n}$. This implies the claim, and we have proved:

Theorem 3.11. *Every real number has a decimal expansion.*

It is a bit unfortunate that the decimal expansion of real numbers is not always unique:

Proposition 3.12. *We have* $0.99999\ldots = 1$.

Proof. By definition,

$$0.99999\ldots = \sum_{n=1}^{\infty} 9 \cdot 10^{-n} = 9 \sum_{n=1}^{\infty} 10^{-9} = 9 \frac{1}{10-1} = 1.$$

□

In fact we have

Proposition 3.13. *Every real number either has a unique g-adic expansion, or it has two: one that terminates and another one ending in infinitely many $g-1$'s.*

Proof. If $r = 0.d_1d_2\ldots d_n > 0$ has a terminating g-adic expansion (with $d_n \neq 0$), then we also have $r = 0.d_1d_2\ldots(d_n-1)(g-1)(g-1)(g-1)\ldots$. The harder problem is showing that reals have at most two g-adic expansions.

It is sufficient to do this for reals in the interval $[0,1]$. So assume that $r = 0.d_1d_2d_3\ldots = 0.e_1e_2e_3\ldots$ are two different g-adic expansions of $r \in [0,1]$, and let m denote the smallest index such that $d_m \neq e_m$. We may assume that $d_m < e_m$.

We claim that $d_{m+1} = d_{m+2} = \ldots = g-1$; in fact, if $d_{m+k} < g-1$, then using the geometric series we get

$$\begin{aligned} r &= \sum_{j=1}^{m} d_m g^{-j} + \sum_{j=m+1}^{\infty} d_j g^{-j} \\ &< \sum_{j=1}^{m} d_m g^{-j} + \sum_{j=m+1}^{\infty} (g-1) g^{-j} \\ &= \sum_{j=1}^{m} d_m g^{-j} + g^{-m} \\ &= 0.d_1d_2\ldots(d_m+1) \\ &\leq 0.e_1e_2\ldots e_m \ < \ r, \end{aligned}$$

which is a contradiction.

In a similar way, we can prove that $e_{m+k} = 0$ for all $k \geq 1$. Thus if a real number has two different g-adic expansions, then one of them is terminating, and there is exactly one other expansion, namely one terminating with $0.d_1\ldots d_m(g-1)(g-1)(g-1)\ldots$. □

Theorem 3.14 (Cantor). *The set of real numbers $(0,1)$ is not countable.*

Proof. Our only chance is a proof by contradiction, so let us assume that we can enumerate the real numbers in $(0,1)$, that is, assume that the sequence of real numbers $r_1, r_2, r_3, \ldots$ contains each and every element of $(0,1)$. Now we write each of these real numbers in decimal notation:

$$\begin{aligned} r_1 &= 0.r_{11}r_{12}r_{13}\ldots, \\ r_2 &= 0.r_{21}r_{22}r_{23}\ldots, \\ r_3 &= 0.r_{31}r_{32}r_{33}\ldots, \end{aligned}$$

and make once more use of the diagonal to construct another real number. What we do is define a real number

$$s = 0.s_1s_2s_3\ldots$$

by putting

$$s_n = \begin{cases} r_{nn} + 1 & \text{if } r_{nn} \leq 5 \\ r_{nn} - 1 & \text{if } r_{nn} \geq 6 \end{cases}$$

This is a genuine real number because $s_n \in \{0, 1, \ldots, 9\}$, and clearly $s \in [0, 1]$. In fact, $s \neq 0, 1$ because $5 \leq s_n \leq 8$.

Now what is so special about our number s? Well, for one thing, $s \neq r_1$, because these numbers differ at the first decimal: $s_1 = r_{11} \pm 1 \neq r_{11}$. Similarly, $s \neq r_2$ because $s_2 \neq r_{22}$, and in general, $s \neq r_n$ because $s_n = r_{nn} \pm 1 \neq r_{nn}$. But this means that our real number s does not occur in the sequence $r_1, r_2, r_3, \ldots$: contradiction. $\square$

Exercise. Every rational number has a decimal expansion; we know that rationals are countable. Where does the proof above go wrong if we attempted to use it to prove that rationals are not countable?

Remark. Given sets A and B, we say that A and B have the same cardinality if there is a bijection between A and B. We say that B has greater cardinality than A if they do not have the same cardinality and if there is a bijection between A and a subset of B. The fact that $\mathbb{Q}$ is countable means that $\mathbb{N}$ and $\mathbb{Q}$ have the same cardinality. The theorem above shows that there are many more real than rational numbers, or more exactly that the cardinality of $\mathbb{R}$ is greater than that of $\mathbb{Q}$. Cantor, who invented the theory of countable sets and cardinal numbers, also asked the following apparently simple question: is there a subset C of $\mathbb{R}$ with $\mathbb{Q} \subset C \subset \mathbb{R}$ such that the cardinality of C is larger than that of $\mathbb{Q}$ but smaller than that of $\mathbb{R}$?

Cantor conjectured that there is no such set (this conjecture became known as the continuum hypothesis; here continuum is Cantor's word for the reals) and thought several times that he had found a proof, but each time found a gap. The striking result of Gödel and Cohen[3] is that this theorem is independent of ZFC, the axiom system of set theory that is currently accepted (named after Zermelo and Frankel, with C representing the axiom of choice). What that means is that not only is it impossible to prove the continuum hypothesis within this system of axioms, it is also impossible to prove that it is wrong. One possible way of making some sense of this affair is to say that our current set of axioms is not strong enough to decide Cantor's question.

Finally let me point out that there is no intuition to guide you through the theory of cardinalities, essentially because the human mind is not used to dealing with infinities. Here's a striking example:

Theorem 3.15. *The interval* $[0, 1) \subset \mathbb{R}$ *and the unit square* $[0, 1) \times [0, 1) \subset \mathbb{R}^2$ *have the same cardinality.*

[3] Gödel proved in 1938 that assuming the continuum hypothesis does not contradict ZFC, that is, you can't prove that it is wrong from ZFC; in 1963, Paul Cohen showed that you also can't prove the continuum hypothesis from ZFC.

Naively, one would have expected the 2-dimensional square to have many more points than the interval. When Cantor stumbled across this result, he sent off a letter to Dedekind and wrote 'I see it, but I don't believe it'. The idea of the proof, as a matter of fact, is so simple that it almost hurts: we have to find a bijection between the two sets above. Write all real numbers as decimals (not allowing them to end with an infinite string of 9's to make the representation unique). Then we define a map $[0,1) \longrightarrow [0,1) \times [0,1)$ by

$$0.a_1a_2a_3\ldots \longmapsto (0.a_2a_4a_6\ldots, 0.a_1a_3a_5\ldots).$$

This map is obviously surjective; unfortunately it is not injective, since e.g.

$$0.01010101\ldots \longmapsto (0, 0.19999\ldots)$$

and

$$0.02 \longmapsto (0, 0.2).$$

Theorem 3.16. *For any real number $a > 0$, there is exactly one real $b > 0$ with $b^2 = a$.*

The element b in the above theorem is denoted by $\sqrt{a}$.

Proof. Consider the sequence (a_n) defined by $a_1 = 1$ and

$$a_{n+1} = \frac{1}{2}\Big(a_n + \frac{a}{a_n}\Big).$$

We find that

$$a_{n+1}^2 - a = \frac{1}{4}\Big(a_n - \frac{a}{a_n}\Big)^2$$

for all $n \geq 1$; in particular, we have $a_n^2 \geq a$ for all $n \geq 2$.

$$a_{n+1}^2 = \frac{1}{4}a_n^2 + \frac{a}{2} + \frac{a}{4a_n^2} < \frac{1}{4}a_n^2 + \frac{3a}{4},$$

hence

$$0 \leq a_{n+1}^2 - a < \frac{1}{4}(a_n^2 - a).$$

Thus a_n^2 converges to a. Finally,

$$a_n - a_{n+1} = \frac{1}{2a_n}(a_n^2 - b),$$

and since a_n is bounded and $|a_n^2 - b| < 4|a_1^2 - a|4^{-n}$, (a_n) is Cauchy, hence converges. This gives $a = \lim a_n^2 = (\lim a_n)^2$ and proves the claim. $\square$

Proposition 3.17. *If $a, b \in \mathbb{R}$ are positive, then $\sqrt{ab} = \sqrt{a}\sqrt{b}$.*

Proof. By Theorem 3.16, it is sufficient to prove that $\sqrt{a}\sqrt{b}$ is the positive root of $y = x^2 - ab$. But this follows from what we know about multiplication of real numbers. $\square$

We remark that the inequality $x^2 \geq 0$ has some beautiful and useful consequences. A very important corollary is the inequality between arithmetic and geometric means:

Proposition 3.18. *For real numbers $a, b \geq 0$, we have $\frac{a+b}{2} \geq \sqrt{ab}$, with equality if and only if $a = b$.*

The proof is left as homework. The inequality is powerful enough to help us solve some simple minimax problems.

Exercise. Among the rectangles with circumference c, determine the one with maximal area.

Let x and y denote the sides of the rectangle; then $x + y = \frac{c}{2}$. We have to maximize the area $A = xy$. By the inequality of arithmetic and geometric means, we have $\sqrt{A} = \sqrt{xy} \leq \frac{x+y}{2} = \frac{1}{4}c$, hence $A \leq \frac{1}{2}\sqrt{c}$ with equality if and only if $x = y$. Thus the rectangle with given circumference and maximal area is the square.

The following result is often useful:

Proposition 3.19 (Sandwich Lemma). *Let (a_n) and (b_n) be sequences of real numbers with the same limit s. If (s_n) is a sequence with $a_n \leq s_n \leq b_n$ for all n larger than some $N \in \mathbb{N}$, then s_n converges, and $\lim s_n = s$.*

Proof. Let an $\varepsilon > 0$ be given. Then there are integers N_1, N_2 such that $|a_n - s| < \varepsilon$ for $n > N_1$ and $|b_n - s| < \varepsilon$ for $n > N_2$. Now $-\varepsilon < a_n - s \leq s_n - s \leq b_n - s < \varepsilon$ shows that $|s_n - s| < \varepsilon$ for all $n > \max\{N_1, N_2\}$. $\square$

e

We now can define Euler's number

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots = \sum_{n=0}^{\infty} \frac{1}{k!},$$

but of course we have to show that this series converges. Since the sequence of partial sums $e_k = \sum_{n=0}^{k} \frac{1}{k!}$ is monotone, it suffices to show that the sequence is bounded. This is quite pretty: just observe that $\frac{1}{n!} = \frac{1}{n(n-1)\cdots 1} \leq \frac{1}{n(n-1)}$, and

the 'do the telescope':

$$\begin{aligned} e_k &= \sum_{n=0}^{k} \frac{1}{k!} \leq 2 + \sum_{n=2}^{k} \frac{1}{n(n-1)} \\ &= 2 + \sum_{n=2}^{k} \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 2 + 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \ldots - \frac{1}{n-1} + \frac{1}{n-1} - \frac{1}{n} \\ &= 2 + 1 - \frac{1}{n} < 3. \end{aligned}$$

For those who are tired of seeing proofs that $\sqrt{2}$ is irrational, let us prove that

Proposition 3.20. *The number e is irrational.*

Proof. Of course we do that by contradiction. Assume that $e = \frac{p}{q}$. Then $q!e = (q-1)! \cdot p$ is an integer. On the other hand,

$$q!e = q! + q! + \frac{q!}{2!} + \frac{q!}{3!} + \ldots + \frac{q!}{q!} + \frac{q!}{(q+1)!} + \ldots$$

is an integer plus

$$X = \frac{q!}{(q+1)!} + \frac{q!}{(q+2)!} + \ldots = \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \frac{1}{(q+1)(q+2)(q+3)} + \ldots.$$

We will have reached a contradiction if we can prove that $0 < X < 1$. Clearly $X > 0$; on the other hand,

$$\begin{aligned} X &< \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \frac{1}{(q+2)(q+3)} + \ldots \\ &= \frac{1}{q+1} + \frac{1}{q+1} - \frac{1}{q+2} + \frac{1}{q+2} - \frac{1}{q+3} + \ldots \\ &= \frac{2}{q-1}, \end{aligned}$$

hence $X < 1$ if $q > 3$. But for $q = 1, 2, 3$ it is easily seen by inspection that $e = 2.71828\ldots \neq p/q$. □

3.2 The Complex Numbers

Constructing complex numbers from the reals is really easy again (the hard part was constructing $\mathbb{R}$ from $\mathbb{Q}$): consider the set

$$\mathbb{C} = \{(x, y) : x, y \in \mathbb{R}\}$$

of pairs of real numbers; we call any such pair a complex number. We define addition and multiplication on $\mathbb{C}$ by

$$(r,s)+(t,u) = (r+t, s+u), \tag{3.5}$$
$$(r,s)\cdot(t,u) = (rt-su, ru+st). \tag{3.6}$$

Since we have defined these operations on elements (not on equivalence classes), there is no well-definedness to prove here. We identify $x \in \mathbb{R}$ with the $(x,0)$; it is then easy to check that $\mathbb{C}$ is a field containing $\mathbb{R}$ as a subfield. More exactly: consider the map $\iota : \mathbb{R} \longrightarrow \mathbb{C}$ defined by $\iota(x) = (x,0)$. Then ι is an injective ring homomorphism, i.e., $\iota(x+y) = \iota(x)+\iota(y)$, $\iota(xy) = \iota(x)\iota(y)$, $\iota(1) = (1,0)$ is the multiplicative identity in $\mathbb{C}$.

We now introduce the abbreviations $1 = (1,0)$ and $i = (0,1)$; then $(x,y) = (x,0)+(0,x) = x\cdot 1+y\cdot i =: x+yi$. Note that $i^2 = -1$ since $(0,1)\cdot(0,1) = (-1,0)$.

It can be shown that $\mathbb{C}$ cannot be ordered in such a way that $z_1 > 0$ and $z_2 > 0$ imply $z_1 z_2 > 0$. On the other hand, we can easily define an absolute value on $\mathbb{C}$ by putting

$$|x+iy| = \sqrt{x^2+y^2}.$$

Proposition 3.21. *For $z, w \in \mathbb{C}$ we have*

1. $|z| \geq 0$, *with equality if and only if* $z = 0$;
2. $|zw| = |z|\cdot|w|$;
3. *Triangle inequality:* $|z+w| \leq |z|+|w|$.

The proof is left as an exercise. Using $|\cdot|$, it is easy to define Cauchy sequences exactly as in $\mathbb{R}$. As a matter of fact, even the proofs carry over without changing a single word; thus we have

Theorem 3.22. *Every Cauchy sequence in $\mathbb{C}$ converges. The set C of Cauchy sequences in $\mathbb{C}$ forms a ring, and* $\lim : C \longrightarrow \mathbb{C}$ *is a ring homomorphism.*

Given any sequence (z_n) of complex numbers, we can form two sequences of real numbers by taking the real and the imaginary part: we simply write $z_n = x_n + iy_n$ with $x_n, y_n \in \mathbb{R}$. What does it mean in terms of (x_n) and (y_n) that (z_n) is Cauchy? Well, the connection is a simple as it can be:

Proposition 3.23. *The sequence (z_n) with $z_n = x_n + iy_n$ is a Cauchy sequence if and only if (x_n) and (y_n) are; in that case,* $\lim z_n = \lim x_n + i \lim y_n$.

Most of the results that we prove for sequences and series of real numbers also hold for complex numbers; the only things that can possibly go wrong are those connected with the order on $\mathbb{R}$, because there is nothing like this in $\mathbb{C}$. Consider e.g. the proof of the existence of square roots of positive reals in $\mathbb{R}$: the proof there fails to hold in $\mathbb{C}$ because we used the order of $\mathbb{R}$. Nevertheless, square roots exist, even for *all* complex numbers:

Proposition 3.24. *For every $z \in \mathbb{C}$ there is a $w \in \mathbb{C}$ such that $w^2 = z$. This number is unique up to sign, that is, w and $-w$ are the only complex numbers with $w^2 = z$.*

Proof. We reduce everything to $\mathbb{R}$ by writing $z = x + iy$ and $w = s + ti$. The equation $w^2 = z$ gives $s^2 - t^2 = x$ and $2st = y$. There are two cases:

- $y = 0$. Then $s = 0$ or $t = 0$.

  If $x < 0$, the first case must hold because of $s^2 - t^2 = x$, and then $t^2 = -x > 0$ has a solution in $\mathbb{R}$. But then $w = 0 + it$ solves our problem.

  If $x > 0$, the second case must hold, and then $s^2 = x > 0$ has a solution in $\mathbb{R}$, so $w = s + 0t$ does it.

- $y \neq 0$. Then $s, t \neq 0$, and plugging $t = y/2s$ into $s^2 - t^2 = x$ gives $s^4 - y^2/4 = xs^2$. This is a biquadratic equation that can be solved easily; first, we put $u = s^2$ and get $u^2 - xu - y^2/4 = 0$, hence $u_{1,2} = \frac{1}{2}(x \pm \sqrt{x^2 + y^2}\,)$. Since we want to extract the square root of u, we must make sure that the solution we get is nonnegative. But $x + \sqrt{x^2 + y^2} > 0$ because $|x| < \sqrt{x^2 + y^2}$. Thus we set

  $$s = \sqrt{\frac{x + \sqrt{x^2 + y^2}}{2}},$$

  then solve $2st = y$ for t, and then verify that $w = s + it$ does what we want.

This proves that square roots of complex numbers always exist. Moreover, if $w^2 = z$, then also $(-w)^2 = z$, so for $z \neq 0$ there are always at least two square roots. But there can't be more than two: If $w_k^2 = z$ for $k = 1, 2$, then $w_1^2 = w_2^2$, hence $(w_1 - w_2)(w_1 + w_2) = 0$. Since $\mathbb{C}$ is a field, a product is 0 if and only if one of its factors is 0, and this implies $w_1 = w_2$ or $w_1 = -w_2$. $\square$

Note that we have used the uniqueness of positive square roots in $\mathbb{R}$ to prove that $\sqrt{ab} = \sqrt{a}\sqrt{b}$. In $\mathbb{C}$, not only does our proof not carry over, not even the result does:

$$-1 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1) \cdot (-1)} = \sqrt{1} = 1,$$

where the first and the last two equality signs hold without doubt, showing that the second equality sign does not hold in this case.

3.3 The p-adic numbers $\mathbb{Q}_p$

The importance of the construction of the reals by 'adjoining' the limits of Cauchy sequences to $\mathbb{Q}$ lies with the fact that the same method can be used to construct other fields containing $\mathbb{Q}$ that are fundamentally different from the real numbers (or, if you don't like number theory: there is a whole branch of

topology dealing with generalizing the above to the construction of completions of uniform spaces; uniform spaces are topological spaces that behave very much like metric spaces, and in fact every metric space is uniform). Actually, the construction itself is exactly the same, the only difference being that the absolute value $|\cdot|$ is replaced by other functions with similar properties. Let us introduce these functions now.

Fix a prime number $p \geq 2$; then every nonzero rational number x can be written uniquely in the form $x = p^a y$, where a is an integer and $y = \frac{r}{s}$ is a fraction with $p \nmid rs$ (we have extracted the power of p from x). Now we define $|x|_p = p^{-a}$ and $|0|_p = 0$. It is seen directly from the definition that $|x|_p$ is small if and only if the numerator of x (when written in lowest terms) is divisible by a high power of p. In particular, $|p|_p = \frac{1}{p}$, $|p^2|_p = \frac{1}{p^2}$ etc.

This p-adic valuation shares the following properties with the usual absolute value:

Proposition 3.25. *For $x, y \in \mathbb{Q}$, we have*

1. $|x|_p \geq 0$, *with equality if and only if* $x = 0$;
2. $|xy|_p = |x|_p |y|_p$;
3. $|x + y|_p \leq |x|_p + |y|_p$; *in fact,* $|\cdot|_p$ *satisfies the stronger (ultrametric) triangle inequality* $|x + y|_p \leq \max\{|x|_p, |y|_p\}$.

Proof. Exercise. □

Thus we have a whole family of absolute values, and in order to be able to make statements about all of them at the same time we put $|\cdot|_\infty := |\cdot|$.

Using any absolute value with this property, we can define a distance on $\mathbb{Q} \times \mathbb{Q}$ by putting

$$d_p(x, y) := |x - y|_p.$$

The usual distance between rational numbers is given by $d_\infty(x, y) = |x - y|$.

This 'distance function' is called a metric: it has the following properties:

Proposition 3.26. *For all $x, y, z \in \mathbb{Q}$ we have*

1. $d_p(x, y) \geq 0$, *with equality if and only if* $x = y$;
2. $d_p(x, y) = d_p(y, x)$;
3. $d_p(x, z) \leq d_p(x, y) + d_p(y, z)$.

The simple proof is left as an exercise. We say that any of these functions d_p make $\mathbb{Q}$ into a metric space.

Now the definition of a Cauchy sequence makes sense in any metric space: a sequence (a_n) is Cauchy if for every $\varepsilon > 0$ there is an integer N such that $d_p(a_m, a_n) < \varepsilon$ for all $m, n > N$.

In fact, for $p = \infty$, this is the usual definition of Cauchy sequences. For $p \neq \infty$, these Cauchy sequences are not at all like the Cauchy sequences we are

used to: for example, $1, p, p^2, p^3, \ldots$ is a Cauchy sequence with respect to d_p; in fact, this sequence converges to 0.

Question: does every Cauchy sequence with respect to d_p converge in $\mathbb{Q}$? Well, the answer is no, and the proof isn't hard using a little number theory. As a matter of fact, not even Cauchy sequences in $\mathbb{Z}$ converge necessarily whenever $p \neq \infty$

Now we can construct the completion of $\mathbb{Q}$ with respect to any of these functions d_p by looking at equivalence classes of Cauchy series modulo Cauchy series converging to 0, and the result turns out to be a field for every p: for $p = \infty$, we get back $\mathbb{R}$, and for primes p we find the field $\mathbb{Q}_p$ of p-adic numbers.

Just as real numbers have a decimal expansion, p-adic numbers can be represented as

$$a_{-n}p^{-n} + \ldots + a_{-1}p^{-1} + a_0 + a_1 p + a_2 p^2 + \ldots = \sum_{k \geq -n} a_k p^k$$

with $a_k \in \{0, 1, \ldots, p-1\}$. These fields are very important fields in number theory. It can be shown, for example, that for odd primes p not dividing an integer a, the field $\mathbb{Q}_p$ contains $\sqrt{a}$ if and only of a is a quadratic residue modulo p.

3.4 Bolzano-Weierstrass

The theorem of Bolzano-Weierstrass asserts that any bounded sequence has a convergent subsequence. The assumption that the sequence is bounded cannot be removed: $a_n = n$ defines a sequence not containing a convergent subsequence. On the other hand, there are sequences containing a whole lot converging subsequences: the sequence $1,\ \frac{1}{2},\ 2,\ 3,\ 1,\ \frac{1}{3},\ \frac{1}{4},\ \ldots$ by which we enumerated the positive rationals has a subsequence converging to any given $\frac{p}{q}$: just take $\frac{p+1}{q+1}$, $\frac{2p+1}{2q+1}$, $\frac{3p+1}{3q+1}$, $\ldots$.

Theorem 3.27. [4] *Every sequence (s_n) of real numbers contains a monotone subsequence.*

Proof. We call the term s_n dominant if it is greater than every term that follows, i.e. if $s_n > s_m$ for all $m > n$. Now there are two cases:

1. (s_n) contains infinitely many dominant terms $s_{n_1}, s_{n_2}, \ldots$; then the s_{n_i} form a monotone decreasing subsequence of (s_n).

2. (s_n) contains only finitely many dominant terms. Then there is an $N \in \mathbb{N}$ such that there are no dominant terms s_n with $n > N$. Pick any $n_1 > N$. Since s_{n_1} is not dominant, there is an $n_2 > n_1$ such that $s_{n_1} \leq s_{n_2}$. For the same reason, there is an $s_3 > s_2$ such that $s_{n_2} \leq s_{n_3}$. Continuing in this way we get a monotone increasing subsequence (s_{n_i}).

[4] Ross, Thm. 11.3.

□

Theorem 3.28 (Bolzano-Weierstrass). *Every bounded sequence of real numbers has a convergent subsequence.*

Proof. Since the sequence is bounded, it contains a monotone subsequence. This subsequence is Cauchy by Theorem 2.15, hence it converges. □

Remark. This result was frequently ascribed to Weierstrass until it was discovered that it already had been published by Bolzano in 1817.

Let us sketch a second proof for the theorem of Bolzano-Weierstrass. Assume that $a_1 < s_n < b_1$ for all $n \in \mathbb{N}$ and constants $a_1, b_1 \in \mathbb{R}$ (these exist because the sequence is bounded).

Now $[a_1, b_1]$ contains infinitely (namely all) s_n. If we split this interval into two parts, at least one of them must contain infinitely many s_n. If $[a_1, \frac{1}{2}(a_1 + b_1)]$ contains infinitely many points, we put $a_2 = a_1$ and $b_1 = \frac{1}{2}(a_1 + b_1)$, otherwise we put $a_2 = \frac{1}{2}(a_1+b_1)$ and $b_2 = b_1$. Repeating this game we construct inductively a sequence of numbers $a_n < b_n$ such that $[a_n, b_n]$ contains infinitely many s_n. Clearly (a_n) is monotone increasing and bounded, whereas (b_n) is monotone decreasing and bounded. Thus (a_n) and (b_n) converge. Moreover, $|a_{n+1} - b_{n+1}| < \frac{1}{2}|a_n - b_n|$, hence $(a_n - b_n)$ converges to 0, and we conclude that $\lim a_n = \lim b_n$.

Now where's our subsequence? Well, just pick some $s_{n_1} \in [a_1, b_1]$, $s_{n_2} \in [a_2, b_2]$ with $n_2 > n_1, \ldots$ (each of these intervals contains infinitely many s_n, so this is possible). Then $a_k \leq s_{n_k} \leq b_k$, hence s_{n_k} converges (to $\lim a_k$) by the sandwich theorem.

The theorem of Bolzano-Weierstrass also holds for complex numbers.

Theorem 3.29. *Each bounded sequence of complex numbers contains a converging subsequence.*

Proof. Let (z_n) be a bounded sequence of complex numbers. Then $z_n = x_n + iy_n$, and since $|z_n| < M$ for some constant M, we find $x_n^2 + y_n^2 < M^2$, so in particular $|x_n|, |y_n| < M$. The first idea now is to apply the real Bolzano-Weierstrass to the sequences (x_n) and (y_n), but this doesn't work: if (x_{n_i}) and (y_{n_k}) are converging subsequences, then $x_{n_i} + y_{n_i}$ need not be a subsequence of (z_n); for example, we might have $x_{n_1} = x_1$ and $y_{n_1} = y_2$.

Thus we have to be a little more careful: the real Bolzano-Weierstrass gives us a converging subsequence (x_{n_i}). Now consider the sequence (y_{n_i}) defined by $z_{n_i} = x_{n_i} + y_{n_i}$. This is a bounded sequence of real numbers, hence contains a converging subsequence $y_{n_{i_k}}$. But now $z_{n_{k_i}} = x_{n_{k_i}} + y_{n_{k_i}}$ is a convergent subsequence of (z_n) because the real and imaginary parts converge.

Instead of applying the real Bolzano-Weierstrass we also could have used the idea of our second proof: In the real case, we kept on dividing the interval $[a_1, b_1]$ into smaller and smaller parts. In the complex case, we have a rectangle $[a_1, b_1] \times [c_1, d_1]$ containing the real and imaginary parts of (z_n): $a_1 < x_n < b_1$ and $c_1 < y_n < d_n$. This follows easily from the fact that (z_n) is bounded. But

now we divide this rectangle into four equal parts; at least one of them contains infinitely many (z_n), and we take this as our second rectangle. The rest of the proof now goes through easily. $\square$

3.5 Absolute Convergence

Consider the series

$$S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - + \dots.$$

Since this is an alternating series whose terms converge to 0, it converges, and it is immediately clear that $0 < S < 1$: indeed, each of the terms $1 - \frac{1}{2}$, $\frac{1}{3} - \frac{1}{4}$ etc is positive, and each of the terms $-\frac{1}{2} + \frac{1}{3}$, $-\frac{1}{4} + \frac{1}{5}$ etc. is negative. As a matter of fact it can be shown that $S = \log 2$, where log is the natural logarithm.

Now we consider $2S$ and rearrange a bit:

$$\begin{aligned}
2S &= \frac{2}{1} - \frac{2}{2} + \frac{2}{3} - \frac{2}{4} + \frac{2}{5} - \frac{2}{6} \pm \dots \\
&= \left(\frac{2}{1} - \frac{2}{2}\right) - \frac{2}{4} + \left(\frac{2}{3} - \frac{2}{6}\right) - \frac{2}{8} + \left(\frac{2}{5} - \frac{2}{10}\right) - \frac{2}{12} \pm \dots \\
&= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} \pm \dots \\
&= S,
\end{aligned}$$

and since $S \neq 0$, we can cancel and get $2 = 1$. Note that we have not cheated: every summand was taken care of in the rearrangement! In fact, the $\frac{2}{m}$ and the $\frac{2}{2m}$ with m odd form the pairs $(\frac{2}{m} - \frac{2}{2m}) = \frac{1}{m}$, whereas the $\frac{2}{m}$ with m divisible by 4 end up as the negative terms in the rearrangement.

Once you know that $1 = 2$, you can prove a whole lot of things; it is known that if A and B are finite sets with $A \subseteq B$ and $\#A = \#B$, then $A = B$. Using that (and $1 = 2$) you can prove that you are the Pope by considering the sets $A = \{\text{you}\}$ and $B = \{\text{you, Pope}\}$. Clearly $A \subseteq B$, $\#A = 1 = 2 = \#B$, hence $A = B$ qed.

Similarly you can prove that any theorem is true if $1 = 2$, which means that the whole of mathematics would be utterly useless. Now fortunately our 'proof' above isn't a proof at all: while we may rearrange (finite) sums $a_1 + \dots + a_n$ by using the laws of associativity and commutativity that we have proved, we have proved no such thing for series! In fact we shall show below that we may rearrange series if and only if they are absolutely convergent.

A series $\sum a_n$ is called absolutely convergent if $\sum |a_n|$ converges. Clearly the sequence above is not absolutely convergent since $\sum |(-1)^{n+1}\frac{1}{n}| = \sum \frac{1}{n}$ is the harmonic series that we know to be divergent. Converging series that are not absolutely convergent are called conditionally convergent. Now we can state the following two theorems:

Theorem 3.30. *If (a_n) is a sequence, (b_n) a rearrangement of (a_n), and if $\sum a_n$ is absolutely convergent, then so is $\sum b_n$, and we have $\sum a_n = \sum b_n$.*

Theorem 3.31 (Riemann). *If $\sum a_n$ is conditionally convergent, then for any real number $s \in \mathbb{R}$ there is a rearrangement (b_n) of (a_n) such that $\sum b_n = s$.*

We call (b_n) a rearrangement of (a_n) if there is a bijection $\pi : \mathbb{N} \longrightarrow \mathbb{N}$ with $a_n = b_{\pi(n)}$ (what this means is the sets $\{a_n : n \in \mathbb{N}\}$ and $\{b_n : n \in \mathbb{N}\}$ are equal, but that their indices are permuted, so the terms appear in a possibly different order.

Proof of Thm. 3.30. Let $s_n = \sum_{k=1}^n a_n$ and $t_n = \sum_{k=1}^n b_n$ denote the partial sums of the two series. Let $\varepsilon > 0$; since $\sum a_n$ converges absolutely, there is an $N \in \mathbb{N}$ such that

$$\sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

In particular,

$$|s_N - s| = \Big| \sum_{k=N+1}^{\infty} a_k \Big| \leq \sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

Let $\pi : \mathbb{N} \longrightarrow \mathbb{N}$ be the bijection such that $a_n = b_{\pi(n)}$, and put $M = \max\{f(1), \ldots, f(N)\}$. Then we have

$$\{a_1, \ldots, a_N\} \subseteq \{b_1, \ldots, b_M\}.$$

Now for any $m \geq M$, the sum $t_m - s_N$ contains only a_k's with $k > N$. Thus

$$|t_m - s_N| \leq \sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

This implies

$$|t_m - s| \leq |t_m - s_N| + |s_N - s| < \varepsilon$$

for all $m \geq N$, and this shows that (b_n) converges absolutely to $\lim t_m = s$. $\square$

Proof of Thm. 3.31. Assume that $\sum a_n$ converges, but not absolutely; define two sequences (r_n) and (s_n) by

$$r_n = \begin{cases} a_n & \text{if } a_n \geq 0 \\ 0 & \text{if } a_n < 0, \end{cases} \quad \text{and} \quad s_n = \begin{cases} 0 & \text{if } a_n \geq 0 \\ -a_n & \text{if } a_n < 0. \end{cases}$$

We claim that $\sum r_n$ and $\sum s_n$ both diverge. In fact, if they both would converge, then so would $\sum r_n - \sum s_n = \sum |a_n|$. If $\sum r_n$ would converge and $\sum s_n$ wouldn't, then $\sum_{n=1}^k a_n = \sum_{n=1}^k r_n - \sum_{n=1}^k s_n$ would show that $\sum a_n$ would not converge, contrary to assumption.

Thus both $\sum r_n$ and $\sum s_n$ diverge, and since $r_n, s_n \geq 0$, the subsequences $\sum_{n=1}^k r_n$ and $\sum_{n=1}^k s_n$ are monotone, $\sum r_n$ and $\sum s_n$ can't be bounded.

The idea behind Riemann's proof is now very simple: assume you want to rearrange the a_n in such a way that the resulting series converges to a given real number $r \geq 0$ (the case $r \leq 0$ follows upon replacing a_n by $-a_n$). Here's what you do: there is an index N_1 such that $B_1 = \sum_{n=1}^{N_1} r_n > r$ because $\sum r_n$ is not bounded; now you subtract terms $s_1, s_2, \ldots, s_{M_1}$ until the sum becomes smaller than r, then you add terms $r_{N_1+1}, \ldots, r_{N_2}$ until it becomes larger than r, and so on. This way you get a rearrangement $b_1 = r_1, \ldots, b_{N_1} = r_{N_1}$, $b_{N_1+1} = -s_1, \ldots, b_{N_1+M_1} = -s_{M_1}$, $b_{N_1+M_1+1} = r_{N_1+1}, \ldots, b_{N_2+M_1} = r_{N_2}$, $b_{N_2+M_1+1} = -s_{M_1+1}, \ldots$, and all you have to do is show that the corresponding series $\sum b_n$ converges to r.

What does this mean? It means that for every $\varepsilon > 0$ there is an index N such that $|r - \sum_{n=1}^{N} b_n| < \varepsilon$. [DO IT]. $\square$

Summary

In this chapter, we have constructed the real numbers and proved a lot of important theorems: $\mathbb{R}$ is a complete ordered field, $\mathbb{R}$ is not countable, the existence of square roots in $\mathbb{R}$ and $\mathbb{C}$, the theorem of Bolzano-Weierstrass, the existence of supremum and infimum of bounded sets of real numbers, and the fact that we may rearrange series if and only if they converge absolutely.

Chapter 4

Continuous Functions

In this chapter, we shall study functions $f : I \longrightarrow \mathbb{R}$ defined on intervals I. In general, f will denote the function and $f(x)$ its value at x.

4.1 Continuity

Let $a < b$ be real numbers and $I = (a, b)$ be an open interval; a function $I \longrightarrow \mathbb{R}$ is said to be continuous at $x_0 \in I$ if for every $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ whenever $|x - x_0| < \delta$. This means that $f(x)$ should be close to $f(x_0)$ if x is close to x_0.

We say that f is continuous on I if it is continuous at every $x \in I$.

The function

$$f : (-1, 1) \longrightarrow \mathbb{R} : x \longmapsto \begin{cases} -1 & \text{if } -1 < x < 0 \\ +1 & \text{if } 0 \leq x < 1 \end{cases}$$

is not continuous at $x_0 = 0$: we have $|f(x) - f(x_0)| = |f(x)| = 2$ for any $x \in (-1, 0)$, so e.g. for $\varepsilon = 1$ we can make δ as small as we want without ever achieving $|f(x) - f(x_0)| < \varepsilon$.

Here's the only example of a continuous function for which we shall apply the definition of continuity directly:

Lemma 4.1. *The functions $f(x) = 1$ and $g(x) = x$ defined on any interval I are continuous on I.*

Proof. Let $x_0 \in I = (a, b)$ and $\varepsilon > 0$ be given. We claim that $\delta = \min\{\varepsilon, \frac{1}{2}|x_0 - a|, \frac{1}{2}|x_0 - b|\}$ does it (the last two elements make sure that we are working entirely within I). In fact, if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| = 0 < \varepsilon$ and $|g(x) - g(x_0)| = |x - x_0| < \delta \leq \varepsilon$. $\square$

In general, we use the ε-δ-criterion for verifying continuity as often as we use the Cauchy criterion for checking convergence: almost never. What we'll

do is prove a couple of results on continuous functions that allows us to reduce everything to Lemma 4.1.

To this end, we have to introduce some structure on the set $\mathcal{F}_I$ of functions $f : I \longrightarrow \mathbb{R}$ defined on an interval I. First, we can introduce an addition on $\mathcal{F}_I$ by letting $f + g$ denote the function defined by $(f + g)(x) = f(x) + g(x)$. Similarly, $(fg)(x) := f(x)g(x)$. With these operations, $\mathcal{F}_I$ is a ring whose zero element is the function that vanishes identically on I, and whose unit element is the constant function $f : x \longmapsto 1$.

Proposition 4.2. *Let $f : I \longrightarrow \mathbb{R}$ be a realvalued function on the open interval I. Then the following conditions are equivalent:*

1. *f is continuous at $x \in I$;*
2. *For any converging sequence (x_n) with $x_n \in I$ and $\lim x_n = x$ we have $\lim f(x_n) = f(x)$.*

Proof. Assume that f is continuous at x_0, and let (x_n) be a sequence of elements $x_n \in I$ converging to x. We have to show that $f(x_n)$ converges to $f(x)$. This means that, given some $\varepsilon > 0$, we must find an N such that $|f(x_n) - f(x)| < \varepsilon$ for all $n > N$.

Here's what we know: given any $\varepsilon > 0$, there is a $\delta > 0$ such that $|f(\xi) - f(x)| < \varepsilon$ whenever $|\xi - x| < \delta$ and $\xi \in I$. This is almost what we want: all we have to do is make sure that the x_n above differ at most by δ from x. But that is easy: since x_n converges to x, there is an N such that $|x_n - x| < \delta$ for all $n > N$. For these n, we have $|f(x_n) - f(x)| < \varepsilon$ from continuity.

Now assume that the sequence $(f(x_n))$ converges to $f(x)$ for any sequence x_n converging to x. We have to show that f is continuous at x, that is, given $\varepsilon > 0$ we have to find $\delta > 0$ such that $|f(\xi) - f(x)| < \varepsilon$ whenever $|\xi - x| < \delta$. We prove this by contradiction.

We know that f is continuous at x if

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall \xi \in (x - \delta, x + \delta) :\ |f(\xi) - f(x)| < \varepsilon.$$

Thus f is not continuous at x if[1]

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists \xi \in (x - \delta, x + \delta) :\ |f(\xi) - f(x)| \geq \varepsilon.$$

Fix an $\varepsilon > 0$ for which this is true. For each positive $\delta = \frac{1}{n}$ we can find a real number $\xi_n \in [x - \frac{1}{n}, x + \frac{1}{n}]$ such that $|f(\xi_n) - f(x)| \geq \varepsilon$. We clearly have $\lim \xi_n = x$, hence $\lim f(\xi_n) = f(x)$ by assumption; but this contradicts the inequality $|f(\xi_n) - f(x)| \geq \varepsilon$. $\square$

[1] Let $P(x)$ be a property of x (say 'x is even') and let $Q(x)$ be its negation (that is, 'x is not even); then the negation of $\forall x \in \mathbb{N} : P(x)$ (x is even for all x) is $\exists x \in \mathbb{N} : Q(x)$ (there is some x such that x is not even'); similarly, the negation of $\exists x \in \mathbb{N} : P(x)$ is $\forall x \in \mathbb{N} : Q(x)$. Applying this observation repeatedly shows that the negation of $\forall x\ \exists y : P(x)$ is $\exists x\ \forall y : Q(x)$ etc.

If $f(x_n)$ converges to $f(\xi)$ for every sequence (x_n) in I converging to ξ, we write $f(\xi) = \lim_{x\to\xi} f(x)$. In particular, continuity at $\xi = \lim x_n$ means that $f(\lim x_n) = \lim f(x_n)$.

Theorem 4.3. *The set $\mathcal{C}_I^0$ of continuous function on an open interval I forms a ring with respect to addition and multiplication; in fact, $\mathcal{C}_I$ is a subring of $\mathcal{F}_I$. For any $x_0 \in I$, the evaluation map $f \longmapsto f(x_0)$ is a ring homomorphism (in fact an $\mathbb{R}$-homomorphism) $\mathcal{C}_I^0 \longrightarrow \mathbb{R}$.*

Proof. We have to show that if $f, g \in \mathcal{C}_I^0$, then so is $f + g$, in other words: if f and g are continuous functions on I, then so is $f + g$. Assume therefore that f and g are continuous, and let an $x_0 \in I$ and a $\varepsilon > 0$ be given. We have to make $|(f + g)(x) - (f + g)(x_0)|$ small; it should be clear how to proceed: there exist $\delta_f > 0$ and $\delta_g > 0$ such that $|f(x) - f(x_0)| < \varepsilon/2$ for all $x \in I$ with $|x - x_0| < \delta_f$, and $|g(x) - g(x_0)| < \varepsilon/2$ for all $x \in I$ with $|x - x_0| < \delta_g$. Now put $\delta = \min\{\delta_1, \delta_2\}$; then

$$\begin{aligned} |(f+g)(x) - (f+g)(x_0)| &= |f(x) - f(x_0) + g(x) - g(x_0)| \\ &\leq |f(x) - f(x_0)| + |g(x) - g(x_0)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

whenever $x \in I$ and $|x - x_0| < \delta$. Thus $f + g$ is continuous at any $x_0 \in I$.

This proof was an exercise in using the ε-δ criterion. Using limits, everything becomes much easier: Assume that f and g are continuous ar x_0. Let (x_n) be a sequence in I converging to x_0. Then

$$\begin{aligned} \lim(f+g)(x_n) &= \lim[f(x_n) + g(x_n)] = \lim f(x_n) + \lim g(x_n) \\ &= f(x_0) + g(x_0) = (f+g)(x_0). \end{aligned}$$

Nice, huh? The proof that fg is continuous is just as easy. Finally, commutativity, associativity and distributivity are inherited from the ring of functions on I. □

Thus e.g. $f(x) = 2x^2 + 7$ is continuous on any interval I because $f(x) = 2g(x) \cdot g(x) + 7 \cdot 1$ is formed by adding and multiplying functions that are continuous on I.

Proposition 4.4. *Assume that f and g are functions $I \longrightarrow \mathbb{R}$ continuous at $x_0 \in I$. If $g(x_0) \neq 0$, then there is an interval $J \subseteq I$ containing x_0 such that f/g is defined on J and continuous at x_0.*

Proof. We want to show that if $g(x_0) \neq 0$, then actually $g(x) \neq 0$ for all $x \in J$, where J is some interval containing x_0.

Assume that this is wrong. Then each interval $[x_0 - \frac{1}{n}, x_0 + \frac{1}{n}]$ contains a ξ_n with $g(\xi_n) = 0$. Clearly $\lim \xi_n = x_0$, and since g is continuous at x_0, we have $g(x_0) = \lim g(x_n) = \lim 0 = 0$: contradiction.

Now f/g is a function $J \longrightarrow \mathbb{R}$. We claim that f/g is continuous at x_0. Let (x_n) be a sequence in J with limit x_0. Then $\lim f(x_n)/g(x_n) =$

$\lim f(x_n)/\lim g(x_n) = f(x_0)/g(x_0)$ by the limit theorems and the fact that f and g are continuous at x_0. $\square$

Theorem 4.5. *If $f : I \longrightarrow R$ and $g : J \longrightarrow R$ are functions such that $f(I) \subseteq J$, then $g \circ f$ is continuous if f and g are.*

Proof. Let $x_0 \in I$; then $y_0 = f(x_0) \in J$. We have to prove that if f is continuous at x_0 and g is continuous at y_0, then $g \circ f$ is continuous at x_0. Let (x_n) be a sequence in I converging to x_0. Then $(f(x_n))$ converges to $y_0 = f(x_0)$ because f is continuous at x_0. But now g is continuous at y_0, and $(f(x_n))$ is a sequence converging to y_0, hence $(g(f(x_n)))$ converges to $g(y_0) = g(f(x_0))$. $\square$

4.2 Properties of Continuous Functions

The results above form the formal part of the investigation of continuous functions. Now we get to more interesting theorems.

Continuous functions on open intervals need not be bounded: $f(x) = \frac{1}{x}$ is continuous on $(0, 1)$, but clearly not bounded. Moreover, f does not assume a maximum or minimum on $(0, 1)$. Continuous functions on closed intervals, on the other hand, have these properties:

Theorem 4.6. [2] *Let $f : [a, b] \longrightarrow \mathbb{R}$ be a real-valued continuous function. Then f is bounded. Moreover, f assumes a maximum and a minimum on $[a, b]$.*

Proof. Assume that f is not bounded on $I = [a, b]$. Then for each $n \in \mathbb{N}$ there must be an $x_n \in I$ with $f(x_n) > n$. Bolzano-Weierstrass gives us a convergent subsequence (ξ_n) of (x_n); put $x = \lim \xi_n$. Since $a \leq \xi_n \leq b$, we must have $a \leq x \leq b$ (here the proof would go wrong if we started with an open interval). Now f is continuous at x, hence $f(x) = \lim f(\xi_n)$. But $f(\xi_n) > n$, so the sequence $f(\xi_n)$ does not converge: contradiction.

Now f is bounded. This implies that the set $S = \{f(x) : x \in I\}$ is a bounded set of real numbers. Such sets have a supremum; let $M = \sup S$. For each $n \in \mathbb{N}$, the number $M - \frac{1}{n}$ can't be an upper bound of S because M is the least upper bound. Thus there must exist some $y_n \in I$ with $M - \frac{1}{n} < f(y_n) \leq M$. By definition of convergence, this implies $\lim f(y_n) = M$. By Bolzano-Weierstrass (you're beginning to see how important this result is), there is a convergent subsequence (η_n) of (y_n); put $y = \lim \eta_n$. Again, $y \in I$, and since f is continuous at y, we find $f(y) = \lim f(\eta_n)$. But $f(\eta_n)$ is a subsequence of the converging sequence $f(y_n)$, hence $\lim f(\eta_n) = \lim f(y_n) = M$. But now we just have proved that $M = f(y)$, that is: the supremum is actually a maximum!

Replacing f by $-f$ and applying the result we just proved, it follows that f also has a minimum because the maximum of $-f$ is the minimum of f. $\square$

Theorem 4.7 (Intermediate Value Theorem). [3] *Let $f : I = [a, b] \longrightarrow \mathbb{R}$ be a continuous functions. If $f(a) < y < f(b)$, then there exists an $x \in (a, b)$ such that $f(x) = y$.*

[2] Ross, Thm. 18.1.
[3] Ross, Thm. 18.2.

Proof. Put $S = \{x \in I : f(x) < y\}$; then $S \neq \varnothing$ since $a \in S$. Since S is bounded (by a and b), it has a supremum: put $\xi = \sup S$. Since b is an upper bound for S and ξ is the least upper bound, we have $\xi \leq b$. Moreover, $a \in S$ implies $\xi \geq a$. Thus $\xi \in I$. Finally, we can't have $\xi = a$ or $\xi = b$ since $f(a) < y$ and $f(b) > y$.

Now for each $n \in \mathbb{N}$. $\xi - \frac{1}{n}$ is not an upper bound for S. Therefore there exists an $s_n \in S$ with $\xi - \frac{1}{n} < s_n \leq \xi$. This implies $\lim s_n = \xi$. Since $f(s_n) < y$, we have $\lim f(s_n) \leq y$, and continuity gives $f(\xi) \leq y$.

It remains to show that $f(\xi) \geq y$. To this end, put $t_n = \min(\xi + \frac{1}{n}, b)$; then$t_n \in I$, and (t_n) converges to ξ. Since $t_n \notin S$, we have $f(t_n) \geq y$; since f is continuous at ξ, we find $f(\xi) = \lim f(t_n) \geq y$. □

A function $f : I \longrightarrow \mathbb{R}$ is called strictly increasing on I if for every pair $x, x' \in I$ with $x < x'$ we have $f(x) < f(x')$.

Lemma 4.8. *Strictly increasing functions are injective.*

Proof. Take $x, x' \in I$ and assume that $x \neq x'$; we want to show that $f(x) \neq f(x')$. But $x \neq x'$ implies $x < x'$ or $x' < x$; since f is strictly increasing, this implies $f(x) < f(x')$ or $f(x') < f(x)$. In either case, $f(x) \neq f(x')$. □

Now we can show that strictly increasing continuous functions have an inverse with respect to composition:

Theorem 4.9 (Existence of inverse function). [4] *Let $f : I \longrightarrow \mathbb{R}$ be a strictly increasing continuous function. Then $f(I)$ is an interval, and there exists a function $f^{-1} : f(I) \longrightarrow I$ such that $f^{-1} \circ f$ is the identity map on I. The function f^{-1} is strictly increasing and continuous.*

Proof. Let $I = [a, b]$; we claim that $f(I) = [f(a), f(b)]$. Note that $a < b$ implies $f(a) < f(b)$, so $[f(a), f(b)]$ really is an interval. For proving the claim we have to show that if $f(a) < y < f(b)$, then $y = f(x)$ for some $x \in I$. But this is the intermediate value theorem.

By the lemma, f is injective, hence $f : I \longrightarrow f(I)$ is a bijection, and we can define a function $f^{-1} : f(I) \longrightarrow I$ by writing $y \in f(I)$ as $y = f(x)$ and mapping $y \longmapsto x$.

We have to show that f^{-1} is strictly increasing and continuous. Assume that $y < y'$ for $y, y' \in f(I)$. Write $y = f(x)$ and $y' = f(x')$; If we had $x = f^{-1}(y) \geq f^{-1}(y') = x'$, then we would also have $y = f(x) \geq f(x') = y'$: contradiction.

Finally, let us prove that f^{-1} is continuous at $\eta \in f(I)$. There are two ways to do this: we can use the limit definition or the $\varepsilon - \delta$-definition of continuity.

First Proof. Let's use the limit definition first. Let (y_n) be a sequence in J with $y_n \longrightarrow \eta$. We have to show that $x_n := f^{-1}(y_n)$ converges to $\xi := f^{-1}(\eta)$. Assume ot doesn't; then

$$\exists\, \varepsilon > 0 \forall\, N \in \mathbb{N}\ \exists n > N :\ |f^{-1}(y_n) - f^{-1}(\eta)| \geq \varepsilon.$$

[4] Ross, Thm. 18.4.

Pick an $\varepsilon > 0$ with this property; then there is an $n_1 > 1$ such that $|f^{-1}(y_{n_1}) - f^{-1}(\eta)| \geq \varepsilon$; next there is an $n_2 > n_1$ such that $|f^{-1}(y_{n_2}) - f^{-1}(\eta)| \geq \varepsilon$, and by repeating this construction we get a subsequence (y_{n_k}) of (y_n) with the property that

$$|f^{-1}(y_{n_k}) - f^{-1}(\eta)| \geq \varepsilon \quad \text{for all } k \in \mathbb{N}. \tag{4.1}$$

Then $f^{-1}(y_{n_k})$ is a sequence in I, hence bounded, so Bolzano-Weierstraß gives us a converging subsequence $(y_{n_{k_l}})$ such that $f^{-1}(y_{n_{k_l}}))$ converges. Let $x \in I$ denote the limit. Then evaluating f at $x = \lim f^{-1}(y_{n_{k_l}})$ gives $f(x) = f(\lim f^{-1}(y_{n_{k_l}}))$. But f is continuous, hence

$$f(\lim f^{-1}(y_{n_{k_l}})) = \lim f(f^{-1}(y_{n_{k_l}})) = \lim y_{n_{k_l}} = y.$$

Thus $f^{-1}(y_{n_{k_l}})$ converges to $x = f^{-1}(y)$, contradicting (4.1): in fact, if all $f^{-1}(y_{n_k})$ differ from $f^{-1}(y)$ by at least some fixed $\varepsilon > 0$, then there can't be a subsequence converging to $f^{-1}(y)$. $\square$

Second Proof.

The second proof uses the $\varepsilon - \delta$-criterion. We use some facts we already have proved: that f^{-1} is a strictly increasing function mapping an interval J onto an interval I (that is, $f^{-1}(J) = I$). The claim thus will follow once we have proved

Proposition 4.10. *If g is a strictly increasing function mapping an interval J onto another interval I, then g is continuous on g.*

Proof. Write $I = [a, b]$ and $J = [c, d]$; since $g : J \longrightarrow I$ is strictly increasing, we actually have $a = g(c)$ and $b = g(d)$. Pick $\eta \in J$; we want to show that g is continuous at η. To this end, we first assume that $\eta \neq c, d$. Since g is strictly increasing, this implies $g(\eta) \neq a, b$. Thus there exists some $\varepsilon_0 > 0$ such that $[g(\eta) - \varepsilon_0, g(\eta) + \varepsilon_0] \subseteq I$.

Now let $\varepsilon > 0$; we have to show that $|g(y) - g(\eta)| < \varepsilon$ for $|y - \eta| < \delta$; it is clearly sufficient to prove this for small ε, so we are free to assume $\varepsilon < \varepsilon_0$.

Since g maps J onto I, there exist $y_1, y_2 \in J$ such that $g(y_1) = g(\eta) - \varepsilon$ and $g(y_2) = g(\eta) + \varepsilon$. Since g is strictly increasing, we have $y_1 < \eta < y_2$. Now whenever $y_1 < y < y_2$, we have $g(y_1) < g(y) < g(y_1)$, hence

$$g(\eta) - \varepsilon = g(y_1) < g(y) < g(y_2) = g(\eta) + \varepsilon.$$

But this implies $|g(y) - g(\eta)| < \varepsilon$. Thus we have proved the inequality we wanted for all y such that $y_1 < y < y_2$. Thus if we put $\delta = \min\{y - y_1, y_2 - y\}$, we see that $|y - \eta| < \delta$ implies $y - \eta < y_2 - y$, hence $2y < \eta + y_2 < 2y_2$ and so $y < y_2$; similarly, $y - \eta > y_1 - y$ implies $2y > y_1 + \eta > 2y_1$, so $y > y_1$. Thus $|y - \eta| < \delta$ implies $y_1 < y < y_2$, which in turn implies $|g(y) - g(\eta)| < \varepsilon$. $\square$

4.3 Uniform Continuity

Consider the function $f : \mathbb{R} \longrightarrow \mathbb{R} : x \longmapsto x^2$. Let us look at the proof that f is continuous. Given $\xi \in \mathbb{R}$, we have to find for every $\varepsilon > 0$ a $\delta > 0$ such that $|x-\xi| < \delta$ implies $|f(x)-f(\xi)| < \varepsilon$. Now $f(x)-f(\xi) = x^2-\xi^2 = (x-\xi)(x+\xi)$; if we assume that $|x-\xi| < \delta$, then we can bound $|x+\xi|$ by $|x+\xi| = |x-\xi+2\xi| \leq |x-\xi| + 2|\xi| < \delta + 2|\xi|$, and we get $|f(x) - f(\xi)| < \delta(\delta + 2|\xi|)$.

Can we make the right hand side small by choosing δ small enough? Certainly, as long as ξ is fixed, because taking

$$\delta = \min\left\{1, \frac{\varepsilon}{1+2|\xi|}\right\}$$

gives $\delta + 2|\xi| < 1 + 2|\xi|$ and

$$|f(x) - f(\xi)| < \delta(\delta + 2|\xi|) < \varepsilon$$

as desired.

Note, however, that our choice of δ depends on ξ: we don't have a uniform choice valid for all ξ, and the larger $|\xi|$, the smaller we have to make our δ.

If $f : I \longrightarrow \mathbb{R}$ is a function for which such a uniform choice of δ is possible, then we say that f is uniformly continuous on I. In other words: f is uniformly continuous on I if

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, x, \xi \in I, |x-\xi| < \delta :\ |f(x) - f(\xi)| < \varepsilon.$$

Compare this with the definition of 'regular' continuity on I: f is continuous on I if

$$\forall\, \xi \in I\ \forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, x \in I, |x-\xi| < \delta :\ |f(x) - f(\xi)| < \varepsilon.$$

The following result is important for defining the Riemann integral (it is also important for showing that functions like exp, sin, cos, log defined by power series are continuous):

Theorem 4.11. *If $f : I \longrightarrow \mathbb{R}$ is continuous on the closed interval $I = [a, b]$, then f is uniformly continuous on I.*

Proof. Assume not. Then

$$\exists\, \varepsilon > 0\ \forall\, \delta > 0\ \exists\, x, \xi \in I, |x-\xi| < \delta :\ |f(x) - f(\xi)| \geq \varepsilon.$$

Pick such an $\varepsilon > 0$; then for each $\delta = \frac{1}{n}$ (you see we are constructing sequences; expect to see Bolzano-Weierstrass soon) there exist $x_n, \xi_n \in I$ such that $|x_n - \xi_n| < \frac{1}{n}$ and

$$|f(x_n) - f(\xi_n)| \geq \varepsilon. \tag{4.2}$$

By Bolzano-Weierstrass there is a convergent sbsequence (x_{n_k}) of (x_n). Since $a \leq x_{n_k} \leq b$, we have $a \leq \lim x_{n_k} \leq b$ (here we use the fact that I is closed).

Since $(x_n - \xi_n)$ is a null sequence, we have $\lim x_{n_k} = \lim \xi_{n_k} =: z$. Using the fact that f is continuous we get

$$f(z) = \lim f(x_{n_k}), \quad f(z) = \lim f(\xi_{n_k}).$$

Thus $f(x_{n_k}) - f(\xi_{n_k})$ converge to 0, contradicting (4.2). $\square$

Chapter 5

The Riemann Integral

Consider a function $f : I \longrightarrow \mathbb{R}$ on a closed interval $I = [a, b]$. The integral of f defines the 'area' between the graph of the function and the x-axis. There are several ways of defining integrals: the most powerful among the simple definition is that of the Riemann integral, and even for them there is a variety of almost equivalent definitions. There are integrals that are more powerful than Riemann's, in particular the Lebesgue integral (which, in some sense, can be shown to be the most powerful); the Lebesgue integral is more powerful than Riemann's because every function that can be integrated by a Riemann integral can also be integrated by Lebesgue's, but not conversely: Dirichlet's function

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

has Lebesgue integral 0 on the interval $I = [0, 1]$, but is not Riemann integrable.

In any case, all definitions of integrals yield the same results for the class of continuous functions, so unless you are interested in functions with lots of discontinuities, the Riemann integral is fine.

5.1 Riemann Sums

Here we want to motivate the definition of the integral of a function $f : [a, b] \longrightarrow \mathbb{R}$. We do this by writing down a couple of properties such an integral should have if it is to give us the area beneath the graph of a function.

No matter how the area is defined, if we add the area below f on $[a, b]$ and the area on $[b, c]$ we should get the area on $[a, c]$: so what we want is

[INT1] $\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx$.

We also want the integral of the function $k \cdot f(x)$ to be k times the integral of $f(x)$:

[INT2] $\int_a^b k \cdot f(x)dx = k \int_a^b f(x)dx$.

Finally, put $m_I(f) = \inf\{f(x) : x \in I\}$ and $M_I(f) = \sup\{f(x) : x \in I\}$. Then we want the area of f on $I = [a, b]$ to be bounded by $(b-a)m_I(f)$ from below and by $(b-a)M_I(f)$ from above:
[INT3] $(b-a)m_I(f) \leq \int_a^b f(x)dx \leq (b-a)M_I(f)$.

The last condition determines the integral of constant functions: if $f(x) = c$ on $I = [a, b]$, then $m_I(f) = M_I(f) = c$, hence [INT3] implies $\int_a^b f(x)dx = (b-a)c$.

There is of course no guarantee that an integral with these three properties exists. As a matter of fact, however, both the Riemann and the Lebesgue integral have these properties, so we can't expect that these properties determine the integral.

The key to the definition of the Riemann integral is the observation that [INT3] gives good approximations if the interval I is small. How can you improve the approximation if I is large? Well, you might try to divide the interval into smaller parts: if $I = [a, b]$, look at 'partitions' $P = \{a = t_0 < t_1 < t_2 < \ldots < t_n = b\}$; by [INT1], our integral should satisfy

$$\int_a^b f(x)dx = \int_{t_0}^{t_1} f(x)dx + \int_{t_1}^{t_2} f(x)dx + \ldots + \int_{t_{n-1}}^{t_n} f(x)dx.$$

Now if we put $m_i(f) = \sup\{f(x) : t_i \leq x \leq t_{i+1}\}$ (sometimes we shall also write $m(f, [t_i, t_{i+1}])$ for $m_i(f)$) and define $M_i(f)$ accordingly, then [INT3] gives us the bounds

$$\begin{aligned} L(f, P) &= (t_1 - t_0)m_0(f) + (t_2 - t_1)m_1(f) + \ldots + (t_n - t_{n-1})m_n(f) \\ &\leq \int_a^b f(x)dx \\ &\leq (t_1 - t_0)M_0(f) + \ldots + (t_n - t_{n-1})M_n(f) \;=\; U(f, P), \end{aligned}$$

and if the intervals $[t_i, t_{i+1}]$ are small, we expect that these bounds are much better than those you get by applying [INT3] directly to f and $[a, b]$.

But if we get better and better approximations by making our partitions finer and finer, then we would expect that the 'true' value of the integral is something like the 'limit' of these approximations $L(f, P)$ and $U(f, P)$, at least in the case when both lower and upper bounds tend to the same limit. Actually, we should not talk about limits here because partitions do not form a sequence: in fact, even the partitions of $[a, b]$ consisting of three points $\{a, t_1, b\}$ are not countable because t_1 can be any real number in (a, b), and the real numbers in an nonempty interval can't be counted!

We can, however, construct sequences of partitions by looking only at partitions that are refinements of one another. Let us call a partition $Q = \{s_0, s_1, \ldots, s_n\}$ of an interval $[a, b]$ finer than a partition $P = \{t_0, t_1, \ldots, t_m\}$ of $[a, b]$ if $P \subseteq Q$ (as sets).

Lemma 5.1. *If $f : I \longrightarrow \mathbb{R}$ is a bounded function, and if P and Q are partitions of I with $P \subseteq Q$, then $L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P)$.*

Proof. The partition Q is finer than P; this means that we can get Q from P by 'adding' some points to P. Using induction it is therefore sufficient to prove the claim for

$$P = \{a = t_0 < t_1 < \ldots < t_k < t_{k+1} < \ldots < t_n = b\}$$

and

$$Q = \{a = t_0 < t_1 < \ldots < t_k < u < t_{k+1} < \ldots < t_n = b\}.$$

The sums $L(f, P)$ and $L(f, Q)$ don't differ much; in fact, if we form the difference, everything cancels except

$$\begin{aligned} L(f,Q) - L(f,P) &= m(f,[t_{k-1},u])(u - t_{k-1}) + m(f,[u,t_k])(t_k - u) \\ &\quad -m(f,[t_{k-1},t_k])(t_k - t_{k-1}). \end{aligned}$$

We have to show that this is nonnegative. To this end, observe that for sets $\varnothing \neq S \subseteq T$ of real numbers, we have $\inf T \leq \inf S \leq \sup S \leq \sup T$ (assuming these exist, e.g. in the case where T (and hence S) is bounded): in fact, since $S \subseteq T$, every lower bound for T is a lower bound for S, in particular the greatest lower bound for T is a lower bound for S, and so $\inf T \leq \inf S$.

This shows that

$$\begin{aligned} m(f,[t_{k-1},t_k])(t_k - t_{k-1}) &= m(f,[t_{k-1},t_k])\{(t_k - u) + (u - t_{k-1})\} \\ &\leq m(f,[u,t_k])(t_k - u) + m(f,[t_{k-1},u])(u - t_{k-1}) \end{aligned}$$

This takes care of $L(f, P) \leq L(f, Q)$. The inequality $U(f, Q) \leq U(f, P)$ can be proved similarly, or, if you want to, by reducing it to the above: in fact, $\sup\{-f(x) : x \in I\} = -\inf\{f(x) : x \in I\}$ implies that $L(-f, P) = -U(f, P)$, hence $L(f, P) \leq L(f, Q)$ if and only if $U(f, Q) \leq U(f, P)$.

The middle inequality finally is obvious in view of $\inf S \leq \sup S$ for bounded nonempty sets S. □

Definition. A bounded function $f : I = [a, b] \longrightarrow \mathbb{R}$ (we need boundedness to ensure that the suprema and infima involved in the definition of $L(f, P)$ and $U(f, P)$ exist; of course continuous functions on closed intervals are bounded, but we want to define our integral for functions that are not necessarily continuous) is called (Riemann) integrable if $L(f) = U(f)$.

Example 1. Constant functions are integrable. In fact, if $f(x) = c$, then we expect the integral of f over $I = [a, b]$ to be the area of the rectangle bounded by $(a, 0)$, $(b, 0)$, (b, c) and (a, c), namely $A = c(a - b)$. And this is what we get: take any partition $P = \{a = t_0 < t_1 < \ldots < t_k = b\}$; then $m_i(f) = \sup\{f(x) : t_i \leq x \leq t_{i+1}\} = c$ since f is constant, and similarly $M_i(f) = c$, so

$$\begin{aligned} L(f,P) &= \sum_{i=1}^{k} c(t_i - t_{i-1}) \;=\; c(t_1 - t_0) + c(t_2 - t_1) + \ldots + c(t_k - t_{k-1}) \\ &= c(t_k - t_0) \;=\; c(b - a). \end{aligned}$$

Similarly, $U(f,P) = c(b-a)$ for any partition, hence $L(f)$ and $U(f)$ exist and both equal $c(b-a)$. This proves that $\int_a^b f(x)dx = c(b-a)$ for $f(x) = c$.

Example 2. Consider $f(x) = x$. Here $m(f,[t_i,t_{i+1}]) = t_i$ and $M(f,[t_i,t_{i+1}]) = t_{i+1}$. Let us use the partitions P_k defined by $t_i = a + \frac{i}{k}(b-a)$; then $t_0 = a$, $t_1 = a + \frac{b-a}{k}$, $\ldots$, $t_k = a + (b-a) = b$, hence $t_i - t_{i-1} = \frac{b-a}{k}$, and

$$\begin{aligned}
U(f,P_k) &= \sum_{i=1}^{k} t_i(t_i - t_{i-1}) \;=\; \sum_{i=1}^{k} \big(a + \frac{i}{k}(b-a)\big)\frac{b-a}{k} \\
&= \frac{b-a}{k}\Big(\sum_{i=1}^{k} a + \frac{b-a}{k}\sum_{i=1}^{k} i\Big) \\
&= \frac{b-a}{k}\cdot ka + \frac{(b-a)^2}{k^2}\cdot\frac{k(k+1)}{2} \\
&= ba - a^2 + \frac{1}{2}(b-a)^2 + \frac{1}{2k}(b-a)^2 \;=\; \frac{1}{2}(b^2-a^2) + \frac{1}{2k}(b-a)^2.
\end{aligned}$$

Clearly $\lim U(f,P_k) = \frac{1}{2}(b^2-a^2)$. In particular, $U(f) \le \frac{1}{2}(b^2-a^2)$. The proof that $\lim L(f,P_k)$ exists and has the same value is left as an exercise. What we may conclude from the above is that $L(f) \ge \frac{1}{2}(b^2-a^2)$ and $U(f) \le \frac{1}{2}(b^2-a^2)$, inparticular $L(f) \ge U(f)$. This does, however, not yet prove that $f(x) = x$ is integrable because we have only considered a very special type of partitions; for proving that the integral of f exists we have to look at all partitions.

Since the number of partitions is not countable, this task is hopeless without the help of a theorem that does this for us. What we would like to have is a theorem telling us that $L(f) \le U(f)$. But isn't this obvious? After all, we have $L(f,P) \le U(f,P)$ for every partition P.

Actually, what might happen is that there exist partitions P and Q such that $L(f,P) > U(f,Q)$: this is not excluded by the inequality above, and if that happened, $L(f) = \sup L(f,P)$ would be larger than $U(f) = \inf U(f,P)$. Fortunately, this doesn't happen. For a proof, we will need to compare $L(f,P)$ and $U(f,Q)$ for different partitions, and this is done by constructing a partition containing both. But this is easy: if P and Q are partitions, then so is $P \cup Q$, and since $P \subseteq P \cup Q$ and $Q \subseteq P \cup Q$, this partition is a common refinement of P and Q.

The following result is exactly what we need:

Proposition 5.2. *If $f : I \longrightarrow \mathbb{R}$ is a bounded function on $I = [a,b]$, then $L(f) \le U(f)$.*

In fact, if we know that $L(f) \le U(f)$ from the lemma and $L(f) \ge U(f)$ by our calculation, then we can conclude that $L(f) = U(f)$, hence $f(x) = x$ is integrable on $[a,b]$, and $\int_a^b f(x)dx = \frac{1}{2}(b^2-a^2)$.

Lemma 5.3. *If $f : I \longrightarrow \mathbb{R}$ is a bounded function on $I = [a,b]$, and if P and Q are partitions of I, then $L(f,P) \le U(f,Q)$.*

Proof. Applying Lemma 5.1 to P, $P \cup Q$ and Q we get

$$L(f,P) \leq L(f,P \cup Q) \leq U(f,P \cup Q) \leq U(f,Q).$$

□

Proof of Prop. 5.2. Let P be a partition of I. For any partition Q of I, we have $L(f,P) \leq U(f,Q)$ by Lemma 5.3. Thus $L(f,P)$ is a lower bound for the set of $U(f,Q)$, hence the greatest lower bound $U(f)$ for this set can't be smaller: $L(f,P) \leq U(f)$.

This inequality now is true for any partition P; thus $U(f)$ is an upper bound for the set of $L(f,P)$, which in turn means that the least upper bound $L(f)$ can't be larger than $U(f)$, i.e., $L(f) \leq U(f)$. □

This concludes the proof of $\int_a^b x\,dx = \frac{1}{2}(b^2 - a^2)$.

Next we prove a criterion for integrability that will come in handy in proofs:

Proposition 5.4. *A bounded function $f : I \longrightarrow \mathbb{R}$ is integrable if and only if for every $\varepsilon > 0$ there exists a partition $P = \{a = t_0 < t_1 < t_2 < \ldots < t_n = b\}$ of $I = [a,b]$ such that $U(f,P) - L(f,P) < \varepsilon$.*

Proof. Note that we don't need absolute values around $U(f,P) - L(f,P)$ since we know that this term is nonnegative.

Assume first that f has a Riemann integral on I, and let $\varepsilon > 0$ be given. Since $L(f) - \frac{\varepsilon}{2}$ is not an upper bound for the set of $L(f,P)$, there must be a partition P_1 of I such that $L(f,P_1) > L(f) - \frac{\varepsilon}{2}$. Similarly, there is a partition P_2 such that $U(f,P_2) < U(f) + \frac{\varepsilon}{2}$.

Now put $P = P_1 \cup P_2$. Then applying Lemma 5.3 gives

$$\begin{aligned} U(f,P) - L(f,P) &\leq U(f,P_2) - L(f,P_1) \\ &< U(f) + \frac{\varepsilon}{2} - L(f) + \frac{\varepsilon}{2} \;=\; U(f) - L(f) + \varepsilon, \end{aligned}$$

but since f is integrable, we have $L(f) = U(f)$. This implies the claim $U(f,P) - L(f,P) < \varepsilon$.

Now suppose that for each $\varepsilon > 0$ there is a partition P of I with $U(f,P) - L(f,P) < \varepsilon$. Then

$$\begin{aligned} U(f) &\leq U(f,P) \;= U(f,P) - L(f,P) + L(f,P) \\ &< \varepsilon + L(f,P) \leq \varepsilon + L(f). \end{aligned}$$

Since this is true for any $\varepsilon > 0$, we must have $U(f) \leq L(f)$. But $L(f) \leq U(f)$ by Prop. 5.2, and our claim follows. □

5.2 Main Properties of Riemann Integrals

The main theorems we shall prove about the Riemann integral is that every continuous function is integrable, and that the Riemann integral has the properties [INT1] - [INT3] (as well as a few more).

Let us start with

Proposition 5.5. *Integrable functions $I \longrightarrow \mathbb{R}$ form a ring.*

Proof. Assume that f and g are integrable. Given $\varepsilon > 0$, we have to find a partitions P of $[a, b]$ such that $U(f + g, P) - L(f + g, P) < \varepsilon$.

Now we know that f is integrable, so there is a partition P_1 of $[a, b]$ such that $U(f, P_1) - L(f, P_1) < \varepsilon/2$; similarly, there is a partition P_2 of $[a, b]$ such that $U(g, P_2) - L(f, P_2) < \varepsilon$. Now consider $P = P_1 \cup P_2$. Then

$$\begin{aligned} U(f, P) - L(f, P) &\leq U(f, P_1) - L(f, P_1) < \frac{\varepsilon}{2}, \\ U(g, P) - L(g, P) &\leq U(g, P_2) - L(g, P_2) < \frac{\varepsilon}{2}. \end{aligned}$$

Now for any subinterval J of I we have (Exercise!)

$$\inf\{f(x) + g(x) : x \in J\} \geq \inf\{f(x) : x \in J\} + \inf\{g(x) : x \in J\},$$

hence $m_J(f+g) \geq m_J(f)+m_J(g)$, which in turn implies $L(f+g, P) \geq L(f, P)+L(g, P)$. Similarly, we can prove $U(f + g, P) \leq U(f, P) + U(g, P)$, hence

$$U(f+g, P)-L(f_g, P) \leq U(f, P)+U(g, P)-L(f, P)-L(g, P) < \tfrac{\varepsilon}{2}+\tfrac{\varepsilon}{2} = \varepsilon.$$

Thus $f + g$ is Riemann integrable. Moreover,

$$\begin{aligned} \int_a^b (f + g)(x)dx &= U(f + g) \leq U(f + g, P) \leq U(f, P) + U(g, P) \\ &< L(f, P) + L(g, P) + \varepsilon \leq L(f) + L(g) + \varepsilon \\ &= \int_a^b f(x)dx + \int_a^b g(x)dx + \varepsilon \end{aligned}$$

and

$$\begin{aligned} \int_a^b (f + g)(x)dx &= L(f + g) \geq L(f + g, P) \geq L(f, P) + L(g, P) \\ &> U(f, P) + U(g, P) - \varepsilon \geq U(f) + U(g) - \varepsilon \\ &= \int_a^b f(x)dx + \int_a^b g(x)dx - \varepsilon, \end{aligned}$$

so we conclude that $\int_a^b (f + g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx$.

Since constant functions are Riemann integrable, the neutral element with respect to additon, the zero function, is integrable. Moreover, if f is Riemann integrable, then so is $-f$. In fact, we have

$$m_J(-f) = \inf\{-f(x) : x \in J\} = -\sup\{f(x) : x \in J\} = -M_J(f),$$

hence $L(-f, P) = -U(f, P)$ and $L(-f) = -U(f)$. Similarly, $U(-f) = -L(f)$, hence $L(f) = U(f)$ implies $L(-f) = U(-f)$.

Now let us show that if f and g have Riemann integrals, then so does their product fg. It is sufficient to prove this for $g = f$: in fact, if we know that f^2

is integrable as long as f is, then $4fg = (f+g)^2 - (f-g)^2$ is integrable since it is a sum (difference) of integrable functions.

Now consider $U(f^2, P)$ and $L(f^2, P)$ for partitions P of I. We know that

$$f(x)^2 - f(y)^2 = [f(x) + f(y)][f(x) - f(y)]$$

□

Our biggest source of integrable functions are the continuous functions:

Theorem 5.6. *Every continuous function $f : I = [a, b] \longrightarrow \mathbb{R}$ is Riemann integrable.*

Proof. We want to apply Prop. 5.4, so let an $\varepsilon > 0$ be given. We have to construct a partition P of I such that $U(f, P) - L(f, P) < \varepsilon$. How can we do this? We have to use the fact that f is continuous somehow. Continuous at x_0 means that we can find a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ whenever $|x - x_0| < \delta$. Thus we we have found a δ-interval around x_0, but this is not exactly what we want; what we'd like to have is a whole lot of δ-intervals that cover $I = [a, b]$ because then we would have a partition of I. But as long as we exploit the usual continuity, our δ will depend on the choice of x_0. Fortunately, continuous functions on closed intervals I are uniformly continuous, so there is a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon/(b-a)$ whenever $x, y \in I$ and $|x - y| < \delta$. Now we choose a partition $P = \{a = t_0 < t_1 < \ldots < t_n = n\}$ such that the lengths of the intervals $[t_k, t_{k+1}]$ are bounded by δ, i.e., such that $t_{k+1} - t_k < \delta$ for $k = 1, 2, \ldots, n$.

On each $I_k = [t_k, t_{k+1}]$, the function f assumes a maximum M_k (at $x_k \in I_k$, say) and a minimum m_k (at $y_k \in I_k$), and then $M_i - m_i = f(x_i) - f(y_i) < \varepsilon/(b-a)$ since $|x_i - y_i| < \delta$. But then

$$U(f, P) - L(f, P) \ < \ \sum_{k=1}^{n} \frac{\varepsilon}{b-a}(t_k - t_{k+1}) = \frac{\varepsilon}{b-a}(t_n - t_0) = \varepsilon,$$

and therefore f is Riemann integrable. □

Proposition 5.7. *If f is integrable, then so is $|f|$, and we have $\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx$.*

Proof. □

Proposition 5.8. *If $f : I = [a, b] \longrightarrow \mathbb{R}$ is integrable, then $F(x) := \int_a^x f(t)dt$ defines a continuous function $I \longrightarrow \mathbb{R}$.*

Proof. We have to show that, given $x_0 \in I$ and $\varepsilon > 0$, there is a $\delta > 0$ such that $|F(x) - F(x_0)| < \varepsilon$ whenever $|x - x_0| < \delta$. We clearly have $F(x) - F(x_0) = \int_{x_0}^x f(t)dt$; assume that $x > x_0$. Since f is bounded, we have $|f(t)| \leq M$ for $t \in I$, hence $\left| \int_{x_0}^x f(t)dt \right| \leq \int_{x_0}^x |f(t)|dt \leq M|x - x_0|$. So if we pick $\delta = \varepsilon/M$, then $|F(x) - F(x_0)| < \varepsilon$. The case $x < x_0$ is treated similarly. □

Proposition 5.9. *If $f, g : I \longrightarrow \mathbb{R}$ are integrable, and if $f(x) \leq g(x)$ for all $x \in I$, then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.*

Proof. $\square$

Theorem 5.10 (Intermediate Value Theorem for Integrals). *If $f : I \longrightarrow \mathbb{R}$ is a continuous function on $I = [a, b]$, then there is an $x \in I$ such that*

$$f(x) = \frac{1}{b-a} \int_a^b f(x)dx.$$

Proof. Since continuous functions on closed intervals assume a maximum M and a minimum m, Proposition 5.9 gives us the bounds

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

Dividing through by $b-a > 0$ shows that $\frac{1}{b-a} \int_a^b f(x)dx$ lies between m and M; by the intermediate value theorem, f assumes this value somewhere in $[a, b]$. $\square$

Chapter 6

Differentiable Functions

6.1 Derivatives

Just as we used two definitions of continuity, there are very good reasons for using two defintions of differentiable functions. The most natural one seems to be the following: a function $f : I \longrightarrow \mathbb{R}$ defined on an open interval is said to be differentiable at $a \in I$ if the limit

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

exists; in this case, we denote it by $f'(a)$ and call it the derivative of f at a.

A simple sketch shows that $f'(a)$ can be interpreted as the slope of a tangent to f at $x = a$.

The derivative of the function $f(x) = x^k$ is computed easily (assuming it exists): we consider the sequence $x_n = a + \frac{1}{n}$ with $x_n \longrightarrow a$; then

$$x_n^k - a^k = (x_n - a)(x_n^{k-1} + a x_n^{k-2} + a^2 x_n^{k-3} + \ldots + a^{k-2} x_n + a^{k-1}),$$

hence

$$\lim_{n \to \infty} \frac{f(x_n) - f(a)}{x_n - a} = \lim_{n \to \infty} (x_n^{k-1} + a x_n^{k-2} + \ldots + a^{k-1}) = k a^{k-1}.$$

Note that for proving $f'(a) = k a^{k-1}$ we would have to consider *every* sequence $x_n \longrightarrow a$, not only the one we have looked at.

As for continuity, there are two definitions of differentiability; as a rule, the following is easier to apply in proofs:

Proposition 6.1. *Let $f : I \longrightarrow \mathbb{R}$ be a function defined on an open interval I. Then the following statements are equivalent:*

1. *f is differentiable at $a \in I$;*
2. *$f(x) = f(a) + (x - a)(f'(a) + u(x)$ for some function $u : I \longrightarrow \mathbb{R}$ such that $\lim_{x \to a} u(x) = 0$.*

Proof. Assume that $f'(a)$ exists; then $u(x) = \frac{f(x)-f(a)}{x-a} - f'(a)$ is a function defined on $I \setminus \{a\}$ such that $\lim_{x\to a} u(x) = 0$.

Conversely, assume that u has the properties in 2. Then $u(x) = \frac{f(x)-f(a)}{x-a} - f'(a)$ for $x \neq a$, and since the left hand side has a limit as $x \to a$, so does the right hand side. This implies that f has a derivative at $x = a$ with value $f'(a)$. □

Corollary 6.2. *If f is differentiable at a, then it is continuous at a.*

Proof. If f is differentiable at a, then

$$f(x) - f(a) = (x - a)(f'(a) + u(x)$$

for some function $u : I \longrightarrow \mathbb{R}$ such that $\lim_{x\to a} u(x) = 0$. Thus if (x_n) is any sequence converging to a, then $f(x_n) - f(a)$ converges to 0. □

Thus the set of functions differentiable at a is a subset of the ring of functions continuous at a. As a matter of fact, it is a subring:

Proposition 6.3. *Let I be an open interval; the functions $f : I \longrightarrow \mathbb{R}$ differentiable at $a \in I$ form a subring of the functions $f : I \longrightarrow \mathbb{R}$ continuous at a. More exactly, we have $(f + g)'(a) = f'(a) + g'(a)$ and the product rule $(fg)'(a) = f(a)g'(a) + f'(a)g(a)$.*

The reason for choosing an open interval is that we don't want to define differentiability from the left or right.

Proof. Let f and g be functions $I \longrightarrow \mathbb{R}$ that are differentiable at $a \in I$. Let (x_n) be a sequence in I converging to a; then

$$\frac{(f+g)(x_n) - (f+g)(a)}{x_n - a} = \frac{f(x_n) - f(a)}{x_n - a} + \frac{g(x_n) - g(a)}{x_n - a},$$

and since the terms on the right hand side have limits as $x_n \to a$ by assumption, so does the term on the left hand side. This shows that $f + g$ has a derivative at $x = a$, and the proof shows that in fact $(f + g)'(a) = f'(a) + g'(a)$.

Now let us look at the product of differentiable functions. We have

$$\begin{aligned}\frac{fg(x_n) - fg(a)}{x_n - a} &= \frac{f(x_n)g(x_n) - f(x_n)g(a) + f(x_n)g(a) - f(a)g(a)}{x_n - a} \\ &= f(x_n)\frac{g(x_n) - g(a)}{x_n - a} + g(a)\frac{f(x_n) - f(a)}{x_n - a}.\end{aligned}$$

Now the terms on the right hand side have limits for $x_n \longrightarrow a$, hence so does the left hand side; this shows that fg has a derivative at $x = a$ if f and g do, and that $(fg)'(a) = f(a)g'(a) + f'(a)g(a)$.

The existence of neutral elements with respect to addition and multiplication is clear since constant functions are differentiable. Moreover, if f is differentiable at $a \in I$, then so is $-f$. □

Since $f(x) = x$ has derivative $f'(x) = 1$, repeated application of the product rule proves that $f(x) = x^n$ has derivative $f'(x) = nx^{n-1}$ for all $n \geq 0$. For negative n we need

Proposition 6.4. *If $f : I \longrightarrow \mathbb{R}$ is differentiable at $a \in I$, and if $f(a) \neq 0$, then $1/f$ is differentiable at a, and we have $(\frac{1}{f})'(a) = -f'(a)/f(a)^2$.*

Proof. Remember when we proved that $1/f$ is continuous at a if f is continuous at a? We proved that, under these assumptions, f is nonzero on some interval containing a, and I complained that Ross didn't prove this. Here's how he begins his proof that $1/f$ is differentiable at a:

> Since $f(a) \neq 0$ and f is continuous at a, there exists an open interval I containing a such that $f(x) \neq 0$ for $x \in I$.

Now we may use this since we proved it, so for any $x \in I$ we have

$$\frac{1}{f(x)} - \frac{1}{f(a)} = \frac{f(a) - f(x)}{f(a)f(x)} = -\frac{f(x) - f(a)}{x - a} \cdot \frac{x - a}{f(x)f(a)},$$

hence

$$\frac{1/f(x) - 1/f(a)}{x - a} = -\frac{f(x) - f(a)}{x - a} \cdot \frac{1}{f(x)f(a)},$$

and letting $x \longrightarrow a$ we see that $1/f$ is differentiable at $x = a$ and has derivative $-f'(a)/f(a)^2$. □

Now we claim

Proposition 6.5 (Chain Rule). *Assume that $f : I \longrightarrow \mathbb{R}$ is differentiable at $a \in I$, and that $g : J \longrightarrow \mathbb{R}$ is differentiable at $f(a) \in J$. Then $g \circ f$ is differentiable at a, and we have $(g \circ f)'(a) = f'(a)g'(f(a))$.*

Proof. Put $h = g \circ f$ and $b = f(a)$. Since f is differentiable at $x = a$ and g at $x = b$, there exist functions u, v such that

$$\begin{aligned} f(x) &= f(a) + (x - a)(f'(a) + u(x), \\ g(y) &= g(b) + (y - b)(g'(b) + v(y), \end{aligned}$$

where $\lim_{x \to 0} u(x) = \lim_{y \to 0} v(y) = 0$. Now

$$\begin{aligned} h(x) - h(b) &= g(f(x)) - g(f(a)) \\ &= [f(x) - f(a)] \cdot [g'(b) + v(y)] \\ &= (x - a)[f'(a) + u(x)] \cdot [g'(b) + v(y)], \end{aligned}$$

so for all $x \neq a$ we have

$$\frac{h(x) - h(a)}{x - a} = [g'(b) + v(y)] \cdot [f'(a) + u(x)].$$

If we let $x \to a$, then $y \to b$ since g is continuous at b, and the right hand side tends to $g'(b)f'(a)$ as claimed. □

Remember the theorem of the existence of inverse functions? We proved that strictly increasing functions $I \longrightarrow J = f(I)$ on intervals I have an inverse function $f^{-1} : J \longrightarrow I$, and that f^{-1} is continuous if f is. Now we shall prove

Theorem 6.6. *Let $f : I \longrightarrow \mathbb{R}$ be an injective function on an open interval I, and let $J = f(I)$. If f is differentiable at $a \in I$, and if $f'(a) \neq 0$, then $f^{-1} : J \longrightarrow I$ is differentiable at $b = f(a)$, and we have $(f^{-1})'(b) = 1/f'(a)$.*

Proof. Here using the second definition of differentiability simplifies the proof considerably: compare with Ross.

Since f is differentiable at a, there is a function $u : I \longrightarrow \mathbb{R}$ with $\lim_{x\to a} u(x) = 0$ and

$$f(x) = f(a) + (x - a)(f'(a) + u(x)). \tag{6.1}$$

Now we are interested in $\frac{g(y)-g(b)}{y-b}$, where $g = f^{-1}$; with $g(y) = x$ and $g(b) = a$, this is nothing but $\frac{x-a}{f(x)-f(a)}$; using (6.1), this becomes

$$\frac{g(y) - g(b)}{y - b} = \frac{x - a}{f(x) - f(a)} = \frac{1}{f'(a) + u(x)}.$$

Now we can write $1/(f'(a) + u(x)) = 1/f'(a) + v(x)$, where $\lim_{x\to a} v(x) = 0$; thus

$$\frac{g(y) - g(b)}{y - b} = \frac{1}{f'(a)} + v(x),$$

which implies that g is differentiable at $y = b$ with $g'(b) - 1/f'(a)$. □

Example: if $f(x) = x^2$, then $g = f^{-1}$ is given by $g(y) = \sqrt{y}$ as long as $y \geq 0$. In order to apply Theorem 6.6 we need f to be injective, so we have to restrict the domain of f to the nonnegativ reals. There $f'(a) \neq 0$ as long as $a \neq 0$, so let us assume that $x > 0$. Then g is differentiable at $b = f(a) = a^2$, and we have $g'(b) = 1/f'(a) = 1/2a = 1/2\sqrt{b}$.

We say that f attains a local maximum at x_0 if there is a $\delta > 0$ such that $f(x) \leq f(x_0)$ for all $x \in (x_0-\delta, x_0+\delta)$. The following criterion gives a necessary condition for the existence of local extrema:

Theorem 6.7. *Let $f : I \longrightarrow \mathbb{R}$ be a differentiable function on some open interval I. If f attains a local maximum or a minimum at $a \in I$, then $f'(a) = 0$.*

Proof. Again, this is a point where the second definition comes in handy: there is a function $u(x)$ with $\lim_{x\to a} u(x) = 0$ and

$$f(x) = f(a) + (x - a)(f'(a) + u(x)).$$

Now if f attains a local maximum at a, then $f(x) \leq f(a)$ for all x close to a. But this inequality is equivalent to $(x - a)(f'(a) + u(x)) \leq 0$. If $x < a$, this implies $f'(a) \geq u(x)$, and letting $x \longrightarrow a$ we find $f'(a) \geq 0$. If $x > a$, on the other hand, we find $f'(a) \leq u(x)$, and letting $x \longrightarrow a$, we find $f'(a) \leq 0$. Thus $f'(a) = 0$.

This takes care of the case when f has a maximum at a. If f has a minimum at a, then $-f$ has a maximum, hence $-f'(a) = 0$ by what we just proved. □

Note that the condition $f'(a) = 0$ does not guarantee the existence of a maximum or minimum at a, as the example $f(x) = x^3$ with $a = 0$ shows.

The mean value theorem for differentiable functions is going to be our key to the fundamental theorem of calculus (it is also the key to finding formulas for the arc length of (smooth) curves, curvature, etc.):

Theorem 6.8 (Mean Value Theorem). *If $f : I \longrightarrow \mathbb{R}$ is differentiable on $I = (a, b)$, then there is an $x \in I$ such that*

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

We shall prove a special case first:

Theorem 6.9 (Rolle's Theorem). *Let $f : I \longrightarrow \mathbb{R}$ be continuous on $I = [a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists an $x \in (a, b)$ such that $f'(x) = 0$.*

Proof. Since f is continuous on a closed interval, it assumes its maximum and minimum somewhere, so there are $x_0, y_0 \in I$ such that $f(x_0) \leq f(x) \leq f(y_0)$. If $x_0 = a$ and $y_0 = b$ or vice versa, then f is constant since $f(a) = f(b)$ by assumption. But then $f'(x) = 0$ for all $x \in (a, b)$. Otherwise, f assumes either a maximum at $y_0 \in (a, b)$ or a minimum in $x_0 \in I$, and then $f'(x) = 0$ for $x = x_0$ or $x = y_0$. □

The Mean Value Theorem now follows by applying Rolle's theorem to $g(x) = f(x) - \frac{f(b)-f(a)}{b-a} f(x)$.

Proposition 6.10. *Let $f : I \longrightarrow \mathbb{R}$ be a differentiable function on some open interval I. Then f is strictly increasing on I if $f'(a) > 0$ for all $a \in I$.*

Proof. Assume that $f'(a) > 0$ for all $a \in I$. If $x_1 < x_2$, then by the mean value theorem

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x)$$

for some $x \in (x_1, x_2)$; since $f'(x) > 0$, this implies $f(x_2) > f(x_1)$, hence f is strictly increasing. □

Note that the converse is not completely true (it is almost true, however): the function $f(x) = x^3$ is strictly increasing on $(-1, 1)$ but $f'(0) = 0$.

6.2 Fundamental Theorem of Calculus

Let $f : [a, b] \longrightarrow \mathbb{R}$ be integrable, and consider a subinterval $[x, x + \Delta]$ of $[a, b]$. The area below f in this subinterval equals the product of the length Δ of the interval and the approximate height $f(x)$. If F is a function such that $F(x + \Delta) - F(x)$ gives the area below f in $[x, x + \Delta]$, then $F(x + \Delta) - F(x) \approx \Delta \cdot f(x)$, thus $\frac{F(x+\Delta)-F(x)}{\Delta} \approx f(x)$. Letting $\Delta \to 0$, we get $F'(x) = f(x)$.

This argument involves lots of hand-waving, and can be made exact by replacing the estimate $F(x+\Delta)-F(x) \approx \Delta \cdot f(x)$ by something exact using the mean value theorem. Of course we have to do this for every small subinterval defined by a partition to get good results.

Theorem 6.11 (Fundamental Theorem of Calculus). *Let $F : [a,b] \longrightarrow \mathbb{R}$ be a function differentiable in (a,b). If $f = F'$ has a Riemann integral, then*

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

Proof. The fact that f has a Riemann integral means that for every $\varepsilon > 0$ there exists a partition $P = \{a = t_0 < t_1 < \ldots < t_n = b\}$ of $[a,b]$ such that $U(f,P) - L(f,P) < \varepsilon$.

In each subinterval $[t_i, t_{i+1}]$ of this partition we apply the mean value theorem to F and get $F(t_{i+1}) - F(t_i) = (t_{i+1} - t_i)f(\xi_i)$ for some $\xi_i \in (t_i, t_{i+1})$. Summing all these equations (telescope sums!) gives

$$F(b) - F(a) = \sum f(\xi_i)(t_{i+1} - t_i).$$

Now $m(f, [t_i, t_{i+1}]) \leq f(\xi_i) \leq M(f, [t_i, t_{i+1}])$, hence $L(f,P) \leq \sum f(\xi_i)(t_{i+1} - t_i) \leq U(f,P)$. Thus $L(f,P) \leq F(b) - F(a) \leq U(f,P)$, and taking the infimum over the left and the supremum over the right inequality we see $L(f) \leq F(b) - F(a) \leq U(f)$. Since f is Riemann integrable, we have $L(f) = U(f) = \int_a^b f(x)dx$, hence this must equal $F(b) - F(a)$ as claimed. $\square$