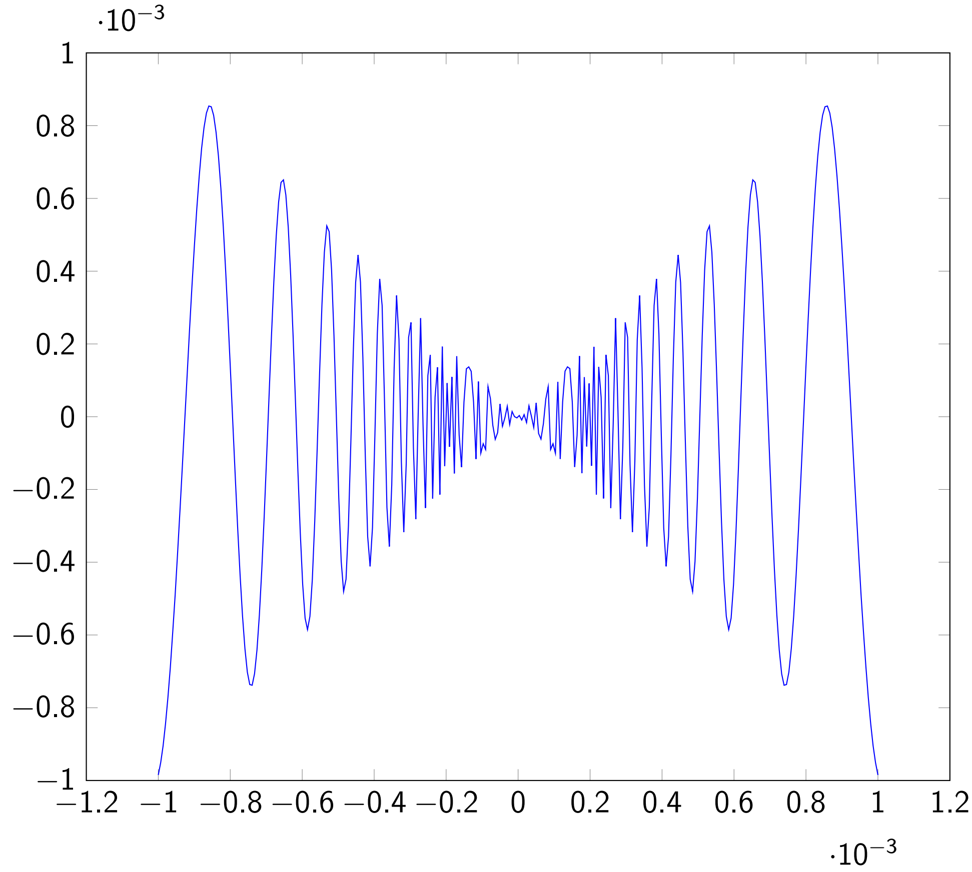
Proofs in Real Analysis



$f(x) = x \sin\left(\frac{1}{x}\right)$

Quasar Chunawala

Contents

Chapter 1

The Natural numbers $\mathbf{N}$.

Real analysis is the study of real numbers, sequences and series of real numbers and real-valued functions. Real analysis is the theoretical foundation which underlies calculus, which is the collection of computational algorithms which one uses to manipulate functions.

1.1 Peano axioms.

One standard way to define the natural numbers $\mathbf{N}$ is in terms of the *Peano axioms*, which were first laid out by Guiseppe Peano (1858-1932). How are we to define what natural numbers are? Informally, we could say -

Definition 1.1. A natural number is any element of the set

$$\mathbf{N} := \{0, 1, 2, 3, 4, 5, 6, \ldots\}$$

which is the set of all numbers created by starting with 0 and then counting forward indefinitely. We call $\mathbf{N}$ the set of natural numbers.

Axiom 1.1.1. 0 *is a natural number.*

Axiom 1.1.2. *If* n *is a natural number, then* $n{+}{+}$ *is also a natural number.*

Thus, for instance from axiom (1) and two applications of axiom (2), we see that $(0{+}{+}){+}{+}$ is a natural number. Of course, this notation gets unwieldy, so we adopt a convention to write these numbers in more familiar notation:

Definition 1.2. We define 1 to be the number $0{+}{+}$, 2 to be the number $(0{+}{+}){+}{+}$, 3 to be the number $((0{+}{+}){+}{+}){+}{+}$ etc. (In other words, we have $1{:=}0{+}{+}$, $2{:=}1{+}{+}$, $3{:=}2{+}{+}$)

***Theorem:* 1.1.1**

3 is a natural number.

Proof. By axiom (1), 0 is a natural number. By axiom (2), the successor of any natural number n is also a natural number. So, $0{+}{+} = 1$ is a natural number. Again, $1{+}{+} = 2$ is a natural number. Finally, $2{+}{+} = 3$ is a natural number. $\square$

It may seem that this is enough to describe the natural numbers. However, we have not completely pinned down the behaviour of $\mathbf{N}$.

Example 1.1.1. Consider a number system which consists of the numbers $0, 1, 2, 3$ in which the increment operation wraps back from 3 to 0. More precisely, $0{+}{+}$ is equal to 1, $1{+}{+}$ is equal to 2, $2{+}{+}$ is equal to 3, $3{+}{+}$ is equal to 0 (and also equal to 4, by the definition of 4). This type of thing actually happens in real life, when one uses a computer to try to store a natural number: if one starts at 0 and performs the increment operation repeatedly, eventually the computer will overflow its memory and the number wraps around back to 0 (though this may take quite a large number of incrementation operations, for instance a 2-byte representation of an integer will wrap around only after $65,536$ increments.) Note that this type of number system obeys axiom (1) and axiom (2) even though it clearly does not correspond to what we intuitively believe the natural numbers to be like.

To prevent this sort of wrap-around issue we will impose another axiom.

Axiom 1.1.3. 0 *is not the successor of any natural number. That is, we have* $n{+}{+} \neq 0$ *for every natural number* n.

***Theorem:* 1.1.2**

4 is not equal to 0.

Don't laugh! Because of the way we have defined 4 - it is the increment of the increment of the increment of the increment of 0 - it is not necessarily true a priori, that this number is not the same as zero, even if it is obvious. Note for instance, in the example discussed before, 4 was indeed equal to 0 and that in a standard two byte computer representation of natural numbers, for instance $65,536$ equals to 0 (using our definition of $65,536$ is equal to 0 increment 65636 times).

Proof. By definition $4 = 3{+}{+}$. By axioms (1) and (2), 3 is a natural number. By axiom (3), 0 is not the successor of any natural number. $n{+}{+} \neq 0$ for any natural number n. So, $3{+}{+} \neq 0$. $\implies 4 \neq 0$. $\square$

However, even with our new axiom, it is still possible that our number system behaves in a pathological way.

Example 1.1.2. Consider a number system consisting of the five numbers $0, 1, 2, 3, 4$ in which the increment operation hits a ceiling at 4. More precisely, suppose that $0{+}{+} = 1$, $1{+}{+} = 2$, $2{+}{+} = 3$, $3{+}{+} = 4$, but $4{+}{+} = 4$ (or in other words, $5 = 4$ and hence $(4{+}{+}){+}{+} = 4{+}{+} = 4$ i.e. $6 = 4$, $7 = 4$ etcetera. This does not contradict the axioms (1), (2) and (3).

There are many ways to prohibit the above types of behaviour from happening, but one of the simplest is to assume the following axiom:

Axiom 1.1.4. *Different natural numbers must have different successors. That is* n, m *are natural numbers and* $n \neq m$*, then* $n{+}{+} \neq m{+}{+}$*. Equivalently, if* $n{+}{+} \neq m{+}{+}$*, then we must have* $n = m$*.*

***Theorem:* 1.1.3**

6 is not equal to 2

Proof. Suppose for the sake of contradiction, assume $6 = 2$. Then, $5{+}{+} = 1{+}{+}$. So, by axiom (4), we have $5 = 1$. so that $4{+}{+} = 0{+}{+}$. By axiom (4) again, we then have $4 = 0$, which contradicts the previous proposition. Our initial assumption is false. □

There is however still one more problem: while axioms (1) and (2) allow us to conform that $0, 1, 2, 3, 4, ...$ are distinct elements of $\mathbf{N}$, there is the problem that there may be other rogue elements in our number system which are not of this form:

Example 1.1.3. Suppose that our number system $\mathbf{N}$ consisted of the following collection of integers and half-integers:

$$\mathbf{N} := \{0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, ...\}$$

(This example is informal, since we are using real numbers which are not supposed to use yet.) One can check that axioms (1)-(4) are still satisfied for this set.

What we want is some axiom which says that the only numbers in $\mathbf{N}$ are those which can be obtained from 0 and the incrementation in order to exclude rogue elements such as 0.5.

Axiom 1.1.5. *Principle of Mathematical Induction. Let* $P(n)$ *be any statement about the natural number* n*. Suppose that* $P(0)$ *is true and suppose that whenever* $P(n)$ *is true, P(n++) is also true. Then* $P(n)$ *is true for every natural number* n*.*

Remark 1.1. We are a little vague on what property means at this point, but some possible examples of $P(n)$ might be n is equal to 3, n solves the equation $(n+1)^2 = n^2 + 2n + 1$ and so forth. Of course we haven't defined many of these concepts yet, but when we do axiom (5) will apply to these properties. A logical remark - because this axiom refers to not just variables, but also properties, it is of a different nature than the other four axioms; indeed axiom (5) should be called an axiom schema rather than an axiom - it is a template for producing an infinite number of axioms, rather than being a single axiom in its own right.

The informal intuition behind this axiom is the following. Suppose, $P(n)$ is such that $P(0)$ is true, and such that whenever $P(n)$ is true, P(n++) is true. Then, since $P(0)$ is true, P(0++)=P(1) is true. Since P(1) is true, P(1++)=P(2) is true. Repeating this indefinitely, we see that $P(0), P(1), P(2), P(3), ...$ are all true - however this line of reasoning will never conclude that $P(0.5)$ for instance is true. Thus, axiom (5) should not hold for number systems which contain unnecessary elements such as 0.5. Indeed, one can give a proof of this fact. Apply axiom (5) to the property $P(n) = n$ is not a half-integer, that is an integer plus 0.5. Then, $P(0)$ is true and if $P(n)$ is true, P(n++) is true. Thus, axiom (5) asserts that $P(n)$ is true for all natural numbers n, that is, there is no natural number that can be a half-integer. In particular, 0.5 cannot be a natural number.

Axioms (1)-(5) are called the Peano axioms for the natural numbers. They are all very plausible and so we shall make the

Assumption. There exists a number system $\mathbf{N}$, whose elements we will call natural numbers, for which axioms (1)-(5) are true.

Here is is another way to view axiom (5).

Proposition 1.1.1. *A subset of* $\mathbf{N}$ *which contains* 0 *and which contains* $n+1$*, whenever it contains* n *must equal* $\mathbf{N}$.

Proof. Assume that $\mathbf{N}$ contains a set S such that:

1) $0 \in S$

2) If $n \in S$, then $n+1 \in S$

and yet, $\mathbf{N} \neq S$.

Consider the set difference set $\{x : x \in \mathbf{N}, x \notin S\}$. Let n_0 be the smallest member in this set. Because of our first assumption, $n_0 \neq 0$. As n_0 is a natural number, it is the successor of number $n_0 - 1$. As n_0 is the smallest number in $\mathbf{N} - S$, $n_0 - 1 \in S$. Because of assumption two, if $n_0 - 1 \in S$, then the successor n_0 belongs to S, which is a contradiction. Our initial assumption is wrong.

$S = \mathbf{N}$. □

Remark 1.2. Tao explains, that our definition of natural numbers is *axiomatic* rather than constructive. We have not told you, what the natural numbers are (so we do not address questions such as what the numbers are made up of, are they physical objects, what do they measure etc.) - we have only listed some things that you can do with them (in fact the only operation we have defined on them right now is the increment one) and some of the properties they have. This is how mathematics works - it treats its objects *abstractly*, caring only about what properties the objects have, not what the objects are or what they mean. If one wants to do mathematics, it does not matter whether a natural number means a certain arrangement of beads on an abacus, or a certain organisation of bits in computer memory, or some more abstract concept with no physical substance; as long as you can increment them, see if two of them are equal, and later on do other arithmetic operations such as add and multiply, they qualify as numbers for mathematical purposes (provided they obey the requisite axioms of course). It is possible to construct the natural numbers from other mathematical objects - from sets, for instance - but there are multiple ways to construct a working model of the natural numbers, and it is pointless atleast from a mathematician's standpoint, as to argue about which model is the "true" one - as long as it obeys all the axioms and does all the right things, that's good enough to do maths.

Remark 1.3. Historically, the realization that numbers could be treated axiomatically is really very recent, not much more than a hundred years old. Before then, numbers were generally understood to be inextricably connected to some external concept, such as counting the cardinality of a set, measuring the length of a line segment, or the mass of a physical object, etc. This worked reasonably well, until one was forced to move from one number system to another, for instance understanding numbers in terms of counting beads, for example is great for conceptualizing the numbers 3 and 5, but doesn't work so well for -3 or $1/3$ or $\sqrt{2}$ or $3+4i$; thus each great advance in the theory of numbers - negative numbers, irrational numbers, complex numbers, even the number zero - led to a lot of unnecessary philosophical anguish. The great discovery of the late nineteenth century was that numbers can be understood abstractly via axioms, without necessarily needing a concrete model; of course a mathematician can use any of these models when it is convenient, to aid his or her intuition and understanding, but they can just as easily be discarded when they begin to get in the way.

***Theorem:* 1.1.4**

Suppose for each natural number n, we have some function $f_n : \mathbf{N} \to \mathbf{N}$ from the natural numbers to the natural numbers. Let c be a natural number. Then we can assign a single value to a_n for all natural numbers n, such that $a_0 = c$ and $a_{n++} = f_n(a_n)$ for each natural number n.

Proof. Our claim is that we define recursive sequences. Specifically, we claim that this procedure gives a single value to each element of the sequence $a_0, a_1, \dots, a_n, a_{n++}, \dots$

We use induction. We first observe, that this procedure assigns a single value to a_0, namely c. None of the other definitions $a_{n++} = f_n(a_n)$ will redefine the value of a_0, because of axiom (3) (0 is not the successor of any natural number).

Now, suppose inductively that the procedure gives a single value to a_n. Then, it gives a single value to $a_{n++} = f_n(a_n)$. None of the other definitions, $a_{m++} = f_m(a_m)$ will redefine the value of a_{n++}, because by axiom (4), if $m \neq n$, then by $m++ \neq n++$. This completes the induction. □

1.2 Addition of natural numbers.

Definition 1.3. Addition of natural numbers. Let m be a natural number. To add zero to m, we define $0 + m := m$. Now suppose, inductively, we have defined how to add n to m. Then, we can add $n{+}{+}$ to m by defining -

$$(n++) + m := (n+m)++$$

For instance, we have

$$\begin{aligned}
2+3 &= (1+3)++ \\
&= ((0+3)++)++ \\
&= ((3)++)++ \\
&= (4)++ \\
&= 5
\end{aligned}$$

From our discussion of recursion in the previous section, we see that have defined addition $n+m$ for every natural number n. Here, we are specialising the previous general discussion to the setting where

$$\begin{aligned}
a_0 &= m \\
a_{n++} &= a_n++
\end{aligned}$$

where $a_n = (n+m)$.

Thus, if $m = 3$, then we can recursively define the addition of natural number n with 3 as

$$\begin{aligned}
&a_0 = 3 \\
&a_1 = a_0++ = 4 \\
&a_2 = a_1++ = 5 \\
&a_3 = a_2++ = 6 \\
&\;\vdots
\end{aligned}$$

Note that the above definition of addition is asymmetric: $3+5$ is incrementing 5 three times, while $5+3$ is incrementing 3 five times. Of course, they both yield the value 8. More generally, it is a fact which we shall prove shortly, that $a+b = b+a$ for all natural numbers a, b although this is not immediately clear from the definition. Notice that we can prove easily, that the sum of two natural numbers is a natural number.

***Theorem:* 1.2.1: Sum of natural numbers.**

The sum of two natural numbers n and m, is a natural number.

Proof. Let $P(n)$ be the proposition, that the sum of n and a fixed natural number$m$ is also a natural number. We will induct on n keeping m fixed.

By axiom (1), 0 is a natural number. By the definition of addition, $0 + m := m$. So, $P(0)$ is true.

Let us assume that $P(n)$ holds. So, suppose that $(n + m)$ is a natural number.

By the definition of addition, $(n++) + m := (n + m) + +$. From Peano's axioms, the successor of a natural number n, $n{+}{+}$ is also a natural number. As $(n + m)$ is a natural number, $(n + m) + +$ is a natural number. Therefore, $P(n + +)$ is true, whenever $P(n)$ is true.

This closes the proof. □

Lemma 1.2.1. *For any natural number n, $n + 0 = n$*

Note that we cannot deduce this immediately from $0 + m = m$, because we do not yet know that $a + b = b + a$.

Proof. We use induction. The base case $0 + 0 = 0$ follows from the definition of addition, since we know that $0 + m := m$, for every natural number m.

Now suppose, inductively that $n + 0 = n$. We wish to show that $(n++) + 0 = (n++)$. But by the definition of addition, $(n++) + 0 := (n + 0) + + = n + +$, since $n + 0 = n$. This closes the proof. □

Lemma 1.2.2. *For any natural numbers n and m, $n + (m++) = (n + m) + +$.*

Again we cannot deduce this yet from $(n++) + m = (n + m) + +$, because we do not know that yet that $a + b = b + a$.

Proof. We induct on n keeping m fixed. We first consider the base case $n = 0$. In this case, we have to prove that $0 + (m++) = (0 + m) + +$. By the definition of addition, the left hand side can be simplified as $m + +$, since zero plus anything, is that thing. On the right hand size, $(0 + m) + + = (m) + +$. So, both sides are equal to $m + +$, and therefore are equal to each other.

We assume inductively that, $n + (m + +) = (n + m) + +$. We now have to show that $(n++) + (m++) = ((n++) + m) + +$. Indeed the left hand side is $(n++) + (m++) = (n + (m++)) + +$ (using the definition of addition). Further, applying the inductive hypotheses, $(n + (m + +)) + + = ((n + m) + +) + +$. Similarly, on the right hand side, applying the definition of addition $((n++) + m) + + = ((n + m) + +) + +$. Thus, both sides are equal to $((n + m) + +) + +$ and are therefore equal to each other. This closes the induction. □

Corollary 1.2.1. *The successor of n is $n + 1$.*

Proof. We can simplify the expression on the left hand side as -

$$\begin{aligned} n++ &= (n+0)++ \\ &= n+(0++) \\ &= n+1 \end{aligned}$$

□

Theorem: 1.2.2: Commutativity of addition.

The addition of natural numbers is commutative. For any natural numbers n and m,

$$n+m=m+n$$

Proof. We induct on n, keeping m fixed.

We start with $n=0$ as the base case. On the left hand side, $0+m := m$ by the definition of addition. Moreover, from lemma (1) $m+0=m$. So, both the sides are equal to m and therefore equal to each other.

We inductively assume that, $n+m=m+n$.

We are interested to prove that $(n++)+m = m+(n++)$. The left hand side $(n++)+m :=$ $(n+m)++$ by the definition of addition. Further, $m+(n++) = (m+n)++$ by lemma (2). Moreover, $n+m=m+n$ from the inductive hypotheses. We also know, from Peano's axioms, that if $a=b$, then their successors are equal; $a++=b++$. So, $(n+m)++=(m+n)++$. This closes the induction. □

Theorem: 1.2.3: Associativity of addition.

For any natural numbers a, b, c, we have -

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

Proof. We shall use induction keeping a, c fixed and inducting on b.

We start with $b=0$ as the base case. We are interested to show that $(a+0)+c = a+(0+c)$. On the left hand side, $a+0=a$ by lemma (1), so the LHS becomes $(a+0)+c = a+c$. On the right hand side, $(0+c) := c$ by the definition of addition. So, $a+(0+c) = a+c$. Both sides are equal to each other. $P(0)$ holds.

We inductively assume $P(b)$. So, $(a+b)+c = a+(b+c)$.

We are interested to prove that $(a+(b++))+c = a+((b++)+c)$.

LHS.

$$\begin{aligned}(a+(b++))+c &= (a+b)+++c && \text{From lemma (2)}\\ &= ((a+b)+c)++ && \text{By the definition of addition}\end{aligned}$$

RHS.

$$\begin{aligned}a+((b++)+c) &= a+(b+c)++ && \text{By the definition of addition}\\ &= (a+(b+c))++ && \text{By lemma (2)}\end{aligned}$$

But, $(a+b)+c = a+(b+c)$ from the inductive step. By Peano's axioms, their successor elements should also be equal to each other. This closes the induction. □

Because of this associativity, we can write sums such as $a+b+c$ without having to worry about which order the numbers are being added together.

***Theorem:* 1.2.4: Cancellation law.**

Let a, b, c be natural numbers, such that $a+b = a+c$. Then, we have $b = c$.

Note that, we cannot use subtraction or negative numbers to prove this proposition, because we have not developed these concepts yet. In fact, this cancellation law is crucial in letting us define subtraction and integers later on, because it allows for a sort of virtual subtraction, even before subtraction is officially defined.

Proof. We prove this by inducting on a.

We consider $a = 0$ as the base case, and show that if $0+b = 0+c$, then $b = c$. By definition of addition, the left hand side is $0+b := b$, the right hand side is $0+c := c$. So, $b = c$.

We inductively assume that if $a+b = a+c$, then $b = c$.

We are interested to prove that if $(a++)+b = (a++)+c$, then $b = c$.

LHS.

$$(a++)+b = (a+b)++ \qquad \text{By the definition of addition}$$

RHS.

$$(a++)+c = (a+c)++ \qquad \text{By the definition of addition}$$

But, $a+b = a+c$. By Peano's axiom's, their successors must be equal. So, $(a+b)++$ must equal $(a+c)++$. Consequently, both sides are equal to each other. This closes the induction. □

Definition 1.4. Positive natural numbers. A natural number is said to be positive if and only if, it is not equal to 0.

***Theorem:* 1.2.5**

If a is positive and b is a natural number, then $a+b$ is also positive (and hence $b+a$ is also positive).

Proof. We induct on b, keeping a fixed.

If $b=0$, then $a+0=a$ is a positive natural number, so this proves the base case.

We inductively assume that if a is positive and b is a natural number, the sum $a+b$ is positive.

We are interested to show that if a is positive and $b++$ is a natural number, the sum $a+(b++)$ is positive. By Lemma (2), $a+(b++)=(a+b)++$. Since, a and b are natural numbers, the sum $a+b$ is also a natural number. From Peano's axioms, 0 is not the successor of any natural number. Therefore, $(a+b)++\neq 0$. So, $a+(b++)$ is positive. This closes the induction. ☐

Corollary 1.2.2. *If a and b are natural numbers, such that $a+b=0$, then $a=0$ and $b=0$.*

Proof. Suppose for the sake of contradiction $a\neq 0$ and $b\neq 0$. If $a\neq 0$, then a is positive and $a+b=0$ is positive by the above theorem. But, this is a contradiction. Similarly, if $b\neq 0$, then b is positive and again $a+b=0$ is positive, a contradiction. Thus if, $a+b=0$, both a and b must be zero. ☐

Lemma 1.2.3. *Let a be a positive number. Then, there exists exactly one natural number b such that $b++=a$.*

Proof. Assume there are two numbers b and c, $b\neq c$, such that $b++=a$ and $c++=a$. Both these quantities are equal to each other: $b++=c++$. By Peano's axioms, if m,n are natural numbers, $m\neq n$, then $m++\neq n++$. The contrapositive of this statement is, if $m++=n++$, it implies that $m=n$. So, in our case, b must be equal to $c/$. This contradicts our initial assumption. Hence, proved. ☐

Definition 1.5. Ordering of natural numbers. Let n and m be natural numbers. We say that, $n\geq m$ or $m\leq n$, if and only if $n=m+a$ for some natural number a. We say that n is strictly greater than m, and write $n>m$ or $m<n$, if and only if $n\geq m$ and $n\neq m$.

Theorem: 1.2.6: Basic properties of the order for natural numbers.

Let a, b, c be natural numbers. Then,
(a) Order is reflexive. $a \geq a$.
(b) Order is anti-symmetric. If $a \geq b$ and $b \geq a$, then $a = b$.
(c) Order is transitive. If $a \geq b$ and $b \geq c$, then $a \geq c$.
(d) Addition preserves order. $a \geq b$ if and only if $a + c \geq b + c$.
(e) $a < b$ if and only if $a{+}{+} \leq b$.
(f) $a < b$ if and only if $b = a + d$ for some positive number d.

Proof. (a) Order is reflexive. We claim that $a \geq a$. We can write the natural number a as,

$$a = a + 0$$

where 0 is a natural number. This proves the claim.

(b) Order is anti-symmetric. We claim that if $a \geq b$ and $b \geq a$, then $a = b$.

$(a \geq b)$ implies that $a = b + k$ for some natural number k, by the definition of $\geq$ relation. Likewise $(b \geq a)$ implies that $b = a + l$ for some natural number l. Adding both the equations, we have $(a+b) = (b+a)+(k+l)$. But, $a+b = b+a$, as addition is commutative. Hence, we can cancel this term on both the sides (cancellation law). That leaves us with $k + l = 0$ where k and l are natural numbers. But, this implies that $k = l = 0$. Therefore, $a = b$.

(c) Order is transitive. Our claim is that if $a \geq b$ and $b \geq c$, then $a \geq c$.

$(a \geq b)$ implies $a = b + k$, where k is a natural number. $(b \geq c)$ implies $b = c + l$, where l is a natural number. Adding we have $a + b = b + c + (k + l)$. Cancelling b from both sides we are left with $a = c + (k + l)$. The sum of two natural numbers $k + l$ is also a natural number. Hence, $a \geq c$.

(d) Addition preserves order. Our claim is that $a \geq b$ if and only if $a + c \geq b + c$.

$\longrightarrow$ direction.

Suppose $a \geq b$. Then $a = b + k$ for some natural number k, by the definition of the order relation. Adding a natural number c to both sides of the equality, we have $(a+c) = (b+c)+k$. So, $(a + c) \geq (b + c)$.

$\longleftarrow$ direction. Suppose $(a + c) \geq (b + c)$. By definition, $(a + c) = (b + c) + k$ for some natural number k. Using the cancellation law, we can cancel the term c from both sides of the equality to obtain $a = b + k$, where k is a natural number. Hence, $a \geq b$.

(e) Our claim is that $a < b$ if and only if $a{+}{+} \leq b$.

$\longrightarrow$ direction.

We are given that $b > a$. This means that $b \geq a$ and $b \neq a$. $b \geq a$ means that $b = a + k$ for some natural number k. Exactly one of the following two possibilities must be true.

Either $k = 0$ or $k \neq 0$. If $k = 0$, this line of reasoning leads us to believe $a = b$, which is a contradiction. We know that $a \neq b$. So, we are left with $k \neq 0$. k is some positive natural number. There exists exactly one natural number m, such that $m++ = k$. So, the equality $b = a + k$ can be written as $b = a + (m++)$ for some natural number m. $(a + (m++)) = (a + m)++ = ((a++) + m)$. Thus, $b = (a++) + m$ for some natural m. $b \geq a++$ as desired.

$\longleftarrow$ direction.

We are given that $b \geq a++$. By definition of the order relation, this implies that $b = (a++) + k = (a + k)++ = (a + (k++))$. From Peano's axioms , $k++ \neq 0$, so $b \neq a$. Therefore, $b \geq a$ and $b \neq a$. So, $b > a$.

(f) Our claim is that $a < b$ if and only if $b = a + d$ for some positive number d.

$\longrightarrow$ direction.

We are given that $b > a$. So, $b \geq a$ and $b \neq a$. Therefore, $b = a + d$, $d \in \mathbf{N}$, and $b \neq a$. If $d = 0$, then $b = a$ which is a contradiction. So, $d \neq 0$. Hence, $b = a + d$, where d is a positive natural number.

$\longleftarrow$ direction.

Conversely, suppose $b = a + d$ where d is a positive natural number. By definition, $d \neq 0$. If $b = a$, then the left hand side becomes $a = a + d$. Cancelling a on both sides, $d = 0$. This line of reasoning leads to a contradiction. So, the only feasible solution is $b \neq a$. Thus, $b \geq a$ and $b \neq a$. Consequently, $b > a$ as desired. $\square$

Theorem: 1.2.7: Trichotomy of order for natural numbers.

Let a and b be natural numbers. Then, exactly one of the following statements is true: $a < b$, $a = b$, $a > b$.

Proof. First we show that, we cannot have more than one of the statements $a < b, a = b, a > b$ holding at the same time. If $a < b$, then $a \neq b$ by definition and if $a > b$ then $a \neq b$ by definition. If $a > b$ and $a < b$, then (i) implies $a = b + k$, (ii) implies $b = a + l$, adding (i) and (ii) yields $a + b = a + b + k + l$, so $k + l = 0$. Therefore, $k = l = 0$. Hence, $a = b$. This is a contradiction, since $a \neq b$. Thus, no more than one of the statements is true.

Now, we show that atleast one of the statements is true. We keep b fixed and induct on a. When $a = 0$, then $0 \leq b$ for all b. This is true, because $b = 0 + b$, hence so $b \geq 0$. This proves the base case. We inductively assume that the proposition for a holds, and now we prove the proposition for $a++$. If $a < b$, then $b = a + d$ where d is a positive natural number. If d is a positive natural number, then there exists atleast one natural c, such that $c++ = d$. So, $b = a + (c++) = (a + c)++ = ((a++) + c$. If $c = 0$, then $a++ = b$. If $c \neq 0$, that is c is positive, then $a++ < b$. Atleast one of these statements must be true.

Next, suppose $a = b$, then $a++ = b + 1$, so $a++ > b$.

Finally, if $a > b$, then $a = b + d$ where d is a positive natural number. $a{+}{+} = (b+d){+}{+} = (b + (d{+}{+}))$ where $d{+}{+} \neq 0$. Consequently, $a{+}{+} > b$. □

1.3 Multiplication.

In the previous section, we have proven all the basic facts that we know to be true about addition and order. To save space and avoid belaboring the obvious, we will now allow ourselves to use all rules of algebra concerning addition and order, that we are familiar with, without further comment. Thus, for instance, we may write things like $a + b + c = c + b + a$ without supplying any further justification. Now, we introduce multiplication. Just as addition is iterated increment operation, multiplication is iterated addition.

Definition 1.6. Multiplication of natural numbers. Let m be a natural number. To multiply zero to m, we define $0 \times m := 0$. Now, suppose inductively we have define how to multiply n to m. Then we can multiply $n{+}{+}$ to m, by defining $(n{+}{+}) \times m := (n \times m) + m$.

Thus, for instance $0 \times m = 0$, $1 \times m = 0 \times m + m = m$, $2 \times m = 0 \times m + m + m$ etc. By induction, one can easily verify that the product of natural numbers is a natural number.

***Theorem:* 1.3.1**

The product of two natural numbers is a natural number.

Proof. Let n and m be natural numbers. Our claim is that the product $n \times m$ is a natural number. We induct on n, keeping m fixed.

By the definition of the multiplication of natural numbers, $0 \times m := 0$. As 0 is a natural number, the proposition is true in the base case $n = 0$.

We inductively assume, that the product of two natural numbers $n \times m$ is a natural number. We are interested to prove that $(n{+}{+}) \times m$ is also a natural number. $(n{+}{+}) \times m := (n \times m) + m$. We know that, $n \times m$ and m are natural numbers. The sum of two natural numbers is a natural number. Hence, the claim is proved. □

Proposition 1.3.1. *For any natural number n, $n \times 0 := 0$.*

Proof. We use induction on n. Let $P(n)$ be the proposition that $n \times 0 := 0$.

$P(0)$ is true. By the definition of multiplication of natural numbers, $0 \times 0 := 0$

We inductively assume that $P(n)$ is true. That is, $n \times 0 = 0$. We are interested to prove that $(n{+}{+}) \times 0 = 0$. By the definition of multiplication, we have $(n{+}{+}) \times 0 = (n \times 0) + 0 = 0 + 0 = 0$. This closes the proof. □

Proposition 1.3.2. *For any natural numbers n, m, $n \times (m{+}{+}) := (n \times m) + n$.*

Proof. We induct on n, keeping m fixed.

$P(0)$ is true. $0 \times (m++) := 0$ by the definition of multiplication. Moreover, $o \times (m++) = (0 \times m) + 0 = 0 + 0 = 0$. The left hand side and the right hand side are equal to 0 and are therefore equal to each other.

We inductively assume that $n \times (m++) = (n \times m) + n$. We are interested to prove that $(n{++}) \times (m{++}) = ((n{++}) \times m) + (n{++})$. The left hand side evaluates to $(n{++}) \times (m{++}) = (n \times (m++)) + (m++) = (n \times m) + n + (m++) = (n \times m) + (n + m)++$. The right hand side evaluates to $((n++) \times m) + (n++) = (n \times m) + m + (n++) = (n \times m) + (m + n)++ = (n \times m) + (n + m)++$. Thus, both sides are equal to $(n \times m) + (n + m)++$ and hence they are equal to each other. This closes the proof. □

Theorem: 1.3.2: Multiplication is commutative.

Multiplication of natural numbers is commutative. Let n, m be natural numbers. Then, $n \times m = m \times n$.

Proof. We induct on n, keeping m fixed.

Let $n = 0$. Then, $0 \times m = m \times 0 = 0$. This proves the base case.

We inductively assume that $n \times m = m \times n$. We are interested to prove that $(n++) \times m = m \times (n++)$. The left hand side evaluates to $(n++) \times m = (n \times m) + m$. The right hand side evaluates to $m \times (n++) = (m \times n) + m$. But, $n \times m = m \times n$. Therefore, the two sides are equal to each other. This closes the proof. □

We will abbreviate $n \times m$ as nm and use the usual convention that multiplication takes precedence over addition, thus for instance $ab + c$ means that $(a \times b) + c$, not $a \times (b + c)$.

Lemma 1.3.1. *Positive natural numbers have no zero divisors. Let n, m be natural numbers. Then, $n \times m = 0$ if and only if atleast one of n, m is equal to zero. In particular, if n and m are both positive, then nm is also positive.*

Proof. We use induction on n, keeping m fixed.

Let $n = 1$. Suppose $1 \times m$ is positive. $0 + + \times m$ is positive. This implies, $(0 \times m) + m$ is positive. But, $0 \times m := 0$. So, m is positive. So, nm is positive $\implies$ n, m are both positive.

We inductively assume that, if n, m are both positive, then nm is positive. We are interested to prove that if $n{++}$ and m are both positive, then $(n{++}) \times m$ is also positive. $(n{++}) \times m = (n \times m) + m$. As m is a positive natural number, the sum $(n \times m) + m$ is also positive. This closes the proof. □

Theorem: 1.3.3: Distributive law.

For any natural numbers a, b, c we have:

$$a(b+c) = ab + ac$$
$$(b+c)a = ba + ca$$

Proof. Since multiplication is commutative, we only need to show the first identity $a(b+c) = ab + ac$. We keep a and b fixed and use induction on c.

Let's prove the the base case $c = 0$.

LHS.

$$a(b+0) = ab$$

RHS.

$$ab + a \times 0 = ab$$

So, we are done.

Now, let us suppose inductively that $a(b+c) = ab + ac$. We are interested to show that $a(b+(c++)) = ab + a(c++)$.

LHS.

$$\begin{aligned} a(b+(c++)) &= a((b+c)++) \\ &= a(b+c) + a \end{aligned}$$

RHS.

$$\begin{aligned} ab + a(c++) &= ab + ac + a \\ &= a(b+c) + a \end{aligned}$$

Both sides are equal to each other and we are done. □

Theorem: 1.3.4: Multiplication is associative.

Multiplication is associative. For any natural numbers a, b, c, we have $(a \times b) \times c = a \times (b \times c)$.

Proof. We induct on c, keeping a and b fixed.

In the base case, $c = 0$. Therefore, $(a \times b) \times 0 = 0$ and $a \times (b \times 0) = a \times 0 = 0$. Both sides are equal to 0 and are hence equal to each other.

We inductively assume that $(a \times b) \times c = a \times (b \times c)$. We are interested to prove that $(a \times b) \times (c++) = a \times (b \times (c++))$.

LHS.

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

RHS.

$$\begin{aligned} a \times (b \times (c++)) &= a \times ((b \times c) + b) \\ &= a \times ((b \times c) + b) \\ &= a \times (b \times c) + (a \times b) \end{aligned}$$

But, $((a \times b) \times c) = (a \times (b \times c))$. So, both the sides are equal to each other and this closes the proof. □

Theorem: 1.3.5: Multiplication preserves order.

If a, b are natural numbers, such that $a < b$, and c is positive, then $ac < bc$.

Proof. $a < b$ means that $b = a + d$ where d is a positive natural number. Let c be a positive natural number, then $bc = (a+d)c = ac+dc$. The product of two positive natural numbers is positive. (To se this, consider two positive natural numbers r, s. There exist natural numbers p, q such that $r = p{+}{+}$, $s = q{+}{+}$. The product $rs = (p{+}{+}) \times (q{+}{+}) = ((p{+}{+}) \times q) + (p{+}{+})$. $p++$ is a positive natural number. So, rs is a positive natural number.) Consequently, $bc > ac$. □

Corollary 1.3.1. *Let a, b, c be natural numbers such that $ac = bc$ and c is non-zero. Then $a = b$.*

Just as earlier we did a virtual subtract, which will eventually let us define a genuine subtraction, this corollary provides virtual division., which will be needed to define genuine division later.

Proof. By trichotomy of order, exactly one of the three cases $a < b$, $a = b$, $a > b$ must be true. Suppose $a < b$, then $b = a + k$, where k is a positive natural number. $bc = (a + k)c = ac + kc$. kc is a positive natural number, so $ac < bc$. Therefore, this possibility can be discarded. Similarly, if $a > b$, then $a = b + l$, where l is a positive natural number. $ac = (b + l)c = bc + lc$. lc is a positive natural number, so $ac > bc$. Therefore, this also leads to a contradiction.

So, the only possibility is that $a = b$ as desired. □

With these propositions, it is easy to deduce all the familiar rules of algebra involving addition and multiplication. The more primitive notion of incrementation will begin to fall by the wayside, and we will see that it is rarely required from now on.

Theorem: 1.3.6: Euclid's division algorithm.

Let n be a natural number, and let q be a positive number. Then, there exists natural numbers m, r such that $0 \leq r < q$ and $n = mq + r$.

Proof. (1) If $n = 0$, then $0 = 0 \times q + 0$, where q is a fixed positive number. So, $P(0)$ is true.

(2) We inductively assume that $P(n)$ is true. Suppose $\exists m, r$ such that $n = mq + r$ where $0 \leq r < q$.

(3) We are interested to prove that $P(n+1)$ is true.

Since, $r < q$, we have a dichotomy - $r + 1 = q$ or $r = q$.

Case I. $r + 1 = q$.

In such case, we can write: $(n+1) = mq + r + 1 = (mq + q) = (m+1)q + 0$. $(m+1)$ and 0 are natural numbers.

Case II. $r + 1 < q$.

In such case, we can write: $(n+1) = mq + r + 1 = mq + s$.

m and s are natural numbers. Further, $0 \leq s < q$.

Hence, the claim is proved. □

Chapter 2

The Integers $\mathbf{Z}$ and rationals $\mathbf{Q}$.

2.1 The integers.

Informally, the integers are what you get by subtracting two natural numbers, for example, $3-5$ should be an integer, as should $6-2$. But, this definition is circular, because it requires the notion of subtraction, which can only adequately be defined, once we the integers are constructed. To get around this issue, we temporarily introduce a meaningless operator $--$. Later on we shall see, that $a - -b$ is in fact equal to $a - b$. It is only needed right now to avoid circularity. (These devices are similar to the scaffolding needed to construct a building; they are only temporarily essential to make sure that the building is built correctly, but once the building is completed, they are thrown away and never used again.)

Definition 2.1. An integer is an expression of the form $a{-}{-}b$, where a and b are natural numbers. Two integers are considered to be equal, $a{-}{-}b = c{-}{-}d$, if and only if $a+d = b+c$. We let $\mathbf{Z}$ denote the set of all integers.

Thus, for instance $3{-}{-}5$ is an integer, and is equal to $2{-}{-}4$, because $3+4 = 2+5$. On the other hand, $3{-}{-}5 \neq 2{-}{-}3$ because $3+3 \neq 2+5$. This notation is strange looking, and has a few deficiencies; for instance, 3 is not yet an integer, because it is not of the form $a{-}{-}b$! We will rectify these problems later.

We have to check that this is a legitimate notion of equality. We need to verify reflexivity, symmetry, transitivity and substitution axioms.

Theorem: 2.1.1: $--$is a RST relation.

The equality of integers satisfies reflexivity, symmetry and transitivity axioms. It is an RST relation.

Proof. (a) Reflexive. $a{-}{-}b = a{-}{-}b$ since $a+b = b+a$. The latter is true, because addition is commutative.

(b) Symmetric. If $a\ {-}{-}b = c\ {-}{-}d$ then $c\ {-}{-}d = a\ {-}{-}b$.

By the definition of integers, if $a{-\!-}b = c{-\!-}d$, then $a+d = c+b$. Therefore, $c+b = a+d$. So, $c{-\!-}d{=}a{-\!-}b$.

(c) Transitive. Suppose $a\ {-\!-}b = c\ {-\!-}d$ and $c\ {-\!-}d = e\ {-\!-}f$. Then, $a+d = c+b$ and $c+f = d+e$. Adding the two equations we obtain, $a+d+c+f = c+b+d+e$. By the cancellation law, $a+f = e+b$. Therefore, $a{-\!-}b{=}e{-\!-}f$.

This closes the proof. □

As for the substitution axiom, we cannot verify it at this stage, because we have not yet defined any operations on the integers. However, when we do define our basic operations on the integers, such as addition, multiplication and order, we will have to verify the substitution axiom at that time to ensure that the definition is valid. (We will only need to do this for the basic operations; more advanced operations on the integers, such as exponentiation, will be defined in terms of the basic ones, and so we do not need to re-verify the substitution axiom for the advanced operations.

Now we define two basic arithmetic operations on integers: addition and multiplication.

Definition 2.2. The sum of two integers, $(a{-\!-}b) + (c{-\!-}d)$, is defined by the formula

$$(a{-\!-}b) + (c{-\!-}d) = (a+c)\ {-\!-}(b+d)$$

The product of two integers, $(a{-\!-}b)(c{-\!-}d)$, is defined by the formula

$$(a{-\!-}b)(c{-\!-}d) = (ac+bd)\ {-\!-}(ad+bc)$$

Thus, for instance, $(3{-\!-}5) + (1{-\!-}4)$ is equal to $4\ {-\!-}9$. There is however one thing we have to check before we can accept these definitions - we have to check that if we replace one of the integers by an equal integer, that the sum or the product does not change. For instance $(3{-\!-}5)$ is equal to $(2\ {-\!-}4)$, so $(3{-\!-}5) + (1{-\!-}4)$ ought to have the same value as $(2{-\!-}4) + (1{-\!-}4)$, otherwise this would not give a consistent definition of addition. Fortunately, this is the case.

***Theorem:* 2.1.2: Addition and multiplication are well defined.**

Let a, b, a', b', c, d be natural numbers. If $(a\ {-\!-}b) = (a'\ {-\!-}b')$ then $(a\ {-\!-}b) + (c\ {-\!-}d) = (a'\ {-\!-}b') + (c\ {-\!-}d)$ and $(a\ {-\!-}b) \times (c\ {-\!-}d) = (a'\ {-\!-}b') \times (c\ {-\!-}d)$, and also $(c\ {-\!-}d) + (a\ {-\!-}b) = (c\ {-\!-}d) + (a'\ {-\!-}b')$ and $(c\ {-\!-}d) \times (a\ {-\!-}b) = (c\ {-\!-}d) \times (a'\ {-\!-}b')$. Thus, addition and multiplication are well-defined operations (equal inputs give equal outputs).

Proof. The claim is $(a\ {-\!-}b) + (c\ {-\!-}d) = (a'\ {-\!-}b') + (c\ {-\!-}d)$. To prove that, we evaluate the both sides as $(a{+}c)\ {-\!-}(b{+}d)$ and $(a'{+}c)\ {-\!-}(b'{+}d)$. To show that these two expressions are equal, has the following implication. $(a+c)\ {-\!-}(b+d) = (a'+c)\ {-\!-}(b'+d)$ means that $(a+c) + (b'+d) = (a'+c) + (b+d)$. But, by the cancellation law, $(a+b') = (a'+b)$. However, we are given that $a\ {-\!-}b = a'\ {-\!-}b'$, so the above expressions are indeed equal. Hence, the claim is true.

The next claim is $(a\ \text{——}b) \times (c\ \text{——}d) = (a'\ \text{——}b') \times (c\ \text{——}d)$. To prove that, we evaluate both the sides as $(ac+bd)\ \text{——}(ad+bc)$ and $(a'c+b'd)\ \text{——}(a'd+b'c)$. If these two expressions are equal, that is $(ac+bd)\ \text{——}(ad+bc) = (a'c+b'd)\ \text{——}(a'd+b'c)$, it implies that $(ac+bd) + (a'd+b'c) = (a'c+b'd) + (ad+bc)$. Since these are all natural numbers, we can pull out common factors, invoking distributivity property. We may write $(a+b')c+(b+a')d = (a'+b)c+(b'+a)d$. Indeed, these two expressions are equal as, $a'+b = a+b'$, which is true because $(a\ \text{——}b) = (a'\ \text{——}b')$. Hence, the claim is proved.

The third claim is $(c\ \text{——}d) + (a\ \text{——}b) = (c\ \text{——}d) + (a'\ \text{——}b')$. To prove this, we again evaluate both the sides. If the two sides are equal, then $(c+a)\ \text{——}(d+b) = (c+a')\ \text{——}(d+b')$. This implies that $c+a+d+b' = c+a'+d+b$. Using the cancellation law, we can cancel c and d. This leaves us with $a+b' = a'+b$, which we know is true. Hence, the claim is proved.

The final claim is $(c\ \text{——}d) \times (a\ \text{——}b) = (c\ \text{——}d) \times (a'\ \text{——}b')$. If the two sides are equal, we must have $(ca+db)\ \text{——}(cb+da) = (ca'+db')\ \text{——}(cb'+da')$. This is true, if $ca+db+cb'+da' = ca'+db'+cb+da$. If we factor out the terms, we are left with $c(a+b')+d(b+a') = c(a'+b)+d(b'+a)$. These expressions are indeed equal, because $a'+b = a+b'$.

This closes the proof. $\square$

The integers $n\ \text{——}0$ behave in the same way as the natural numbers n; indeed one can check that $(n\ \text{——}0) + (m\ \text{——}0) = (n+m)\ \text{——}0$. Because by definition, $(n\ \text{——}0) + (m\ \text{——}0) = (n+m)\ \text{——}0+0 = (n+m)\ \text{——}0$. And $(n\ \text{——}0) \times (m\ \text{——}0) = (nm+0)\ \text{——}0+0 = nm\ \text{——}0$. Furthermore, $(n\ \text{——}0) = (m\ \text{——}0)$ implies that $n+0 = m+0$, so $n = m$. The converse is also true. Therefore, $(n\ \text{——}0) = (m\ \text{——}0)$ if and only if $n = m$. The mathematical term for this is that there is an *isomorphism* between the natural numbers n and those integers of the form $n\ \text{——}0$. Thus, we may identify natural numbers with integers by setting $n = n\ \text{——}0$. This does not affect our definitions of addition, multiplication or equality since they are consistent with each other. For instance, the natural number 3 is now considered to be the same as the integer $3\ \text{——}0$, thus $3 = 3\ \text{——}0$. In particular, 0 is equal to $0\ \text{——}0$ and 1 is equal to $1\ \text{——}0$. Of course, if we set n equal to $n\ \text{——}0$, then it will also be equal to any other integer which is equal to $n\ \text{——}0$, for instance 3 is equal not only to $3\ \text{——}0$, but also to $4\ \text{——}1$, $5\ \text{——}2$ etc.

We can now define incrementation on the integers by defining $x{+}{+} := x+1$ for any integer x; this is of course consistent with our definition of the increment operation for natural numbers. However, this is no longer an important operation for us, as it has now been superseded by the more general notion of addition. Now, we consider other basic operations on the integers.

Definition 2.3. Negation of integers. If $(a\ \text{——}b)$ is an integer, we define the negation $-(a\ \text{——}b)$ to be the integer $(b\ \text{——}a)$. In particular, if $n{=}n\text{——}0$ is a positive natural number, we can define its negation $-n{=}0\text{——}n$.

For instance, $-(3\text{——}5){=}(5\text{——}3)$. One can check that this definition is well-defined.

Theorem: 2.1.3: Negation is well-defined.

Let a, b, a', b' be natural numbers. If $(a{-\!\!-}b){=}(a'{-\!\!-}b')$ then $-(a{-\!\!-}b){=}-(a'{-\!\!-}b')$. Thus, the negation operation on integers is well-defined.

Proof. To prove this claim, we evaluate both the sides. By definition, the left hand side equals $(b{-\!\!-}a)$, whereas the right hand side equals $(b'{-\!\!-}a')$. If the two sides are equal, it implies that $b + a' = b' + a$. We know that this equality holds since $(a{-\!\!-}b){=}(a'{-\!\!-}b')$. Hence, the claim is proved. □

We can now show that integers correspond to exactly what we expect.

Lemma 2.1.1. *Trichotomy of integers. Let x be an integer. Then, exactly one of the following three statements is true.*

(a) x is zero.

(b) x is equal to a positive natural number.

(c) x is the negation $-n$ of a positive natural number n.

Proof. We first show that atleast one of (a),(b),(c) is true. As x is an integer, by definition, $x = a{-\!\!-}b$ where a, b are natural numbers. From the trichotomy property, we know that exactly one of the three cases $a = b$, $a > b$ or $a < b$ must hold. If $a = b$, then $a + 0 = 0 + b$, so $a{-\!\!-}b = 0$. Therefore, $x = 0$. If $a > b$, then $a = b + c$ for some positive natural number c. Therefore, $a + 0 = c + b$. So, $a{-\!\!-}b = c{-\!\!-}0 = c$. Therefore, $x = c$ a positive natural number. Lastly, if $a < b$, then $b = a + d$ for some positive natural number d. So, $0 + b = a + d$. Therefore $0{-\!\!-}d = a{-\!\!-}b$. Therefore, $x = 0{-\!\!-}d$. x is the negation $-d$ of some positive natural number d.

Now, we show that no more than one of (a),(b),(c) is true. Suppose $x = 0$. Then, $x \neq 0$ is false. x is not a positive natural number. Therefore, (a) and (b) cannot be simultaneously true. If $x = 0$, then x is also not the negation of a positive natural number. This is so because, if $0\ {-\!\!-}n = 0 = 0\ {-\!\!-}0$, then $n + 0 = 0$, so $n = 0$. If the negation of a natural number is 0, then the number equals 0. So, x cannot be the negation of a positive natural number. Therefore, (a) and (c) cannot be simultaneously true. Suppose that x is equal to a positive natural number p, and it is the negation of a positive natural number q. $x = p$ and $x = 0\ {-\!\!-}q$. Then $p{-\!\!-}0 = 0\ {-\!\!-}q$, so $p + q = 0$. Therefore, $p = q = 0$. This is a contradiction. Therefore, (b) and (c) cannot hold simultaneously.

Thus, exactly one of (a), (b), (c) is true for any integer x. □

If n is a positive natural number, we call the negation of n, $0{-\!\!-}n = -n$ a *negative integer*. Thus, every integer is positive, zero, or negative, but not more than one of these at a time.

We now summarize the algebraic properties of the integers.

Theorem: 2.1.4: Laws of algebra for integers.

Let x, y, z be integers. Then, we have

$$\begin{aligned} x + y &= y + x \\ (x + y) + z &= x + (y + z) \\ x + 0 = 0 + x &= x \\ x + (-x) = (-x) + x &= 0 \\ xy &= yx \\ (xy)z &= x(yz) \\ x1 = 1x &= x \\ x(y + z) &= xy + xz \\ (y + z)x &= yx + zx \end{aligned}$$

Proof. Let $x{=}a --b$, $y{=} c--d$ and $z{=}e--f$.

(A1) **LHS.**

$$(a - -b) + (c - -d) = (a + c) - -(b + d)$$

RHS.

$$\begin{aligned} (c - -d) + (a - -b) &= (c + a) - -(d + b) \\ &= (a + c) - -(b + d) \end{aligned}$$

(A2) **LHS.**

$$\begin{aligned} ((a - -b) + (c - -d)) + (e - -f) &= ((a + c) - -(b + d)) + (e - -f) \\ &= ((a + c) + e) - -((b + d) + f) \\ &= (a + c + e) - -(b + d + f) \end{aligned}$$

RHS.

$$\begin{aligned} (a - -b) + ((c - -d) + (e - -f)) &= (a - -b) + ((c + e) - -(d + f)) \\ &= (a + (c + e)) - -(b + (d + f)) \\ &= (a + c + e) - -(b + d + f) \end{aligned}$$

(A3) **LHS.**

$$\begin{aligned} 0 + (a - -b) &= (0 - -0) + (a - -b) \\ &= (0 + a) - -(0 + b) \\ &= a - -b \end{aligned}$$

RHS.

$$\begin{aligned}(a--b)+0 &= (a--b)+(0--0)\\ &= (a+0)--(b+0)\\ &= a--b\end{aligned}$$

(A4) Firstly, we deduce that $0 = 0+0 = (n--n)$ because, $0+n = n+0$ for any natural number n.

LHS.

$$\begin{aligned}(a--b)+(-(a--b)) &= (a--b)+(b--a)\\ &= (a+b)--(b+a)\\ &= (a+b)--(a+b)\\ &= 0\end{aligned}$$

RHS.

$$\begin{aligned}(-(a--b))+(a--b) &= (b--a)+(a--b)\\ &= (b+a)--(a+b)\\ &= (a+b)--(a+b)\\ &= 0\end{aligned}$$

(M1) We claim that $xy = yx$.

LHS.

$$(a--b)(c--d) = (ac+bd)--(ad+bc)$$

RHS.

$$\begin{aligned}(c--d)(a--b) &= (ca+db)--(cb+da)\\ &= (ac+bd)--(ad+bc)\end{aligned}$$

(M2) We claim that $(xy)z = x(yz)$.

LHS.

$$\begin{aligned}((a--b)(c--d))(e--f) &= ((ac+bd)--(ad+bc))(e--f)\\ &= (ace+bde+adf+bcf)--(acf+bdf+ade+bce)\end{aligned}$$

RHS.

$$\begin{aligned}(a--b)((c--d)(e--f)) &= (a--b)((ce+df)--(cf+de))\\ &= (ace+adf+bcf+bde)--(acf+ade+bce+bdf)\end{aligned}$$

Both sides are equal to each other, and hence the claim is proved.

(M3). We claim that $1x = x1 = x$. We know, that the integers $n - -0$ behave in the same way as the natural numbers n. So, $(1 - -0)(a - -b) = (1a + 0b) - -(1b + 0a) = a - -b$. This proves that $1x = x$. Since, we have proved that $xy = yx$ for integers x, y, $1x = x1 = x$. This closes the proof.

(D1)

LHS.

$$\begin{aligned}(a - -b)((c - -d) + (e - -f)) &= (a - -b)((c + e) - -(d + f)) \\ &= (ac + ae + bd + bf) - -(ad + af + bc + be)\end{aligned}$$

RHS.

$$\begin{aligned}(a - -b)(c - -d) + (a - -b)(e - -f)) &= ((ac + bd) - -(ad + bc)) + ((ae + bf) - -(af + be)) \\ &= (ac + bd + ae + bf) - -(ad + bc + af + be)\end{aligned}$$

The two sides are equal to each other. This closes the proof.

(D2) The addition of two integers is a well-defined integer. Let x, y, z be integers. The sum $(y + z)$ is well defined. $x(y + z) = (y + z)x = xy + xz$ since $ab = ba$ for two integers a, b. $\square$

Remark 2.1. The above set of nine identities (A1)-(A4), (M1)-(M3), (D1)-(D2) have a name; they are asserting that the integers form a *commutative ring*. Note that, some of these identities were already proven for natural numbers, but this does not automatically mean that they hold for the integers because the integers are a larger set than the natural numbers. On the other hand, this proposition supercedes many of the propositions derived earlier for the natural numbers.

We now define the operation of subtraction $x - y$ of two integers.

Definition 2.4. The subtraction of two integers $x - y$ is defined by the formula

$$x - y = x + (-y)$$

We do not need to verify the substitution axiom for this operation, since we have defined subtraction in terms of two other operations on integers, namely addition and negation, and we have aready verified that those operations are well-defined.

One can easily check now that if a and b are natural numbers, then

$$a - b = a + (-b) = (a - -0) + (0 - -b) = (a + 0) - -(0 + b) = a - -b$$

and so $a - -b$ is just the same thing as $a - b$. Because of this we can now discard the $--$ notation, and use the familiar operation of subtraction instead. (As remarked before, we could not use subtraction immediately because it would be circular.)

Lemma 2.1.2. *Integers have no zero divisors. Let a and b be integers such that $ab = 0$. Then, either $a = 0$ or $b = 0$ (or both).*

Proof. **Claim 1** : Let x, y be integers. $x - y \neq 0$ if and only if $x \neq y$.

($\longrightarrow$) direction. If $x = y$, then $x + 0 = 0 + y$. Thus, $x - y = 0$.

($\longleftarrow$) direction. If $x - y = 0$, then $x - y = 0 - 0$. Thus, $x + 0 = 0 + y$. So, $x = y$.

Claim 2 : Let a, b be integers. If $ab = 0$, then atleast one of the integers a, b is zero.

Suppose $a = (p - q)$, $b = (r - s)$ and $a \neq 0$. Therefore, $ab = (p - q)(r - s) = (pr + qs) - (ps + qr)$. But, $ab = 0$. So, $(pr + qs) - (ps + qr) = 0$. From claim (1), this means $pr + qs = ps + qr$, and $p \neq q$. Using the trichotomy property of natural numbers, there arise two cases:

Case I. $p < q$

Let $q = p + k$, where k is a positive natural number. Then, $pr + (p + k)s = ps + (p + k)r$. So, $ks = kr$. By cancellation law, $s = r$. So, $b = 0$.

Case II. $p > q$

Let $p = q + l$, where l is a positive natural number. Then, $(q + l)r + qs = (q + l)s + qr$. So, $lr = ls$. By cancellation law, $r = s$. So, $b = 0$. This closes the proof. □

Corollary 2.1.1. *Cancellation law for integers. If a, b, c are integers such that $ac = bc$ and c is non-zero, then $a = b$.*

Proof. As c is an integer, by the trichotomy property, we have three cases : (a) c is the negation of a positive natural number (b) $c = 0$ (c) c is equal to a positive natural number.

Let us express the integers a, b, c in the form -

$$\begin{aligned} a &= m - n \\ b &= p - q \\ c &= r - s \end{aligned}$$

Since $c \neq 0$, we have the below two cases.

Case I. c is the negation of a positive natural number.

If $c = -(r - s)$, then $s - r = k$, where k is a positive natural number. Therefore, $s = r + k$.

$$
\begin{aligned}
ac &= bc \\
(m-n)(r-s) &= (p-q)(r-s) \\
(mr+ns)-(ms+nr) &= (pr+qs)-(ps+qr) \\
(mr+ns)+(ps+qr) &= (pr+qs)+(ms+nr) \\
(mr+n(r+k))+(p(r+k)+qr) &= (pr+q(r+k))+(m(r+k)+nr) \\
mr+nr+nk+pr+pk+qr &= pr+qr+qk+mr+mk+nr \\
nk+pk &= qk+mk \\
(n+p)k &= (q+m)k \\
n+p &= q+m \\
p+n &= m+q \\
p-q &= m-n \\
a &= b
\end{aligned}
$$

Case II. c is equal to a positive natural number.

If $c = r - s = k$, Therefore, $r = s + k$. This can be proved similar to case (I). □

We now extend the order relation, which was defined on the natural numbers, to integers by repeating the definition verbatim.

Definition 2.5. Ordering of integers. Let n and m be integers. We say that, n is greater than or equal to m, and write $n \geq m$ or $m \leq n$, iff we have $n = m + a$ for some natural number a. We say that n is strictly greater than m, and write $n > m$ or $m < n$, iff $n \geq m$ and $n \neq m$.

Using the laws of algebra of integers, it is not hard to show the following properties of order:

***Theorem:* 2.1.5: Properties of the order relation on integers.**

Let a, b, c be integers.
(a) $a > b$ if and only if $a - b$ is a positive natural number.
(b) (Addition preserves order) If $a > b$, then $a + c > b + c$.
(c) (Positive multiplication preserves order) If $a > b$ and c is positive, then $ac > bc$.
(d) (Negation reverses order) If $a < b$, then $-a < -b$.
(e) (Order is transitive) If $a > b$ and $b > c$, then $a > c$.
(f) (Order trichotomy) Exactly one of the statements $a < b, a = b, a > b$ is true.

Proof. (a) **Claim:** $a > b$ if and only if $a - b$ is a positive natural number.

($\longrightarrow$) direction.

$a > b$. Therefore, $a = b + k$ for some natural number k and $a \neq b$. Simplifying, $a + (-b) = b + (-b) + k \implies a - b = k$. $(a - b)$ is a natural number. But, $a \neq b \implies a - b \neq 0$. So, $a - b$ is a positive natural number.

($\longleftarrow$) direction.

Assume that $a-b$ is equal to a positive natural number k. Then, $a-b=k$ such that $k \neq 0$. This means that $a=b+k$ such that $k \neq 0$. $k \neq 0 \implies (a-b) \neq 0 \implies a \neq b$. We conclude that, $a=b+k$ such that $a \neq b$. Thus, $a>b$.

(b) **Claim:** Addition preserves order. If $a>b$, then $a+c>b+c$.

$a>b \implies a=b+k$ where k is a positive natural number by property (a). Therefore, $a+c=b+k+c \implies (a+c)=(b+c)+k$ where k is a positive natural number. Hence, $a+c>b+c$.

(c) **Claim:** Positive multiplication preserves order. If $a>b$ and c is positive, then $ac>bc$.

$a>b \implies a=b+k$ where k is a positive natural number by property (a). Therefore, $ac=(b+k)c \implies ac=bc+kc$. c is a positive integer, so it is equal to a positive natural number. The product kc of two positive natural numbers k,c is positive. Therefore, $ac>bc$.

(If n,m are positive natural numbers, the product nm is positive. We induct on n keeping m fixed. $1 \times m := m$. m is a positive natural number, so $1m$ is positive. This proves the base case. We inductively assume that, nm is positive. We are interested to show that $(n{+}{+}) \times m$ is positive. $(n{+}{+}) \times m = (n \times m) + m$. Addition of two positive natural numbers is positive. This closes the proof.)

(d) **Claim:** Negation reverses order. If $a>b$, then $-a<-b$.

$a>b \implies a=b+k$ where k is a positive natural number by property (a). Multiplying throughout by -1, $(-1)a=(-1)(b+k)$. But it can be shown that, $(-1)x=-x$. Clearly, $(-1)x+x=(-1)x+(1)x=(-1+1)x=0x=0$. Thus, $(-1)x$ is the negation of x. Hence, $(-1)x=-x$. Thus, $(-1)a=(-1)b+(-1)k \implies -a=-b-k \implies (-b)=(-a)+k$, such that k is a positive natural number. Thus, $(-a)<(-b)$.

(e) **Claim:** Order is transitive. If $a>b$ and $b>c$, then $a>c$.

$a>b$ and $b>c$. So, $a=b+k$ and $b=c+l$ where k,l are positive natural numbers. Adding these two equations, we have : $a+b=b+c+k+l$. So, $a=c+(k+l)$. $(k+l)$ is a positive natural number. So, $a>c$.

(f) **Claim:** Order trichotomy. Exactly one of the statements $a<b, a=b, a>b$ is true.

We look at the difference between the integers $a-b$. We know, that exactly one of three cases holds true (i) an integer is equal to zero (ii) an integer is equal to a positive natural number (c) an integer is the negation of a positive natural number.

If $a-b=0$, then $a=b$.

If $a-b=k$, where k is a positive natural number, $a=b+k$, so $a>b$.

If $a-b=-c$, where c is a positive natural number, $b=a+c$, so $a<b$.

This closes the proof. □

Problem 2.1. Verify that the definition of equality of integers is both reflexive and symmetric.

Exercise 4.1.1 - page 81, Analysis I, Tao

Proof. Let a, b, c, d be natural numbers. We define integers $x = a - b$ and $y = c - d$.

The equality relation $=_{\mathbf{Z}}$ on the integers $\mathbf{Z}$ is defined to be the set of all ordered pairs (x, y) given by

$R := \{(x, y)|(a + d) = (b + c);$ such that $x = a - b, y = c - d\}$

(1) **Reflexive.** Clearly, $a+b = a+b$ for natural numbers a, b, so $a-b = a-b$. This implies $x = x$ for all integers x.

(2) **Symmetric.** Moreover, if $x = y$, then $a + d = b + c$, it implies $b + c = a + d$, so $y = x$. □

Problem 2.2. Show that $(-1) \times a = -a$ for every integer a.

Exercise 4.1.3 - page 81, Analysis I, Tao

Proof. Consider the fact that $(-1)a + a = (-1)a + 1a = ((-1) + 1)a = 0a$. Further, let us write the integer a in the form $a = n - m$, where m, n are natural numbers. Multiplying by 0, yields $(0 - 0)(n - m) = (0m + n0) - (0n + 0m) = 0$. So, $(-1)a + a = 0$. $(-1) \times a$ is the negation of a. $(-1) \times a = -a$. □

2.2 The rationals.

We have now constructed the integers, with the operations of addition, subtractions, multiplication and order and verified all the expected algebraic and order theoretic properties. Now, we will use a similar construction to build rationals, adding division to our mix of operations.

In analogy with the integers, we create a new meaningless symbol // which will eventually be superceded by division and define -

Definition 2.6. A rational number is an expression of the form $a//b$, where a and b are integers and b is non-zero; $a//0$ is not considered to be a rational number. Two rational numbers are considered to be equal, $a//b = c//d$, if and only if $ad = bc$. The set of all rational numbers is denoted $\mathbf{Q}$.

This is a valid definition of equality. We can prove this fact as follows.

***Theorem:* 2.2.1**

Equality of rational numbers is well-defined.

Proof. We are interested to prove that, $=_{\mathbf{Q}}$ is a legitimate notion of equality.

(1) $=_{\mathbf{Q}}$ is reflexive.

If $(a, b) =_{\mathbf{Q}} (a, b)$, this implies that $ab =_{\mathbf{Z}} ba$. The last equality holds, because multiplication in integers is commutative.

(2) $=_{\mathbf{Q}}$ is symmetric.

$(a, b) =_{\mathbf{Q}} (c, d)$
$\implies ad =_{\mathbf{Z}} bc$
$\implies bc =_{\mathbf{Z}} ad$
$\implies cb =_{\mathbf{Z}} da$
$\implies (c, d) =_{\mathbf{Z}} (a, b)$

(3) $=_{\mathbf{Q}}$ is transitive.

Suppose $(a, b) =_{\mathbf{Q}} (c, d)$ and $(c, d) =_{\mathbf{Q}} (e, f)$.

$\implies ad =_{\mathbf{Z}} bc, cf =_{\mathbf{Z}} de$
$\implies adcf =_{\mathbf{Z}} bcde$
$\implies acdf =_{\mathbf{Z}} bcde$
$\implies af =_{\mathbf{Z}} be$
$\implies (a, b) =_{\mathbf{Q}} (e, f)$ □

Definition 2.7. If $a//b$ and $c//d$ are rational numbers, we define their sum

$$a//b + c//d := (ad + bc)//bd$$

their product

$$(a//b) \times (c//d) := ac//bd$$

and the negation

$$-(a//b) := (-a)//b$$

***Theorem:* 2.2.2: Well-definedness of $+, -, \times$ operations**

The sum, product, and negation operations on the rational numbers are well-defined, in the sense that if one replaces $a//b$ with another rational number $a'//b'$ which is equal to $a//b$, then the output of the above operations remains unchanged, and similarly for $c//d$.

Proof. Let $(a, b) =_{\mathbf{Q}} (a', b')$. Then, $ab' =_{\mathbf{Z}} a'b$.

(1) Addition is well-defined.

$(a, b) + (c, d) = (ad + bc, bd)$
$(a', b') + (c, d) = (a'd + b'c, b'd)$

Observe that, $(ad + bc)b'd = ab'd^2 + bb'cd$ and $(a'd + b'c)bd = a'bd^2 + bb'cd$. The right hand sides of the two equations are equal to each other. So, $(ad + bc)b'd =_{\mathbf{Z}} (a'd + b'c)bd$. Thus, $(ad + bc)//bd = (a'd + b'c)//b'd$. Hence, $a//b + c//d = a'//b' + c//d$. Equal inputs give equal outputs.

(2) Multiplication is well-defined.

$(a, b) \times (c, d) = (ac, bd)$
$(a', b') \times (c, d) = (a'c, b'd)$

Now, $(ac)(b'd) = (ab')(cd) = (a'b)(cd) = (a'c)(bd)$. Therefore, $ac//bd = a'c//b'd$. Consequently, $(a//b) \times (c//d) = (a'//b') \times (c//d)$. Equal inputs give equal outputs.

(3) Negation is well-defined.

$-(a//b) := (-a)//b$
$-(a'//b') := (-a')//b'$

Note that, $(-a)b' = -(ab') = -(a'b) = (-a')b$. So, $(-a)//b = (-a')//b'$. Consequently, $-(a//b) = -(a'//b')$. Equal inputs yield equal outputs. □

We note that the rational numbers $a//1$ behave in a manner identical to the integers a:

$(a//1) + (b//1) = (a + b)//1$
$(a//1) \times (b//1) = ab//1$
$-(a//1) = (-a)//1$

Also, $a//1$ and $b//1$ are only equal when a and b are equal. Because of this, we will identify a with $a//1$ for each integer $a := a//1$; the above identities then guarantee that the arithmetic of integers is consistent with the arithmetic of the rational numbers. Thus, just as we embedded the natural numbers inside the integers, we embed the integers inside the rational numbers. In particular, all natural numbers are rational numbers, for instance 0 is equal to $0//1$ and 1 is equal to $1//1$.

Observe that a rational number $a//b$ is equal to $0 = 0//1$ if and only if $a \times 1 = b \times 0$ i.e. if the numerator a is equal to zero. Thus, if a and b are non-zero then so is $a//b$.

We now define a new operation on the rationals : the reciprocal. If $x = a//b$ is a non-zero rational (so that $a, b \neq 0$), then we define the reciprocal x^{-1} of x to be the rational number $x^{-1} := b//a$. It is easy to check that this operation is consistent with our notion of equality : if two rational numbers $a//b, a'//b'$ are equal, consider their reciprocals $b//a, b'//a'$; we see that $ba' = a'b = ab' = b'a$. Consequently, $b//a = b'//a'$. So, if two rational numbers are equal, their reciprocals are also equal. We however, leave the reciprocal of 0 undefined.

We now summarize the algebraic properties of rational numbers.

***Theorem:* 2.2.3: Laws of algebra for rationals.**

et x, y, z be rationals. Then the following laws of algebra hold:

$$\begin{aligned} \mathrm{x + y} &= \mathrm{y + x} \\ \mathrm{(x+y)+z} &= \mathrm{x + (y+z)} \\ \mathrm{x + 0 = 0 + x} &= \mathrm{x} \\ \mathrm{x + (-x) = (-x) + x} &= 0 \\ \mathrm{xy} &= \mathrm{yx} \\ \mathrm{(xy)z = x(yz)} & \\ \mathrm{x1 = 1x} &= \mathrm{x} \\ \mathrm{x(y+z)} &= \mathrm{xy + xz} \\ \mathrm{(y + z)x} &= \mathrm{yx + zx} \end{aligned}$$

If x is non-zero, we also have:

$$xx^{-1} = x^{-1}x = 1$$

Remark 2.2. The above set of ten identities have a name; they are asserting that the rationals **Q** form a *field*. This is better than being a commutative ring because of the tenth identity. $xx^{-1} = x^{-1}x = 1$.

Proof. Let $x = a//b, y = c//d, z = e//f$.

(A1) **Claim.** $x + y = y + x$

LHS.

$$a//b + c//d = (ad + bc)//bd$$

RHS.

$$\begin{aligned} c//d + a//b &= (cb + da)//db \\ &= (ad + bc)//(bd) \end{aligned}$$

Consequently, addition of rational numbers is commutative.

(A2) **Claim.** $(x + y) + z = x + (y + z)$

LHS.

$$\begin{aligned} (a//b + c//d) + e//f &= (ad + bc)//bd + e//f \\ &= (adf + bcf + bde)//bdf \end{aligned}$$

RHS.

$$\begin{aligned} a//b + (c//d + e//f) &= a//b + (cf + de)//df \\ &= (adf + bcf + bde)//bdf \end{aligned}$$

Both equations are equal to $(adf + bcf + bde)//bdf$ and are hence equal to each other. Consequently, addition of rational numbers is associative.

(A3) **Claim.** $x + 0 = 0 + x = x$

LHS.

$$\begin{aligned} (a//b + 0//1) &= (a \times 1) + (b \times 0)//(b \times 1) \\ &= a//b \end{aligned}$$

RHS.

$$\begin{aligned} (0//1 + a//b) &= (0b + 1a)//(1b) \\ &= a//b \end{aligned}$$

Both expressions are equal to $a//b$ and are hence equal to each other. Hence, the claim is proved.

(A4) **Claim.** $x + (-x) = (-x) + x = 0$

LHS.

$$\begin{aligned} (a//b + (-(a//b))) &= a//b + (-a)//b \\ &= (ab + b(-a))//(b \cdot b) \\ &= (ab + (-a)b)//(b \cdot b) \\ &= (ab - ab)//(b \cdot b) \\ &= 0//(b \cdot b) \\ &= 0 \end{aligned}$$

RHS.

$$\begin{aligned} (-(a//b) + a//b) &= (-a)//b + a//b \\ &= ((-a)b + ba)//(b \cdot b) \\ &= (-ab + ab)//(b \cdot b) \\ &= 0//(b \cdot b) \\ &= 0 \end{aligned}$$

There exists a negative element for every element $x \in \mathbf{Q}$.

(M1) **Claim.** $xy = yx$

LHS.

$$(a//b)(c//d) = ac//bd$$

RHS.

$$\begin{aligned}(c//d)(a//b) &= ca//db\\ &= ac//bd\end{aligned}$$

Multiplication of rational numbers is commutative.

(M2) **Claim.** $(xy)z = x(yz)$

LHS.

$$\begin{aligned}((a//b)(c//d))(e//f) &= (ac//bd)(e//f)\\ &= ace//bdf\end{aligned}$$

RHS.

$$\begin{aligned}(a//b)((c//d)(e//f)) &= (a//b)(ce//df)\\ &= ace//bdf\end{aligned}$$

Multiplication of rational numbers is associative.

(M3) **Claim.** $x1 = 1x = x$

LHS.

$$\begin{aligned}(a//b)(1//1) &= a1//b1\\ &= a//b\end{aligned}$$

RHS.

$$\begin{aligned}(1//1)(a//b) &= 1a//1b\\ &= a//b\end{aligned}$$

(D1) **Claim.** $x(y+z) = xy + xz$

LHS.

$$\begin{aligned}(a//b)(c//d + e//f) &= (a//b)(cf + de)//df\\ &= (acf + ade)//bdf\end{aligned}$$

RHS.

$$\begin{aligned}(a//b)(c//d) + (a//b)(e//f) &= ac//bd + ae//bf\\ &= (abcf + abde)//bbdf\end{aligned}$$

Note that, $(acf + ade)bbdf = (abcf + abde)bdf$. So, the two sides are equal to each other.

(D2) **Claim.** $(y + z)x = yx + zx$

By commutativity of the multiplication of rational numbers, $(y + z)x = x(y + z)$. Hence, the claim is proved.

Existence of reciprocals.

LHS.

$$\begin{aligned} xx^{-1} &= (a//b)(b//a) \\ &= ab//ba \\ &= ab//ab \end{aligned}$$

$ab//ab = 1//1 = 1$ since by cross-multiplying, we see $ab(1) = 1(ab)$.

RHS.

$$\begin{aligned} x^{-1}x &= (b//a)(a//b) \\ &= ba//ab \\ &= ab//ab \\ &= 1 \end{aligned}$$

This completes the proof. □

We can now define the **quotient** x/y of two rational numbers x and y, provided that y is non-zero by the formula

$$x/y := x \times y^{-1}$$

Thus, for instance, $(3//4)/(5//6) = (3//4)(6//5) = (18//20) = (9//10)$.

Using this formula, it is easy to see that $a/b = a//b$ for every integer a and every non-zero integer b. Thus, we can now discard the // notation and use the more customary a/b instead of $a//b$.

In a similar spirit, we define subtraction on the rational numbers by the formula:

$x - y := x + (-y)$,

just as we did with integers.

The **field** axioms allow us to use all the normal rules of algebra; we will now proceed to do so without further comment.

Definition 2.8. A rational number x is said to be positive, if and only if we have $x = a/b$ for some positive integers a and b. It is said to be negative if and only if $x = -y$ for some positive rational number y.

Thus, for instance, every positive integer is a positive rational number and every negative integer is a negative rational number, and so our new definition is consistent with our old one.

Lemma 2.2.1. *Trichotomy of rationals. Let x be a rational number. Then exactly, one of the following three statements is true: (a)x is equal to 0. (b) x is a positive rational number. (c) x is a negative rational number.*

Proof. (1) We first prove that at most one of the statements is true. Let $x = a/b$. Suppose x is equal to 0 and x is a positive rational number. Then, $a = 0$. But, x is a positive rational number, so a and b are positive integers. This is a contradiction. Similarly, x equal to 0 and x being a negative rational number cannot be true simultaneously. Lastly, if x is a positive rational number then $x = a/b$ where a and b are positive integers. If x is a negative rational number, then $x = -y = -c/d$ where y is a positive rational number. $a/b = -c/d$. So, $ad = -bc$. But, a, b, c, d are positive integers. This is a contradiction. x equal to a positive rational number and a negative rational number cannot be true simultaneously.

(2) Next, we can show that atleast one the statements is true.

$x = a/b$. Assume that $b > 0$. By the trichotomy of integers, either $a = 0$, $a > 0$ or $a < 0$. Hence, the rational a/b is either equal to 0, a positive rational number or a negative rational number. Atleast one of these is true. □

Definition 2.9. Ordering of the rationals. Let x and y be rational numbers. We say that $x > y$ if and only if $x - y$ is a positive rational number, and $x < y$ if and only if $x - y$ is a negative rational number. We write $x \geq y$ if and only if $x > y$ or $x = y$, and similarly define $x \leq y$.

***Theorem:* 2.2.4: Basic properties of order on rationals.**

Let x, y, z be rational numbers. Then the following numbers hold:
(a) (Order trichotomy). Exactly one of the three statements $x = y, x < y, x > y$ is true.
(b) (Order is anti-symmetric). One has $x < y$ if and only if $y > x$.
(c) (Order is transitive). If $x < y$ and $y < z$, then $x < z$
(d) (Addition preserves order). If $x < y$, then $(x + z) < (y + z)$.
(e) (Positive multiplication preserves order). If $x < y$ and z is positive, then $xz < yz$.

Proof. (a) **Reflexive**. Consider the rational number $x - y$. By the lemma (2.2.1) on trichotomy of rational numbers, exactly one of the statements is true:

(i) $x - y$ is equal to zero.

(ii) $x - y$ is equal to a positive rational number.

(iii) $x - y$ is equal to a negative rational number.

By definition (2.9), exactly one of the statements $x = y$, $x < y$ or $x > y$ must be true.

(b) **Anti-Symmetric.**

Let me rigorously establish that $-(x - y) = y - x$. Let $z = (x - y)$.

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

(i) ($\longrightarrow$) direction. From definition (2.9), it follows that, if $x < y$, then $x - y$ is a negative rational number. From definition (2.8), we suppose $x - y$ be of the form $(x - y) = -(a/b)$, a/b is a positive rational number. Taking negation on both sides, $-(x - y) = a/b$. So, $y - x$ is a positive rational number. By definition (2.9), $y > x$.

(ii) ($\longleftarrow$) direction. From definition (2.9), it follows that, if $y > x$, then $y - x$ is a positive rational number. Let $y - x = (a/b)$. Where a and b are positive integers. Taking negation on both sides, $-(y - x) = -(a/b)$ where a/b is a positive rational number. So, $x - y$ is a negative rational number. By definition (2.9), $x < y$.

(c) Transitive.

Claim. If $x < y$ and $y < z$, then $x < z$. By definition (2.9), $x - y = -(m/n)$ and $y-z = -(p/q)$ where m/n and p/q are positive rational numbers. $(x-y)+(y-z) = -(m/n)+(-(p/q)) = (-m)/n + (-p)/q = ((-mq) + (-pn))/nq = -(mq + pn)/nq$. $(mq + pn)/nq$ is a positive rational number, for positive integers m, n, p, q. Consequently, $x - z$ is a negative rational number. Hence, $x < z$.

(d) Addition preserves order.

Claim. If $x < y$, then $(x + z) < (y + z)$.

By definition (2.8), if $x < y$, then $x - y = -(m/n)$ where m/n is a positive rational number. $x + z - y + (-z) = -(m/n)$. Therefore, $(x + z) - (y + z) = -(m/n)$. Hence, $x + z < y + z$.

(e) Positive multiplication preserves order.

Claim. If $x < y$, and z is positive, then $xz < yz$.

$x < y$, so $(x - y) = -(m/n)$ where m/n is a positive rational number. Since z is a positive rational number, Let $z = p/q$ where p, q are positive integers. $(x - y)z = -(m/n)(p/q) = (-m)p/qn = -(mp)/qn = -(mp/qn)$. Both mp and qn are positive integers. Hence, $xz - yz$ is a negative rational number. Consequently, $xz < yz$. $\square$

Remark 2.3. The above five properties of order on the rational numbers, combined with the field axioms for $\mathbf{Q}$ have a name: they assert that the rational numbers $\mathbf{Q}$ form an ordered field. It is important to keep in mind, that the proposition only works when z is positive.

Problem 2.3. Show that if x, y, z are rational numbers, such that $x < y$ and z is negative, then $xz > yz$.

Proof. Since $x < y$, by definition, $x - y = -(m/n)$ where m/n is a positive rational number. Also, $z = -(p/q)$ since z is a negative rational number. $(x - y)z = -(m/n) \times -(p/q) = (-m)(-p)/nq$.

For integers x, y, we can easily show that $(-x)(-y) = xy$. To see this, consider the sum $-(xy) + (-x)(-y) = (-x)y + (-x)(-y) = (-x)(y + (-y)) = (-x)0 = 0$. Thus, $(xy) + (-xy) + (-x)(-y) = 0 + xy$. So, $(-x)(-y) = xy$.

Consequently, $(-m)(-p)/nq = mp/nq$ which is a positive rational number. So, $xz - yz$ is a positive rational number. Therefore, $xz > yz$. $\square$

2.3 Absolute value and exponentiation.

We have already introduced the four basic arithmetic operations of addition, subtraction, multiplication and division on the rationals. We also have a notion of order $<$, and have organised the rationals on a number line, as negative rationals, zero and the positive rationals. In short, we have shown, that the rationals $\mathbf{Q}$ form an *ordered field*.

One can now use these basic operations to construct more operations. There are many such operations we can construct, but we shall just introduce two particularly useful ones: absolute value and exponentiation.

Definition 2.10. Absolute value. If x is a rational number, the absolute value $|x|$ of x is defined as follows. If x is positive, then $|x| := x$. If x is negative, then $|x| := -x$. If x is zero, then $|x| := 0$.

Definition 2.11. Distance. Let x and y be rational numbers. The quantity $|x - y|$ is called the distance between x and y and is sometimes denoted $d(x, y)$, thus $d(x, y) := |x - y|$. For instance, $d(3, 5) = 2$.

***Theorem:* 2.3.1: Basic properties of absolute value and distance.**

Let x, y, z be rational numbers.
(a) (Non-degeneracy of absolute value). We have $|x| \geq 0$. Also, $|x| = 0$ if and only if x is 0.
(b) (Triangle inequality for absolute value). We have $|x + y| \leq |x| + |y|$.
(c) We have the inequalities $-y \leq x \leq y$ if and only if $y \geq |x|$. In particular, we have $-|x| \leq x \leq |x|$.
(d) (Multiplicativity of absolute value). We have $|xy| = |x||y|$. In particular, $|-x| = |x|$.
(e) (Non-degenaracy of distance). We have $d(x, y) \geq 0$. Also, $d(x, y) = 0$ if and only if $x = y$.
(f) (Symmetry of distance). $d(x, y) = d(y, x)$.
(g) (Triangle inequality for distance). $d(x, z) \leq d(x, y) + d(y, z)$

Proof. (a) Claim. $x \geq 0$.

Case I. $x \geq 0$. $|x| = x \geq 0$. Case II. $x < 0$. $|x| = -x > 0$. In both cases, $|x| \geq 0$.

Moreover, if $|x| = 0$ then $|x| = x = 0$. And if $x = 0$, then taking absolute value on both sides, $|x| = 0$.

(b) Triangle inequality. $|x + y| \leq |x| + |y|$.

Case I. $x \geq 0$ and $y \geq 0$.

$x + y \geq 0$. $|x + y| = x + y = |x| + |y|$

Case II. $x \leq 0$ and $y \leq 0$.

$x + y \leq 0$. $|x + y| = -(x + y) = (-x) + (-y) = |x| + |y|$.

Case III. $x \leq 0$ and $y \geq 0$.

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

Case IV. $x \geq 0$ and $y \leq 0$.

$x + y = |x| + (-|y|) \leq |x| + |y|$. So, $|x + y| \leq ||x| + |y|| = |x| + |y|$.

In all the cases, $|x + y| \leq |x| + |y|$.

(c) Claim. $-y \leq x \leq y$ if and only if $y \geq |x|$. In particular, $-|x| \leq x \leq |x|$.

($\longrightarrow$) direction.

Case I. $x \geq 0$

$y \geq x = |x|$.

Case II. $x \leq 0$.

$-y \leq x$
$\implies -y \leq -|x|$
$\implies y \geq |x|$

($\longleftarrow$) direction.

Case I. $x \geq 0$.

$y \geq |x| = x$

Further, $y \geq |x| \geq -x$. So, $-y \leq x$. Consequently, $-y \leq x \leq y$.

Case II. $x \leq 0$.

$y \geq |x| = -x$, so $-y \leq x$. Moreover, $y \geq |x| \geq x$. Consequently, $-y \leq x \leq y$.

In all cases, $-y \leq x \leq y$.

In particular, we know that $|x| = \max\{x, -x\} \geq x$. And, $|x| \geq -x$, so $-|x| \leq x$. So, $-|x| \leq x \leq |x|$.

(d) Multiplicativity of absolute value.

Claim. $|xy| = |x||y|$

Case I. $x \geq 0$ and $y \geq 0$.

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

Case II. $x \leq 0$ and $y \leq 0$.

$|xy| = xy = (-|x|)(-|y|) = |x||y|$.

Case III. $x \leq 0$ and $y \geq 0$.

$|xy| = -xy = -(-|x|)(|y|) = |x||y|$.

Case IV. $x \geq 0$ and $y \leq 0$.

$|xy| = -xy = -(|x|)(-|y|) = |x||y|$.

In all cases, $|xy| = |x||y|$.

In particular, $|-x| = |-1||x| = 1|x| = |x|$.

(e) Non-degeneracy of distance. **Claim** We have $d(x, y) \geq 0$. Also, $d(x, y) = 0$ if and only if $x = y$.

Let $z = x - y$. $d(x, y) = |x - y| = |z| \geq 0$. Moreover, if $|z| = 0$, then $z = 0$, so $x - y = 0 \implies x = y$.

(f) Symmetry of distance.

Let $z = x - y$. $d(x, y) = |x - y| = |z| = |-z| = |y - x| = d(y, x)$.

(g) Triangle inequality for distance.

$|x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z| = d(x, y) + d(y, z).$

This completes the proof. □

Absolute value is useful for measuring how **close** two numbers are.

Definition 2.12. ϵ-closeness. Let $\epsilon > 0$ be a rational number, and let x, y be rational numbers. We say that y is ϵ-close x if and only if we have $d(y, x) \leq \epsilon$.

Remark 2.4. This definition is not standard in mathematics textbooks; we will use it as scaffolding to construct the more important notion of limits (and of Cauchy sequences) later on, once we have those more advanced notions, we will discard the notion of $\epsilon-$closeness.

Example 2.3.1. The numbers 0.99 and 1.01 are $0.1-$close, but they are not 0.01-close, because $d(1.01, 0.99) = |1.01 - 0.99| = 0.02$, which is larger than 0.01. The numbers 2 and 2 are ϵ-close for every positive ϵ.

Some basic properties of ϵ-closeness are the following.

***Theorem:* 2.3.2**

Let x, y, z, w be rational numbers.

(a) If $x = y$, then x is ϵ-close to y for every $\epsilon > 0$. Conversely, if x is ϵ-close to y for every $\epsilon > 0$, then we have $x = y$.

(b) Let $\epsilon > 0$. If x is ϵ-close to y, then y is ϵ-close to x.

(c) Let $\epsilon, \delta > 0$. If x is ϵ-close to y and y is δ-close to z, then x and z are $(\epsilon + \delta)$-close.

(d) Let $\epsilon, \delta > 0$. If x and y are ϵ-close, and z and w are δ-close, then $x + z$ and $y + w$ are $(\epsilon + \delta)$-close, and $x - z$ and $y - w$ are also $(\epsilon + \delta)$-close.

(e) Let $\epsilon > 0$. If x and y are ϵ-close, then they are also ϵ'-close for every $\epsilon' > \epsilon$.

(f) Let $\epsilon > 0$. If y and z are both ϵ-close to x, and w is between y and z (i.e. $y \leq w \leq z$ or $z \leq w \leq y$), then w is also ϵ-close x.

(g) Let $\epsilon > 0$. If x and y are ϵ-close, and z is non-zero, then xz and yz are $\epsilon|z|$-close.

(h) Let $\epsilon, \delta > 0$. If x and y are ϵ-close, and z and w are δ-close, then xz and yw are $(\epsilon|z| + \delta|x| + \epsilon\delta)$-close.

Proof. (a)($\longrightarrow$) direction.

If $x = y$, then $d(x, y) = |x - y| = 0 < \epsilon$, $\forall \epsilon > 0$.

($\longleftarrow$) direction.

If $d(x, y) \leq \epsilon$, $\forall \epsilon > 0$, it implies $d(x, y) = 0 \implies |x - y| = 0 \implies (x = y)$.

(b) Clearly $d(x, y) = d(y, x)$. So, if $d(x, y) \leq \epsilon \implies d(y, x) \leq \epsilon$ for some $\epsilon > 0$.

(c) $d(x, y) \leq \epsilon$ for some $\epsilon > 0$.

$d(y, z) \leq \delta$ for some $\delta > 0$.

$d(x, y) + d(y, z) \leq (\epsilon + \delta)$ since addition preserves order. Therefore, $d(x, z) \leq d(x, y) + d(y, z) \leq (\epsilon + \delta)$.

(d) $|x - y| \leq \epsilon$. $|z - w| \leq \delta$. $|x - y| + |z - w| \leq (\epsilon + \delta)$. Thus, $|x + z - (y + w)| = |x - y + z - w| \leq |x - y| + |z - w| \leq (\epsilon + \delta)$.

(e) $d(x, y) \leq \epsilon < \epsilon'$, $\forall \epsilon' > \epsilon$.

(f) $d(y, x) \leq \epsilon$. So, $|y - x| \leq \epsilon$. $d(z, x) \leq \epsilon$. So, $|z - x| \leq \epsilon$. So, we have the inequalities: $-\epsilon \leq y - x \leq \epsilon$ and $-\epsilon \leq z - x \leq \epsilon$.

Case I. Suppose $y \leq w \leq z$. Thus, $y - x \leq w - x \leq z - x$. Hence, $-\epsilon \leq y - x \leq w - x \leq z - x \leq \epsilon$.

Case II. Suppose $z \leq w \leq y$. Thus, $z - x \leq w - x \leq y - x$. Hence, $-\epsilon \leq z - x \leq w - x \leq y - x \leq \epsilon$.

(g) $|x - y| \leq \epsilon$. Then, $|z||y - x| \leq \epsilon|z|$. So, $|yz - xz| \leq \epsilon|z|$.

(h) Let $a := y - x, b := w - z$. Then, $y = x + a, w = z + b$ and $|a| \leq \epsilon$, $|b| \leq \delta$. $yw = (x+a)(z+b) = xz+bx+az+ab$. So, $|yw-xz| \leq |b|x+a|z|+|a||b| = \delta|x|+\epsilon|z|+\delta\epsilon$. $\square$

Remark 2.5. One should compare statements (a)-(c) of this theorem with the reflexive, symmetric and transitive axioms of equality. It is often useful to think of the notion of ϵ-closeness as an approximate substitute for that of equality in analysis.

Definition 2.13. Exponentiation to a natural number. Let x be a rational number. To raise x to the power 0, we define $x^0 := 1$; in particular we define $0^0 := 1$. Now, suppose inductively that x^n has been defined for some natural number n, then we define $x^{n+1} := x^n \times x$.

***Theorem:* 2.3.3: Properties of exponentiation.**

Let x, y be rational numbers, and let n, m be natural numbers.

(a) We have $x^n x^m = x^{n+m}$, $(x^n)^m = x^{nm}$, $(xy)^n = x^n y^n$.

(b) Suppose $n > 0$. Then, we have $x^n = 0$ if and only if $x = 0$.

(c) If $x \geq y \geq 0$, then $x^n \geq y^n \geq 0$. If $x > y \geq 0$ and $n > 0$, then $x^n > y^n \geq 0$.
(d) $|x^n| = |x|^n$.

Proof. (a1) $x^n x^m = x^{n+m}$

We induct on n keeping m fixed.

Suppose $n = 0$. Then, $x^0 \times x^m = 1 \times x^m = x^m$. On the right hand side, $x^{0+m} = x^m$. The two sides are equal to x^m and are hence equal to each other. This proves the base case.

We inductively assume that $x^n x^m = x^{n+m}$.

We are interested to prove that $x^{n+1}x^m = x^{n+1+m}$. $x^{n+1}x^m = x \cdot x^n \cdot x^m = x \cdot x^{n+m}$. On the right hand side, the sum of two natural numbers, n, m, $(n+m)$ is a natural number. By definition of exponentiation, $x^{(n+m)+1} := x^{n+m} \cdot x$. So, the two sides are equal to $x^{n+m} \cdot x$ and are hence equal to each other.

(a2) $(xy)^n = x^n y^n$.

Suppose $n = 0$. Then, $(xy)^0 := 1$. And $x^0 y^0 = 1 \times 1 = 1$.

We inductively assume that, $(xy)^n = x^n y^n$.

We are interested to prove that $(xy)^{n+1} = x^{n+1}y^{n+1}$. $(xy)^{n+1} := (xy)(xy)^n = (xy)(x^n y^n) = (x \cdot x^n)(y \cdot y^n) = x^{n+1}y^{n+1}$.

(a3) $(x^n)^m = x^{nm}$.

Clearly, $1^n = 1$. To see this, $1^0 := 1$. Assume that, $1^n = 1$. Then, $1^{n+1} = 1 \times 1^n = 1 \times 1 = 1$.

To prove the main result, we induct on n keeping m fixed.

Suppose $n = 0$. Then, $(x^0)^m := 1^m = 1$. $x^{nm} = x^0 = 1$.

We inductively assume that, $(x^n)^m = x^{nm}$.

We are interested to prove that $(x^{(n+1)})^m = x^{(n+1)m}$.

$(x^{n+1})^m = (x \cdot x^n)^m = (x^m)(x^n)^m = (x^m)(x^n m) = x^{(m+nm)} = x^{m(n+1)} = x^{(n+1)m}$

(b) Firstly, it is easy to see that, if $n > 0$, $0^n = 0$. To prove this elementary fact, observe that $0^1 := 0 \times (0^0) = 0 \times 1 = 0$. We inductively assume that $0^n = 0$. Now, $0^{n+1} = (0)(0^n) = (0)(0) = 0$.

In the opposite direction, if $x^1 = 0$, it implies $(x)(x^0) = 0$, so $x1 = 0 \implies x = 0$. Suppose $x^n = 0$, $n > 0$. If $x^{n+1} = 0$, then $(x)(x^n) = 0$. Atleast one of x or x^n must be zero. $x^n = 0 \implies x = 0$. $x = 0 \implies x = 0$. In both cases, $x = 0$. This proves the claim.

(c) Suppose $n = 0$. Then, if $x \geq y \geq 0$ it implies $x^0 \geq y^0 \geq 0$, since $1 \geq 1 \geq 0$.

We inductively assume that $(x \geq y \geq 0) \implies (x^n \geq y^n \geq 0)$.

We are interested to prove that $(x \geq y \geq 0) \implies (x^{n+1} \geq y^{n+1} \geq 0)$. Clearly, $x^n \geq y^n \implies (x)(x^n) \geq (x)(y^n) \geq (y)(y^n)$, because positive multiplication preserves order. So, $x^{(n+1)} \geq y^{(n+1)}$. Moreover, $y^n \geq 0 \implies (y)(y^n) \geq 0 \implies y^{(n+1)} > 0$. Hence, $x^{(n+1)} \geq y^{(n+1)} \geq 0$.

(d) Suppose $n = 0$. Then, $|x^0| := |1| = 1$ and $|x|^0 := 1$, since $|x| \in \{x, -x\}$.

We inductively assume that, $|x^n| = |x|^n$.

We are interested to prove that, $|x^{(n+1)}| = |x|^{(n+1)}$. The left hand side can be simplified as: $|x^{(n+1)}| = |(x)(x^n)| = |x||x^n| = |x||x|^n = |x|^{(n+1)}$, which equals the expression on the right hand side. This proves the claim. $\square$

Definition 2.14. (Exponentiation to a negative number.) Let x be a non-zero rational number. Then, for any negative integer $-n$, we define $x^{-n} := 1/x^n$.

Thus, for instance $x^{-3} = 1/x^3 = 1/(x \times x \times x)$. We now have x^n defined for any integer n, whether n is positive, negative or zero. Exponentiation with integer exponents has the all of the properties, that hold for natural number exponents. It is an easy exercise to prove this.

2.4 Gaps in the rational numbers.

***Theorem:* 2.4.1: Interspersing of integers with rational numbers.**

Let x be a rational number. Then, there exists an integer n such that $n \leq x < (n+1)$. In fact, this integer is unique. (For each x, there is one and only one integer n for which $n \leq x < (n+1)$). In particular, there exists a natural number N such that $N > x$ (that is there is no such thing as a rational number which is larger than all the natural numbers).

Remark 2.6. The integer n for which $n \leq x < (n+1)$ is sometimes referred to as the integer part of x and is sometimes denoted $n = \lfloor x \rfloor$.

Proof. Let the rational number $x := a/b$ where a, b are integers. By the Euclidean algorithm, there exists integers q, r such that

$$a = qb + r$$

where $0 \leq r < b$.

$0 \leq r < b$
$\implies qb \leq qb + r < qb + b$
$\implies qb \leq a < (q+1)b$
$\implies q \leq a/b < (q+1)$

Hence, for every rational number a/b there exists a unique integer n such that $n \leq a/b < (n+1)$. $\square$

Also, between any two rational numbers, there is atleast one additional rational number.

Theorem: 2.4.2: Interspersing of rationals by rationals.

If x and y are two rational numbers such that $x < y$, there exists a third rational z such that $x < z < y$.

Proof. We set $z := (x+y)/2$. Since, $x < y$, and $1/2 = 1//2$ is positive, using the property that positive multiplication preserves order, $x/2 < y/2$. Since, addition also preserves order, if we add $y/2$ to both sides, $x/2 + y/2 < y$. So, $z = (x+y)/2 < y$. On the other hand, if we add $x/2$ to both sides, we obtain $x/2 + x/2 < x/2 + y/2 \implies x < (x+y)/2 = z$. Thus, $x < z < y$ as desired. □

Despite the rationals having this denseness property, they are still incomplete; there are still an infinite number of "gaps" or "holes" between the rationals, although this denseness property does ensure that these holes are in some sense infinitely small. For instance, we will show that the rational numbers do not contain any square root of two.

Theorem: 2.4.3: $\sqrt{2}$ is irrational.

There does not exist any rational number x for which $x^2 = 2$.

Proof. Suppose, for the sake of contradiction that we had a rational number x for which $x^2 = 2$. Clearly, $x \neq 0$. We may assume that x is positive, for if x were negative, we could just replace x by $-x$ (since $x^2 = (-x)^2$). Thus, suppose, x is a positive rational number, $x := p/q$ where p, q are positive integers. So, $(p/q)^2 = 2$, which we can rearrange as $p^2 = 2q^2$. Define a natural number p to be even if $p = 2k$ for some natural number k, and odd if $p = 2k+1$ for some natural number k. Every natural number is either even or odd, but not both. If p is odd then p^2 is also odd, because $p^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1 = 2l + 1$, where l is a natural number, which contradicts $p^2 = 2q^2$. Thus, p is even, that is, $p = 2k$ for some natural number k. Since, p is positive, k is also positive. Inserting $p = 2k$ into $p^2 = 2q^2$ we obtain, $4k^2 = 2q^2$. So, $q^2 = 2k^2$.

To summarize, we started with a pair (p, q) of positive integers such that $p^2 = 2q^2$ and ended up with a pair (q, k) of positive integers such that $q^2 = 2k^2$. Since, $p^2 = 2q^2$, $q^2 < p^2 \implies q < p$. If we rewrite $p' := q$ and $q' := k$, we thus can pass from one solution (p, q) to the equation $p^2 = 2q^2$ to a new solution (p', q') to the same equation which has a smaller value of p. But, then we can repeat this procedure again and again, obtaining a sequence $(p'', q''), (p''', q''')$ etc. of solutions to $p^2 = 2q^2$, each one with a small value of p than the previous, and each one consisting of positive integers. But this contradicts the principle of infinite descent (see problems). This contradiction shows that our initial assumption was wrong, we could not had a rational x for which $x^2 = 2$. □

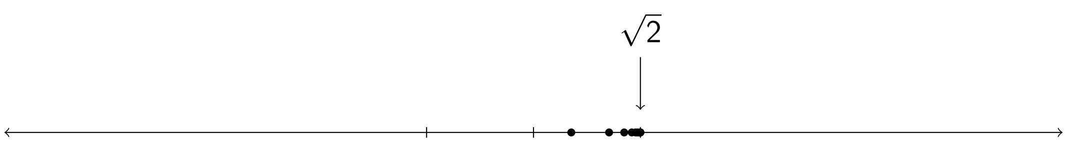
On the other hand, we can get rational numbers which are arbitrarily close to $\sqrt{2}$.

***Theorem:* 2.4.4**

For every rational number $\epsilon > 0$, there exists a non-negative rational number x such that $x^2 < 2 < (x+\epsilon)^2$.

Proof. Let $\epsilon > 0$ be rational. Suppose for the sake of contradiction, there is no non-negative rational number x for which $x^2 < 2 < (x+\epsilon)^2$. This means that whenever x is non-negative and $x^2 < 2$, we must also have $(x+\epsilon)^2 < 2$ (note that $(x+\epsilon)^2$ cannot equal 2, by the above proposition). If we set $x = 0$, since $x^2 = 0^2 < 2$, then $(x+\epsilon)^2 = (0+\epsilon)^2 = \epsilon^2 < 2$. Which then implies $(\epsilon+\epsilon)^2 = (2\epsilon)^2 < 2$, and indeed a simple induction shows that $(n\epsilon)^2 < 2$ for every natural number n. Not that $n\epsilon$ is non-negative for every natural number n. But, by the theorem on the density of rationals, we can find an integer n, such that $n > 2/\epsilon$, which implies that $n\epsilon > 2$, which implies that $(n\epsilon)^2 > 4 > 2$, contradicting the claim that $(n\epsilon)^2 < 2$ for all natural numbers n. This contradiction gives the proof. $\square$

Example 2.4.1. If $\epsilon = 0.001$, we can take $x = 1.414$, since $x^2 = 1.999396$ and $(x+\epsilon)^2 = 2.002225$.



$\sqrt{2}$ as a limit of a sequence of rational numbers.

The above proposition indicates that while the set **Q** does not actually have $\sqrt{2}$ as a member, we can get as close as we wish $\sqrt{2}$. For instance, the sequence of rationals

$$1.4, 1.41, 1.414, 1.4142, 1.41421, 1.414212, 1.4142128$$

seem to get closer and closer to $\sqrt{2}$, as their squares indicate:

$$1.96, 1.9881, 1.999396, 1.99996164, 1.9999899241, 1.99999598094, 1.999999784368$$

Thus, it seems that we can take a square root of 2, by taking a limit of a sequence of rationals. This is how we shall construct the real numbers in the next chapter. There is another way to do so using something called Dedekind cuts.

Chapter 3

The Real numbers $\mathbf{R}$.

We have rigorously constructed three fundamental number systems : the natural number system $\mathbf{N}$, the integers $\mathbf{Z}$ and the rational number system $\mathbf{Q}$. We defined the natural numbers using the five Peano axioms and postulated that such a system existed; this is plausible, since the natural numbers correspond to the very intuitive and fundamental notion of *sequential counting.* Using that number system one could then recursively define addition and multiplication, and verify that they obeyed the usual laws of algebra. We then constructed the integers by taking formal differences of the natural numbers $a--b$. We then constructed the rationals by taking formal quotients of the integers $a//b$.

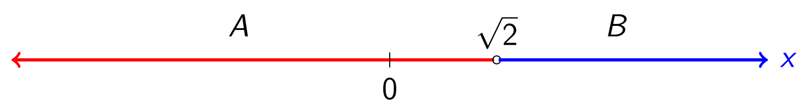
The rational number system is already sufficient to do a lot mathematics - much of high school algebra works just fine, if one only knows about the rationals. However, there is a fundamental area of mathematics where the rational number system does not suffice - that of geometry. For instance, a right-angled triangle with both sides equal to 1, gives a hypotenuse of $\sqrt{2}$., which is an **irrational number**. Things get even worse when one sees numbers such as π or $\cos 1$, which turn out to be in some sense "even more" irrational then $\sqrt{2}$. These numbers are known as **transcendental numbers**. Thus, in order to have a number system which can adequately describe geometry - or even something as simple as measuring lengths on a line - one needs to replace the rational number system with the real number system. Since the differential and integral calculus is also intimately tied up with geometry - think of slopes of tangents or areas under a curve - calculus also requires the real number system to function properly.

However, a rigorous way to construct the reals from the rationals turns out to be somewhat difficult - requiring a bit more machinery than what was needed to pass from naturals to integers, or from the integers to the rationals. In those constructions, the task was to introduce one more algebraic operation to the number system e.g. one can get integers from rationals by introducing subtraction and get the rationals from the integers by introducing division. But, to get the reals from the rationals is to pass from a discrete system to a continuous one and requires the introduction of a somewhat different notion - that of a *limit.* The limit is a concept which on one level is quite intuitive, but to pin down rigorously turns out to be difficult. (It was only in the 19th century, that mathematicians such as Cauchy and Dedekind figured out how to deal with limits rigorously).

In the last chapter, we explored the **gaps** in rational numbers; now we shall fill these gaps using limits to create real numbers. The real number system will end up being a lot like the rational numbers, but will have some new operations - notably that of **supremum**, which can be used to define limits and thence everything else that calculus needs.

3.1 Cut.

The set $\mathbf{Q}$ of rational numbers is incomplete. It has "gaps", one of which occurs at $\sqrt{2}$. These gaps are more like pinholes, they have infinitely small width. Incompleteness is what is *wrong* with $\mathbf{Q}$. Our goal is to complete $\mathbf{Q}$ by filling it's gaps. An elegant method to arrive at this goal is **Dedekind cuts** in which one visualises real numbers as places at which a line may be cut with scissors. See the figure below.



A Dedekind cut

Definition 3.1. *Cut.* A cut in $\mathbf{Q}$ is a pair of subsets A, B of $\mathbf{Q}$ such that :

(a) *Non-trivial.* $A \cup B = \mathbf{Q}$, $A \neq \phi$, $B \neq \phi$, $A \cap B = \phi$.

(b) *Closed downwards.* If $p \in A$ and $q < p$, then $q \in A$.

(c) *No largest element.* If $p \in A$, there exists a $r \in A$, such that $p < r$.

A is the left-hand part of the cut and B is the right-hand part. We denote the cut $x = A|B$. Making a semantic leap, we now answer the question, "what is a real number?"

Definition 3.2. A *real number* is a cut in $\mathbf{Q}$.

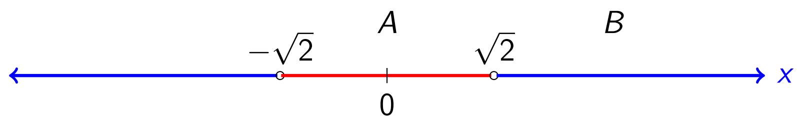
$\mathbf{R}$ is the class of all real numbers $x = A|B$. We will show in a natural way $\mathbf{R}$ is a complete ordered field containing $\mathbf{Q}$. Before spelling out what this exactly means, here are two examples of cuts.

Example 3.1.1. $A|B = \{r \in \mathbf{Q} : r < 1\}|\{r \in \mathbf{Q} : r \geq 1\}$.

It is convenient to say that $A|B$ is a **rational cut**, if it is a cut like in example (3.1.1): For some fixed rational number c, A is the set of all rational numbers $< c$ while B is the rest of $\mathbf{Q}$. The B-set of a rational cut contains the smallest element c, and conversely, if $A|B$ is a cut in $\mathbf{Q}$ and B contains the smallest element c then $A|B$ is a rational cut at C.

Example 3.1.2. $A|B = \{r \in \mathbf{Q} : r^2 < 2\}|\{r \in \mathbf{Q} : r^2 \geq 2\}$.

$A|B$ is not a cut, because it is not closed downwards. A pictorial representation of this is shown below.



$$A|B := \{r \in \mathbb{Q} : r^2 < 2\}|\{r \in \mathbb{Q} : r^2 \geq 2\}$$

There is also an order relation $x \leq y$ on cuts.

Definition 3.3. If $x = A|B$ and $y = C|D$ are cuts such that $A \subseteq C$, then x is **less than or equal to** y and we write $x \leq y$. If $A \subset C$ and $A \neq C$, then x is **less than** y and we write $x < y$.

The distinguishing property of $\mathbf{R}$ from $\mathbf{Q}$ and which is at the bottom of every significant theorem about $\mathbf{R}$ involves upper bounds and least upper bounds or, equivalently lower bounds and greatest lower bounds.

Definition 3.4. Upper Bound. Let S be an ordered set. Suppose $E \subset S$. If there exists a $\beta \in S$, such that $x \leq \beta$ for all $x \in E$, we say that E is *bounded above*, and call β the *upper bound* of E.

We also say that E is **bounded above** by β.

Definition 3.5. Lower Bound. Suppose $E \subset S$. The set E is *bounded below* if there exists a lower bound l such that $l \leq x$ for all $x \in E$.

Definition 3.6. Least Upper Bound. A real number α is said to be the *least upper bound* for a set $E \subseteq \mathbf{R}$ if it meets the following two criteria:

1) α is an upper bound for E. This implies $x \leq \alpha$ for all $x \in E$.

2) if β is any upper bound for E, then $\alpha \leq \beta$.

The least upper bound is also frequently called the *supremum* of the set E. Although the notation $\alpha = \sup E$ is still common, we will always write $\alpha = \sup E$ for the least upper bound.

Example 3.1.3. 3 is an upper bound for the set of negative integers.

-1 is an upper bound for the set of negative integers.

1 is the least upper bound for the set of rational numbers $1 - \frac{1}{n}$with $n \in \mathbf{N}$.

-100 is an upper bound for the empty set ϕ.

A least upper bound for S may or may not belong to S. That is why, you should say *least upper bound for* S rather than *least upper bound of* S.

Example 3.1.4. Let $A|B := \{r \in \mathbb{Q} : r^2 < 2\}|\{r \in \mathbb{Q} : r^2 \geq 2\}$. Show that $\frac{3}{2}$ is an upper bound for the A-set.

Solution. Suppose for the sake of contradiction, there exists $x \in A$ such that $x \geq \frac{3}{2}$. Then, $x^2 \geq \frac{3}{2}x \geq (\frac{3}{2})^2$. This follows from the fact that positive multiplication preserves order in

$\mathbb{Q}$. So, $x^2 \geq \frac{9}{4} > 2$. This is a contradiction, because $x^2 < 2$ for all $x \in A$. Hence, our initial assumption is wrong. $x \leq \frac{3}{2}$ for all $x \in A$. So, $\frac{3}{2}$ is an upper bound for A.

Example 3.1.5. Let $E := \{\frac{1}{2}, 1, 2\}$. Show that the $\sup E = 2$.

Proof. (1) 2 is an upper bound for E. If $x \in E$, $x \leq 2$.

(2) Suppose there were an upper bound $\gamma < 2$. Well, then γ wouldn't be bigger than 2. More precisely, if $x \in E$ and $x = 2$, x is not less than or equal to γ. So, γ cannot be an upper bound for A. □

Example 3.1.6. Let $E := \mathbb{Q}_-$ = the set of all negative rational numbers. Show that the $\sup \mathbb{Q}_- = 0$.

Proof. (1) By definition, if x is any negative rational number, $x < 0$. So, 0 is an upper bound for the set of negative rational numbers.

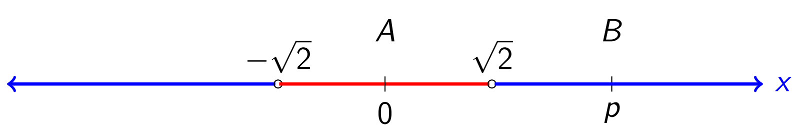
(2) Suppose, for the sake of contradiction, let's assume there is an upper bound, an itty-bitty tiny number to the left of 0, that's greater than all the negative rationals. Let γ be an upper bound less than 0. γ is a negative rational number. Since γ is an upper bound for $\mathbb{Q}_-$, it is bigger than all negative rational numbers. So, $x \leq \gamma$ for all $x \in \mathbb{Q}_-$. But, consider the rational $\frac{\gamma}{2}$. $\frac{1}{2} < 1 \implies \frac{\gamma}{2} > \gamma$. Moreover, $\gamma/2 \in \mathbb{Q}_-$, so γ is not an upper bound. This contradicts our initial assumption. $\sup E = 0$. □

Example 3.1.7. What if $E := \mathbb{Q}$ = the set of all rational numbers? What is $\sup \mathbb{Q}$?.

Proof. E is unbounded above, so the supremum does not exist. We often say, in practice that $\sup E = +\infty$. □

Example 3.1.8. Consider again the set $A := \{r \in \mathbb{Q} : r^2 < 2\}$. We saw that $\frac{3}{2}$ is an upper bound for A. But, does it have a least upper bound in Q?

Proof. Here, the set A is not unbounded. Even though A is bounded, A does not have a least upper bound in $\mathbb{Q}$, because there is no rational number $r \in \mathbb{Q}$, such that $r^2 = 2$. One way to show rigorous, A does not have a least upper bound is to show that, no matter what upper bound you give me, it is always possible to find a smaller one.



$$A|B := \{r \in \mathbb{Q} : r^2 < 2\}|\{r \in \mathbb{Q} : r^2 \geq 2\}$$

Let p be an upper bound of A. Then, for all $r \in A$, $r \leq p$. Can we find a q lying between the (right) edge of A and p? There is a temptation to take an average of p and $\sqrt{2}$ and come up with a number like $q = \frac{p+\sqrt{2}}{2}$, but we haven't defined $\sqrt{2}$ yet. So, we use a clever trick. We subtract a rational length from p, so $q^2 > 2$, but $q < p$. Define,

$$q := p - \frac{p^2 - 2}{p + 2}$$

By the trichotomy property, there are three cases to consider $p^2 > 2$, $p^2 = 2$ and $p^2 < 2$. $p^2 = 2$ is ruled out, since p is rational.

(i) $p^2 > 2$.

$(\frac{p^2-2}{p+2})$ is positive. So, $q < p$. Upon simplification, we have :

$$\begin{aligned} q &= p - \frac{p^2 - 2}{p + 2} \\ &= \frac{p(p+2) - (p^2 - 2)}{p + 2} \\ &= \frac{p^2 + 2p - p^2 + 2}{p + 2} \\ &= \frac{2p + 2}{p + 2} \end{aligned}$$

We would also like to show that $q^2 > 2$.

$$\begin{aligned} q^2 - 2 &= \left[\frac{(2p+2)}{(p+2)}\right]^2 - 2 \\ &= \left[\frac{4(p^2 + 2p + 1)}{(p^2 + 4p + 4)}\right] - 2 \\ &= \frac{4p^2 + 8p + 4 - 2p^2 - 8p - 8}{(p+2)^2} \\ &= \frac{2p^2 - 4}{(p+2)^2} \\ &= \frac{2(p^2 - 2)}{(p+2)^2} \end{aligned}$$

Since $p^2 > 2$, the right hand side of the equation is positive. So, $q^2 > 2$.

(ii) $p^2 < 2$. Clearly, we see that $q > p$ and $q^2 < 2$.

Note that, our upper bound p was arbitrary. We showed that one can always find q such that $q < p$, $q^2 > 2$. Hence, A has no least upper bound in $\mathbb{Q}$. $\square$

The set A will have a supremum in $\mathbb{R}$.

3.2 The completeness property of $\mathbb{R}$.

This brings us to the final, and the most distinctive, assumption about the real number system. We must clearly find some way to articulate what we mean by insisting that $\mathbb{R}$ does not contain the gaps that permeates $\mathbb{Q}$. Because this is the defining difference between the rational numbers and the real numbers, we will be excessively precise about how we phrase this assumption, hereafter referred to as the *Axiom of completeness.*

Axiom 3.2.1. Axiom of completeness. *Every non-empty subset of S that is bounded above has a least upper bound.*

***Theorem:* 3.2.1: $\mathbb{R}$ is complete.**

The set $\mathbb{R}$ constructed by means of Dedekind cuts, is ***complete*** in the sense that it satisfies the least upper bound property:

"Every non-empty subset of real numbers that is bounded above has a least upper bound".

Proof. Let $\mathcal{C} \subset \mathbb{R}$ be any non-empty collection of cuts which is bounded above, say by the cut $X|Y$. Define,

$C := \{a \in \mathbb{Q} :$ for some cut $A|B \in \mathcal{C}$ we have $a \in A\}$ and $D =$ rest of $\mathbb{Q}$

It is easy to see that $z := C|D$ is a cut.

(i) *Non-trivial.* $\mathcal{C}$ is non-empty. If an A-set belongs to $\mathcal{C}$, then $A \subset C$. So, C is non-empty.

(ii) *Closed downward.* Consider $p \in C$. Then, $p \in A_1$, where A_1 is the half-set of some cut in $\mathcal{C}$. Since, $A_1|B_1$ is a cut, it is close downward. If we were to choose $q < p$, then $q \in A_1$. Since $A_1 \subset C$, $q \in C$.

(iii) *No largest element.* Consider an element $p \in c$. Then, p belongs to the half-set A of some cut. Suppose $p \in A_1$ for some $A_1 \in \mathcal{C}$. $A_1|B_1$ is a cut. Hence, A_1 has no largest element. There exist $r \in A_1$ such that $r > p$. So, we see that $r \in C$. C has no largest element.

Clearly, it is an upper bound for $\mathcal{C}$, since the A-set for every element of $\mathcal{C}$ is a subset of C. Hence, $x \leq z$ for all $x \in \mathcal{C}$.

We are interested to prove that $z = C|D$ is the least upper bound for the set $\mathcal{C}$. Let $z' = C'|D'$ be any upper bound of $\mathcal{C}$ such that $z' < z$. In such case, $C' \subset C$, so there exists an element $s \in C, s \notin C'$. But, $s > z'$. Therefore, z' is not an upper bound for $\mathcal{C}$.

That is among all upper bounds, $z = C|D$ is the least. □

3.3 Order properties of $\mathbb{R}$.

We have gone from $\mathbb{Q}$ to $\mathbb{R}$ by pure thought. We have defined $\mathbb{R}$ to be the set of all cuts $x = A|B$. We have defined an order on $\mathbb{R}$; if $x := A|B, y = C|D$ are two real numbers, we say that x is ***less than*** y, if and only if $A \subset C$. Let's prove that, this is indeed a legitimate definition of order.

Theorem: 3.3.1: Order trichotomy.

If $x := A|B$ and $y := C|D$ are any two real numbers, exactly one of the below statements must hold true:
(i) $x = y$
(ii) $x < y$
(iii) $x > y$

Proof. (i) We first prove that at most one of the properties could hold true.

Suppose that $x < y$ and $x = y$. Then $A \subset C$ and $A = C$. But, if A is a proper subset of C, $A \subset C$, it implies $A \neq C$. So, both $x < y$ and $x = y$ cannot be true simultaneously. Similarly, $x > y$ and $x = y$ cannot both hold true simultaneously. Finally, suppose that whenever $x < y$, we have $x > y$. By definition, whenever $A \subset C$, we have $C \subset A$. This is impossible, if $A \neq C$. Hence, at most one of statements (i), (ii) and (iii) is true.

(ii) We show that atleast one of the properties must hold true. Since, x and y are cuts, they are closed downward. Clearly, either of $A \subset C$, $A = C$, $C \subset A$ must be true. Hence, exactly one $A|B < C|D, A|B = C|D, A|B > C|D$ must be true. □

Theorem: 3.3.2: Order properties.

If $x := A|B, y := C|D, z := E|F$ be real numbers. The following order properties are satisfied in $\mathbb{R}$:

(O1) *Reflexive.* $x \leq x$.

(O2) *Anti-symmetric.* If $x \leq y$ and $y \leq x$, then $x = y$.

(O3) *Transitive.* If $x < y$ and $y < z$, then $x < z$.

Proof. (O1) *Reflexive.* Let $\alpha := A|B$ be a cut. Since a set is a subset of itself, $A \subseteq A$. So, $\alpha \leq \alpha$.

(O2) *Anti-symmetric.* Let $\alpha := A|B$ and $\beta := C|D$ be cuts. $\alpha \leq \beta \implies A \subseteq C$. $\beta \leq \alpha \implies C \subseteq A$. $(A \subseteq C) \wedge (A \subseteq C) \implies A = C$. Therefore, $\alpha = \beta$.

(O3) *Transitive.* Let $\alpha := A|B$, $\beta := C|D$, $\gamma := E|F$ be cuts. The proper subset of a proper subset is also a proper subset. So, $(A \subset C) \wedge (C \subset E) \implies (A \subset E)$. We therefore conclude $\alpha < \gamma$. □

We can prove the remaining order properties (O4)-(O5), once we have defined addition and multiplication.

3.4 Field properties of $\mathbb{R}$.

To be more complete, we describe the natural arithmetic of cuts. Let cuts $x = A|B$ and $y = C|D$ be given. How do we add them? Subtract them? ... Generally, the answer is to do the corresponding operation to the elements comprising the two halves of the cuts, being careful about the negative numbers. The sum of x and y is $x + y = E|F$ where

$$E = \{r \in \mathbb{Q} : \text{ for some } a \in A \text{ and for some } c \in C, r = a + c\}$$
$$F = \text{the rest of } \mathbb{Q}$$

Lemma 3.4.1. *The sum of two cuts α, β; $\alpha + \beta$ is also a cut.*

Proof. Non-trivial. We have $\alpha \neq \phi$, $\beta \neq \phi$. If $x \in \alpha$ and $y \in \beta$, by definition $(x+y) \in \alpha+\beta$. Hence, $\alpha + \beta \neq \phi$.

Closed downward. If $p \in \alpha + \beta$, then $p = r + s$ where $r \in \alpha$, $s \in \beta$. Suppose $q < p$. Then, $q - s < r$. This implies that $q - s \in \alpha$. But, $q = (q - s) + s$. We deduced that, $q - s \in \alpha$ and $s \in \beta$, so $q \in \alpha + \beta$.

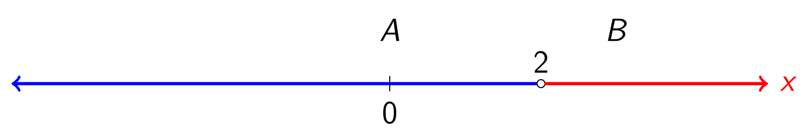
No largest element. Suppose $p \in \alpha + \beta$. We know that, $p = r + s$ for some $r \in \alpha, s \in \beta$. Since, β is a cut, it has not largest element. There exists $t \in \beta$, such that $t > s$. So, $r + t > r + s$. Therefore, $r + t > p$ and $r \in \alpha, t \in \beta$. Thus, $\alpha + \beta$ has no largest element. ☐

The zero cut is 0^*, it is the set of all negative rational numbers and $0^* + x = x$ for all $x \in \mathbb{R}$. The additive inverse of $x = A|B$ is $-x = C|D$ where

$$C = \{p \in \mathbb{Q} : \exists r > 0, -p - r \notin A\}$$
$$D = \text{the rest of } \mathbb{Q}$$

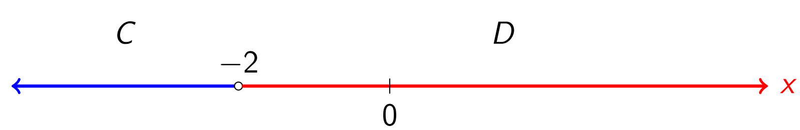
If you look at a pictorial representation of x and $-x$, it helps to see why $x + (-x) = 0^*$.

Let the rational cut $x := \{r \in \mathbb{Q} : r < 2\}|\{r \in \mathbb{Q} : r \geq 2\}$. Then, x can be visualised as,



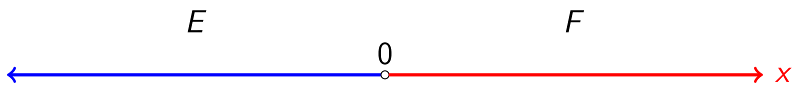
$$x := \{r \in \mathbb{Q} : r < 2\}|\{r \in \mathbb{Q} : r \geq 2\}$$

Whereas, $(-x)$ looks like the below set,



$$(-x) := \{r \in \mathbb{Q} : r < -2\}|\{r \in \mathbb{Q} : r \geq 2\}$$

Clearly, addition of the two cuts (pair-wise addition of all blue points) should yield 0^*.



$$0^* := \{r \in \mathbb{Q} : r < 0\} | \{r \in \mathbb{Q} : r \geq 0\}$$

Theorem: 3.4.1: Field axioms for addition in ℝ.

Let x, y, z be real numbers. $\mathbb{R}$ satisfies the following field axioms for addition.
(A1) *Commutative.* $x + y = y + x$
(A2) *Associative.* $(x + y) + z = x + (y + z)$
(A3) *Zero element.* $x + 0 = x$
(A4) *Additive inverse.* $x + (-x) = 0$.

Proof. (A1) Let α, β be cuts. Then we have,

$$\begin{aligned}
\alpha + \beta &:= \{r + s : r \in \alpha, s \in \beta\} \\
&= \{s + r : r \in \alpha, s \in \beta\} \\
&= \beta + \alpha
\end{aligned}$$

$r +_{\mathbb{Q}} s = s +_{\mathbb{Q}} r$, because, addition of rational numbers is commutative.

(A2) Let α, β, γ be cuts. Then we have,

$$\begin{aligned}
(\alpha + \beta) + \gamma &:= \{(r + s) + t : r \in \alpha, s \in \beta, t \in \gamma\} \\
&= \{r + s + t : r \in \alpha, s \in \beta, t \in \gamma\} \\
&= \{r + (s + t) : r \in \alpha, s \in \beta, t \in \gamma\} \\
&= \alpha + (\beta + \gamma)
\end{aligned}$$

(A3) Let $0^* =$ the set of all negative rational numbers. This is the real number 0. Let α be any cut.

$$\begin{aligned}
\alpha + 0 &= \{r + s : r \in \alpha, s \in 0^*\} \\
&= \{r + s : r \in \alpha, s < 0\} \\
&= \{r + s : (r + s) \in \alpha\} \\
&= \alpha
\end{aligned}$$

To elaborate, $s < 0$. So, $r + s < r$. Since, $r \in \alpha \implies r + s \in \alpha$, since α is a cut and it is closed downwards.

(A4) Let α be a cut. If r is a small positive rational number, $r > 0$, then the set $\beta = -\alpha$ contains all points p such that $-p - r \notin \alpha$. We are interested to show that $\alpha + \beta = 0^*$.

β is a cut.

Non-trivial.

If $s \notin \alpha$ and $p = -s - 1$, then $-p - 1 \notin \alpha$. Hence, $p \in \beta$. So, β is not empty.

Closed downward.

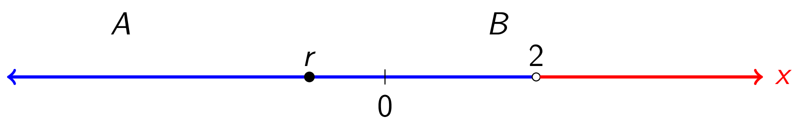
Pick $p \in \beta$, $r > 0$. We know that, $-p - r \notin \alpha$. If $q < p$, then $-q - r > -p - r$, hence $-q - r \notin \alpha$. So, $q \in \beta$.

No largest element.

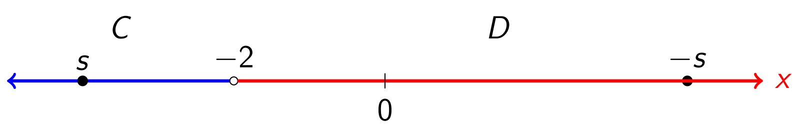
Put $t = p + r/2$. Then $t > p$, and $-t - r/2 = -p - r \notin \alpha$, so that $t \in \beta$. Hence, β has no largest element.

$\alpha + \beta = 0^*$.

If $r \in \alpha$, $s \in \beta$, then $-s \notin \alpha$. Hence, $r < -s$. Therefore, $r + s < 0$ for all r, s.



$$\alpha := \{r \in \mathbb{Q} : r < 2\}|\{r \in \mathbb{Q} : r \geq 2\}$$



$$\beta := \{r \in \mathbb{Q} : r < -2\}|\{r \in \mathbb{Q} : r \geq -2\}$$

□

Multiplication is trickier to define. It helps to first say that the cut $x = A|B$ is **positive** if $0^* \leq x$ or **negative** $x < 0^*$. Since 0 lies in A or B, a cut is either negative, zero or positive. If addition was defined as the sum of all pairs $a + b$, $a \in \alpha$, $b \in \beta$, what do you think multiplication of cuts would look like? We just need to be careful multiplying the negative rationals in both cuts. If α, β are positive cuts, then:

$$\alpha \cdot \beta = \{r \in \mathbb{Q} : (r = ab, a > 0, b > 0, a \in \alpha, b \in \beta) \vee (r \leq 0)\}$$

If α is positive and β is negative then we define the product to be $-(\alpha \cdot (-\beta))$. Since α and $-\beta$ are positive cuts this makes sense and is a negative cut. Similarly, when α is negative and β is positive, then by definition their product is the negative cut $-((-\alpha)\cdot(\beta))$, while if α and β both are negative then their product is the positive cut $((-\alpha)\cdot(-\beta))$. Finally, if α or β is the zero cut 0^* we define $\alpha\cdot\beta = 0^*$. This makes five cases in the definition.

Lemma 3.4.2. *Product of reals. The product of two real numbers is a real number.*

If α, β are cuts, their product $\alpha\cdot\beta$ is also a cut.

Proof. Non-trivial.

Take $a \in \alpha$, $b \in \beta$, $a > 0, b > 0$. Then, $ab \in \alpha\cdot\beta \implies \alpha\cdot\beta \neq \phi$. Take $a' \notin \alpha$ and $b' \notin \beta$, then $a' > a, b' > b$. This implies, $a'b' > a'b > ab$ for all choices of ab. So, $a'b' \notin \alpha\cdot\beta$. Hence, $\alpha\cdot\beta \neq \mathbb{Q}$.

Closed downward.

Pick $p \in \alpha\cdot\beta$. Let $p = ab$ for some $a \in \alpha, b \in \beta$. If $q < p$, then $q < ab \implies q/a < b$. But, $b \in \beta$. β is a cut and closed downward. So, $q/a \in \beta$. q can be written as, $q = (q/a)\cdot a$, where $q/a \in \beta$, $a \in \alpha$. This means, $q \in \alpha\cdot\beta$.

No largest element.

Take $r > p$. Then, $r > ab$. So, $r/b > a$. But, $a \in \alpha$. α is cut and has no largest element. Therefore, $\exists (r/b) \in \alpha$, such that $r/b > a$. The rational number r can be expressed as the product $r = (r/b)\cdot b$, where $r/b \in \alpha$, $b \in \beta$. Hence, $r \in \alpha\cdot\beta$. □

***Theorem:* 3.4.2: $\mathbb{R}$ satisfies multiplication axioms.**

The set of real numbers satisfies the multiplication axioms of a field. Let $x, y, z \in \mathbb{R}$.

(M1) *Commutative.* $xy = yx$
(M2) *Associative.* $(xy)z = x(yz)$
(M3) *Multiplicative identity.* $1x = x1 = x$
(M4) *Multiplicative inverse.* $xx^{-1} = 1 = x^{-1}x$

Proof. (M1) *Commutative.*

Let $x = A|B, y = C|D$ be positive cuts. The multiplication cut $x\cdot y = \{r : (r = ac, a \in A, c \in C, a > 0, c > 0) \vee (r \leq 0)\}$. But, $ac = ca$ for all rational numbers $a, c \in \mathbb{Q}$. Hence, $x\cdot y = y\cdot x$.

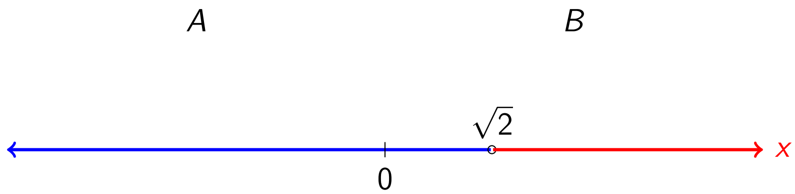
(M2) *Associative.* Let $x = A|B, y = C|D, z = E|F$ be positive cuts. Since $(ac)e = a(ce)$ for all rational numbers $a, c, e \in \mathbb{Q}$, multiplication of cuts is associative. $(xy)z = x(yz)$.

(M3) *Multiplicative identity* Let $x = A|B$ be a positive cut. Pick $a \in A$ and $r \in 1^*$. $r < 1 \implies ar < a$ for all choices of a. So, the members of the product $1^*\cdot x$ are the same as elements of A. Hence, $1x = x$. Similarly, $x1 = x$.

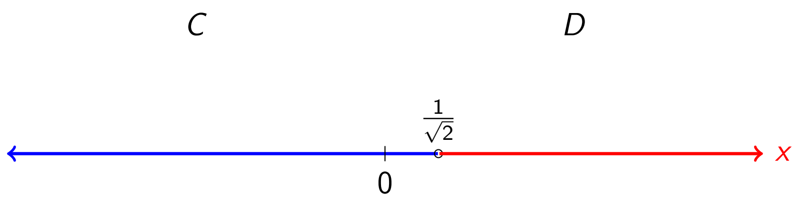
(M4) *Multiplicative inverse* Let $\alpha = A|B$ be a positive cut. The inverse cut $\alpha^{-1} = C|D$ can intuitively be defined as,

$$\beta = \alpha^{-1} := \{r \in \mathbb{Q} : (r \leq 0) \vee (\exists b \in B, b \text{ not the smallest element in B}, r = \frac{1}{b})\}$$

It is an easy exercise to prove that α^{-1} is a cut.



$$\alpha := \{r \in \mathbb{Q} : (r^2 < 2) \vee (r < 0)\}|\text{rest of } \mathbb{Q}$$



$$\alpha^{-1} := \{r \in \mathbb{Q} : (r^2 < \tfrac{1}{2}) \vee (r < 0)\}|\text{rest of } \mathbb{Q}$$

Non-trivial.

Since $B \neq \phi$ and $B \neq \mathbb{Q}$, C is non-trivial.

Closed downward.

Let $x > 0$ be an arbitrary element in C. If $y < x$, then $\frac{1}{y} > \frac{1}{x}$. But, $\frac{1}{x} \in B$. Therefore, $\frac{1}{y} \in B$. So, $y \in C$. Hence, if α^{-1} is closed downwards.

No largest element.

If C has a largest member $c > 0$ and $\frac{1}{c} \in B$, then $\frac{1}{c}$ is the smallest element in B.

Without loss of generality, suppose $\beta < \alpha$. Let $r \in \alpha$, $s \in \beta$. Then, $\frac{1}{s} \notin \alpha$. So, $\frac{1}{s} > r$. Therefore, $rs < 1$. Consequently, $x \cdot \frac{1}{x} = 1$.

$\square$

***Theorem:* 3.4.3: $\mathbb{R}$ satisfies distributive law.**

The set of real numbers satisfy the distributive law. Let $x, y, z \in \mathbb{R}$.

(D) x(y + z) = xy + xz

Proof. Let $x = A|B, y = C|D, z = E|F$ be positive cuts. We know, $x(y + z)$ is the union of the set of all rationals $a(c + e)$ such that $a \in A, c \in C, e \in E, a > 0, c > 0, e > 0$ with $\{r \in \mathbb{Q} : r \leq 0\}$. In the rational numbers, $a(c + e) = ac + ae \in xy + xz$. Hence, $x(y + z) = xy + xz$. □

The set of real numbers $\mathbb{R}$ is a **complete, ordered field**.

We associate with each $r \in \mathbb{Q}$, a rational cut r^* which consists of all $p \in \mathbb{Q}$, such that $p < r$. It is clear that r^* is a cut. Cut arithmetic is consistent with $\mathbb{Q}$-arithmetic. That is, if we replace rational numbers r by the corresponding rational cuts $r* \in \mathbb{R}$, it preserves sums, products and order. So, the ordered field $\mathbb{Q}$ is isomorphic to the ordered field $\mathbb{Q}^*$ whose elements are cuts. By definition, this is what we mean, when we say that $\mathbb{Q}$ is a subfield of $\mathbb{R}$.

Suppose we try the same cut construction in $\mathbb{R}$ that we did in $\mathbb{Q}$. Are there gaps in $\mathbb{R}$ that can be detecting by cutting $\mathbb{R}$ with scissors? The natural definition of a cut in $\mathbb{R}$. The natural definition of a cut in $\mathbb{R}$ is a division $\mathcal{A}|\mathcal{B}$ where $\mathcal{A}$ and $\mathcal{B}$ where $\mathcal{A}$ and $\mathcal{B}$ are disjoint, non-empty subcollections of $\mathbb{R}$ with $\mathcal{A} \cup \mathcal{B} = \mathbb{R}$, and $a < b$ for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$. Further, $\mathcal{A}$ contains no largest element. Each b is an upper bound for $\mathcal{A}$. Therefore, $\sup \mathcal{A}$ exists and $a \leq y \leq b$ for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$.

3.5 Applications of Supremum and Infimum.

Lemma 3.5.1. *Assume $s \in \mathbf{R}$ is an upper bound for the set $S \subseteq \mathbf{R}$. Then, $s = \sup S$ if and only if, for every choice of $\epsilon > 0$, there exists an element $s_\epsilon \in S$ satisfying $s - \epsilon < s_\epsilon$.*

Proof. Here is a short rephrasing of this lemma: Given that s is an upper bound, s is the least upper bound if and only if any number small than s is not an upper bound. Putting it this way, qualifies as a proof, but we will expand on what is being said in each direction.

($\rightarrow$) For the forward direction, we assume $s = \sup S$ and consider $s - \epsilon$, where $\epsilon > 0$ has been arbitrarily chosen. Because, $s - \epsilon < s$, part (ii) of the definition (3.6) for the supremum of a set implies $s - \epsilon$ cannot be an upper bound for S. If this is the case, then there must be some number in S, let's call it s_ϵ which is greater than $s - \epsilon$ (because otherwise $s - \epsilon$ must be an upper bound). This proves the lemma in one direction.

($\leftarrow$) Conversely, assume s is an upper bound for S, with the property that no matter how $\epsilon > 0$, $s - \epsilon$ is no longer an upper bound for S. Notice, that what this implies is that if b is any other number less than s, b is not an upper bound. We put $\epsilon = s - b$. $[s - \epsilon = s - (s - b) = \epsilon]$. To prove that, $s = \sup A$, we must prove part (II) of the definition (3.6). Because we have just argued that any number smaller than s cannot be an upper bound, it follows that if b is some other upper bound for S, $b \geq s$. □

Problem 3.1. (a) Write the formal definition in the style of definition (3.6) for the infimum or the greatest lower bound of a set. (b) Now, state and prove a version of the lemma (3.5.1) for greatest lower bounds.

Solution.

(a) A real number l is said to be the infimum for a set $S \subseteq \mathbf{R}$ if it meets the following two criteria:

1) l is the lower bound for S. This implies $l \leq x$ for all $x \in S$.

2) If b any lower bound for S, $l \geq b$.

(b) Assume $l \in \mathbf{R}$ is an lower bound for the set $S \subseteq \mathbf{R}$. Then, $l = \inf S$ if and only if for every choice of $\epsilon > 0$, there exists $l_\epsilon \in S$ such that $l_\epsilon < l + \epsilon$.

Proof.

($\rightarrow$) In the forward direction, we assume $l = \inf S$. If we choose an arbitrary small but fixed $\epsilon > 0$, we have $l < l + \epsilon$. From the part (ii) of the definition of infimum, we can infer that $l + \epsilon$ is not the lower bound for S. So, there must exist a number in S smaller than $l + e$, call it, l_ϵ.

($\leftarrow$) Conversely, assume that l is a lower bound for S. Given that no matter what $\epsilon > 0$ is chosen, $l + \epsilon$ is not a lower bound for S. In other words, if b is a number greater than l, b is not a lower bound. We put $\epsilon = b - l$. To prove that, $l = \inf S$, we must prove part (ii) of the definition of an infimum. Because, we have just argued that any number greater than l cannot be a lower bound, it follows that if b is some lower bound, $b \leq l$. Hence, $l = \inf S$.

Problem 3.2. Let $A \subset \mathbb{R}$ be a non-empty subset. Define $-A = \{x : -x \in A\}$. Show that

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

Proof. (1)(a) Suppose u is an upper bound for $-A$.

Then, $u \geq y$ for all $y \in -A$. There is no other $y \in -A$, except those of the form $y = -x$ where $x \in A$. So, $u \geq -x$ for all $x \in A$. Therefore, $-u \leq x$ for all $x \in A$. Therefore, $-u$ is a lower bound for A.

(b) Let v be an arbitrary upper bound for $-A$. Then, $-v \leq x$ for all $x \in A$. That is $-v$ is a lower bound for A. Let's assume that u is the supremum of $-A$. Then, $u \leq v$ for all choices of v. Therefore, $-v \leq -u$ for all $-v$. Hence, $-u$ is an infimum for A.

Hence, $\sup -A = -\inf A$.

(2)(a) Suppose L is an lower bound for $-A$.

Then, $l \leq y$ for all $y \in -A$. There is no other $y \in -A$, except those of the form $y = -x$ where $x \in A$. So, $l \leq -x$ for all $x \in A$. Therefore, $-l \geq x$ for all $x \in A$. Therefore, $-l$ is an upper bound for A.

(b) Let m be an arbitrary lower bound for $-A$. Then, $-m \geq x$ for all $x \in A$. That is $-m$ is an upper bound for A. Let's assume that l is the infimum of $-A$. Then, $m \leq l$ for all choices of m. Therefore, $-m \geq -l$ for all $-m$. Hence, $-l$ is a supremum for A.

Hence, $\inf -A = -\sup A$.

□

Problem 3.3. Let $A, B \subset \mathbb{R}$ be non-empty. Define

$$A + B = \{z = x + y : x \in A, y \in B\},$$
$$A - B = \{z = x - y : x \in A, y \in B\}$$

Show that

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

Establish analogous formulas for $\inf(A + B)$ and $\inf(A - B)$.

Solution.

(1a) Let $a = \sup A$, $b = \sup B$. By definition, $x \leq a$ for all $x \in A$. And $y \leq b$ for all $y \in B$. There are no z in $A+B$ other than those of the form, $z = x+y$. Therefore, $z = x+y \leq a+b$. Hence, $a + b$ is an upper bound for the set $A + B$.

(1b) Let's choose an arbitrary small but fixed positive number $\epsilon > 0$. We would like to show any number smaller than $a + b$ cannot be an upper bound. Consider $a + b - \epsilon$.

$$a + b - \epsilon = (a - \frac{\epsilon}{2}) + (b - \frac{\epsilon}{2})$$

Since, $a - \frac{\epsilon}{2} < a$ it cannot be an upper bound for A. Thus, there exists some $a' \in A$, such that $a' > a - \frac{\epsilon}{2}$. Similarly, there exists $b' \in B$ such that $b'b - \frac{\epsilon}{2}$. Thus, $a' + b'$ is member of $A + B$ and $a' + b' > a + b - \epsilon$. So, $a + b - \epsilon$ cannot be an upper bound for $A + B$.

Thus, $\sup A + \sup B = \sup(A + B)$

(2a) Let $a = \sup A$ and $b = \inf B$.

We have, $a \geq x$ for all $x \in A$. Moreover $b \leq y$ for all $y \in B$. So, $-b \geq -y$ for all $y \in B$. We have, $a - b \geq x - b \geq x + (-y) = z$. There are no other elements z in $A - B$ other than those of the form $z = x + (-y)$. So, $a - b \geq z$ for all $z \in A - B$. Thus, $a - b$ is an upper bound for the set $A - B$.

(2b) Let's choose an arbitrary small but fixed positive number $\epsilon > 0$. We would like to show any number smaller than $a - b$ cannot be an upper bound. Consider $a - b - \epsilon$.

$$a + b - \epsilon = (a - \frac{\epsilon}{2}) - (b + \frac{\epsilon}{2})$$

Since, $a - \frac{\epsilon}{2} < a$ it cannot be an upper bound for A. There exists some number a' in A, that is greater than $a - \frac{\epsilon}{2}$. Similarly, $b + \frac{\epsilon}{2} > b$. Therefore, $b + \frac{\epsilon}{2}$ cannot be a lower bound. $\exists b'$ such that $b' < b + \epsilon/2$ or $-b' > -(b + \epsilon/2)$. So, there exists $a' - b' \in A - B$ such that $a' - b' > a - b - \epsilon$. Hence, $a - b - \epsilon$ cannot be an upper bound for $A - B$.

Thus, $\sup A - \inf B = \sup(A - B)$

Analogously,

$$\begin{aligned}\inf(A + B) &= \inf A + \inf B \\ \inf(A - B) &= \inf A - \sup B\end{aligned}$$

Problem 3.4. (a) Let A be bounded below, and define $B = \{b \in \mathbf{R} : b \text{ is a lower bound for } A\}$. Show that $\sup B = \inf A$. (b) Use (a) to explain why there is no need to assert that greatest lower bounds exist as part of the Axiom of Completeness. (c) Propose another way to use the Axiom of Completeness to prove that sets bounded below have greatest lower bounds.

Proof. (a) B is the set of all lower bounds for A. By definition, if b is any lower bound for A, $b \leq \inf A$. Hence, $\inf A$ is an upper bound for B.

Moreover, $A := \{a : a \geq b, b \in B\}$. That is, A is the set of all upper bounds for B. $\inf A \leq a$ for all $a \in A$. Hence, $\inf A$ is the least among all upper bounds for B.

Consequently, $\inf A = \sup B$.

(b) There is no need to assert that greatest lower bounds exist as part of the axiom completeness. For any subset B of the real numbers $\mathbb{R}$ that is bounded above, we can define A as the rest of $\mathbb{R}$, that is $A := \mathbb{R} - B$. From (a) it follows, that any subset of real numbers bounded below has a greatest lower bound that equals $\sup B$.

Problem 3.5. Assume that A and B are non-empty, bounded above and satisfy $B \subseteq A$. Show that $\sup B \leq \sup A$.

Proof. B is a subset of A. All elements of the set B also belong to A. Mathematically, $(x \in A) \implies (x \in B)$.

By part(i) of the definition (3.6) of the supremum of a set, $x \leq \sup A$ for all $x \in A$. Since, $B \subseteq A$, $x \leq \sup A$ for all x in A, contained in B. Hence, $\sup A$ is an upper bound for B.

By part(i) of the definition (3.6) of the supremum of a set, if u is any upper bound for the set B, $\sup B \leq u$. Take $u = \sup A$. Then $\sup B \leq \sup A$.

This closes the proof. □

Problem 3.6. Let $A \subseteq \mathbf{R}$ be bounded above, and let $c \in \mathbf{R}$. Define the sets $c + A$ and cA by $c + A = \{c + a : a \in A\}$ and $cA = \{ca : a \in A\}$.

(a) Show that $\sup(c + A) = c + \sup A$.

(b) If $c \geq 0$, show that $\sup cA = c \sup A$.

Solution.

(a) To properly verify this, we focus separately on each part of the definition (3.6) of the supremum of a set. We know that, $a \leq \sup A$ for all elements $a \in A$. Therefore, $c + a \leq c + \sup A$. The set $c + A$ has no other elements, except those of the form $c + a, a \in A$. Hence, $c + \sup A$ is an upper bound for $c + A$. Moreover, I would like to show that any number small than $c + \sup A$ is not an upper bound for $c + A$. Let $\epsilon > 0$ be a arbitrary small but fixed positive real number. Then, $\sup A - \epsilon < \sup A$. So, $\sup A - \epsilon$ is not an upper bound for A. Hence, there exists $n_\epsilon \in A$, such that $\sup A - \epsilon < n_\epsilon$. Thus, $c + \sup A - \epsilon < c + n_\epsilon$. And $c + n_\epsilon \in c + A$. Therefore, $c + \sup A - \epsilon$ is not an upper bound for $c + A$. Hence, $\sup(c + A) = c + \sup A$.

(b) $a \leq \sup A$ for all members a belonging to A. Therefore, $ca \leq c \sup A$ for all ca belonging to cA. $c \sup A$ is an upper bound for cA. Let ϵ/c be a arbitrary small but fixed positive number, $c \neq 0$. We have, $\sup A - \frac{\epsilon}{c} < \sup A$. So, $\sup A - \frac{\epsilon}{c}$ is not an upper bound for A. There exists a number in A, say, n_ϵ such that $\sup A - \frac{\epsilon}{c} < n_\epsilon$. Therefore, $c \sup A - \epsilon < cn_\epsilon$. Thus, $c \sup A - \epsilon$ is not an upper bound for cA. Consequently, $\sup(cA) = c \sup A$.

Problem 3.7. Give an example of each of the following or state that the request is impossible.

(a) A set B with $\inf B \geq \sup B$.

(b) A finite set that contains it's infimum but not it's supremum.

(c) A bounded subset of $\mathbf{Q}$ that contains its supremum but not its infimum.

Solution. (a) It is impossible for a set B to satisfy $\inf B > \sup B$, because of the order $<$ in $\mathbf{R}$. $\inf B \leq b \leq \sup B$ for all $b \in B$. $\inf B = \sup B$ when we have a singleton set, for example, $\{0\}, \{1\}$ etc.

(b) This is impossible. Every finite set must contain both its infimum and supremum.

(c) Consider the set $A := \{r \in \mathbf{Q} : (r > 0) \wedge (r \leq 3)\}$. This is a bounded subset, that contains its supremum but not its infimum.

Problem 3.8. Let $A_1, A_2, A_3, \ldots$ be a collection of non-empty sets, each of which is bounded above.

(a) Find a formula for $\sup(A_1 \cup A_2)$. Extend this to $\sup(\cup_{k=1}^n A_k)$.

(b) Consider $\sup(\cup_{k=1}^{n} A_k)$. Does this formula extend to the infinite case?

Proof. (a) Claim: $A_1 \cup A_2$ is bounded above.

We know that, $A_1 \cup A_2 := \{x : (x \in A_1) \vee (x \in A_2)\}$. There exists upper bounds u_i, such that $x \leq u_i (\forall x \in A_i)$, $i = 1, 2$. By the trichotomy property, atleast one $u_1 < u_2, u_1 = u_2, u_1 > u_2$ holds. Let us assume, for simplicity, $u_2 > u_1$. Then, $x \leq u_1 < u_2$ for all $x \in A_1$ and $x \leq u_2$ for all $x \in A_2$. Thus, $x \leq \max(u_1, u_2)$ for all elements $x \in A_1 \cup A_2$.

Claim: $m = \max\{\sup A_1, \sup A_2\}$ is an upper bound for $A_1 \cup A_2$.

We can argue as above m is an upper bound for $A_1 \cup A_2$. This verifies part (i) of the definition (3.6) of the supremum of a set.

Claim: If v is any other upper bound of $A_1 \cup A_2$, then $m \leq v$.

Let v be any other upper bound for $A_1 \cup A_2$. Then, $x \leq v$ for all elements $x \in (A_1 \cup A_2)$. This implies two things : $x \leq v$ for all elements x in A_1 and $x \leq v$ for all elements $x \in A_2$. So, v an upper bound for A_1 and A_2. But, we know that, $\sup A_1 \leq v$. And $\sup A_2 \leq v$. Hence, $m \leq v$.

This closes the proof.

In general, we can extend this formula to n terms. $\sup(\cup_{k=1}^{n} A_k) = \max\limits_{1 \leq k \leq n} \{\sup A_k\}$.

(b) The supremum of any union is the supremum of the suprema of those sets. Even for uncountbale unions!

$\square$

3.6 Consequences of Completeness.

The first application of the axiom of completeness is a result that may look like a more natural way to mathematically express the sentiment that the real line contains no gaps.

Theorem: 3.6.1: Nested Interval Property.

For each $n \in \mathbf{N}$, assume we are given a closed interval $I_n = [a_n, b_n] = \{x \in \mathbf{R} : a_n \leq x \leq b_n\}$. Assume also that each I_n contains I_{n+1}. Then, the resulting nested sequence of closed intervals

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq I_4 \ldots$$

has a nonempty intersection; that is

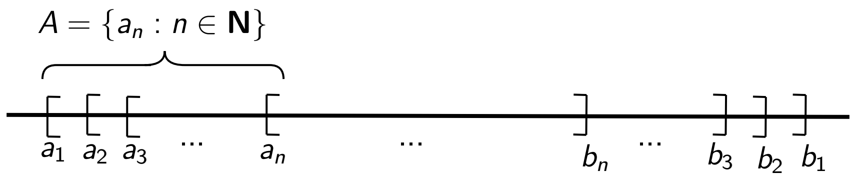
$$\bigcap_{n=1}^{\infty} I_n \neq \phi$$

Proof. In order to show that $\bigcup_{n=1}^{\infty} I_n$ is not empty, we are going to use the Axiom of Completeness (AoC) to produce a single real number x satisfying $x \in I_n$ for every $n \in \mathbf{N}$.

Now, AoC is a statement about bounded sets, and the one we want to consider is the set

$$A = \{a_n : n \in \mathbb{N}\}$$

of left-hand endpoints of the intervals.



Because the intervals are nested, we see that every b_n serves as an upper bound for A. Thus, we are justified in setting

$$x = \sup A$$

Now, consider a particular $I_n = [a_n, b_n]$. Because x is an upper bound for A, we have $a_n \leq x$. The fact that each b_n is an upper bound for A and that x is the least upper bound implies that $x \leq b_n$.

Altogether then, we have $a_n \leq x \leq b_n$, which means that $x \in I_n$ for every choice of $n \in \mathbb{N}$. Hence, $x \in \bigcap_{n=1}^{\infty} I_n$, and the intersection is not empty. □

3.6.1 The density of **Q** in **R**.

The set **Q** is an extension of **N**, and **R** in turn is an extension of **Q**. The next few results indicate how **N** and **Q** sit inside of **R**.

Theorem: 3.6.2: Archimedean Property.

(i) Given any number $x \in \mathbf{R}$, there exists $n \in \mathbf{N}$, satisfying $n > x$.
(ii) Given any real number $y > 0$, there exists an $n \in \mathbf{N}$ satisfying $1/n < y$.

Proof. Part(i) of the proposition states that $\mathbb{N}$ is not bounded above. There has never been any doubt about the truth of this, and it could be reasonably argued that we should not have to prove it at all.

The counterargument is that there is still a great deal of mystery about what the real numbers actually are. What we have said so far, is that **R** is an extension of **Q**, that maintains algebraic and order properties of the rationals but also possesses the least upper bound property articulated in the Axiom of Completeness. In the absence of any other information about $\mathbf{R}$, we have to consider the possibility that in extending $\mathbf{Q}$, we unwittingly acquired some new numbers which are upper bound for $\mathbf{N}$. This theorem asserts that real numbers do not contain such exotic creatures. The axiom of completeness, which we adopted to patch up the holes in $\mathbf{Q}$, carries with it the implication that $\mathbf{N}$ is an unbounded subset of $\mathbf{R}$.

And so to the proof. Assume, for contradiction that $\mathbf{N}$ is bounded above. By the Axiom of Completeness (AoC), $\mathbf{N}$ should have a least upper bound and we can set $\alpha = \sup \mathbf{N}$. If we consider $\alpha - 1$, we no longer have an upper bound and therefore there exists an $n \in \mathbf{N}$ satisfying $\alpha - 1 < n$. But, this is equivalent to $\alpha < n+1$. Because, $n+1$ is a natural number, $n + 1 \in \mathbf{N}$, we have a contradiction to the fact that α is supposed to be an upper bound for $\mathbf{N}$. (Note that the contradiction here depends only on AoC and the fact that $\mathbf{N}$ is closed under addition.)

Part (ii) follows from (i) by letting $x = \frac{1}{y}$. □

This familiar property of $\mathbf{N}$ is the key to an extremely important fact about how $\mathbf{Q}$ fits inside of $\mathbf{R}$.

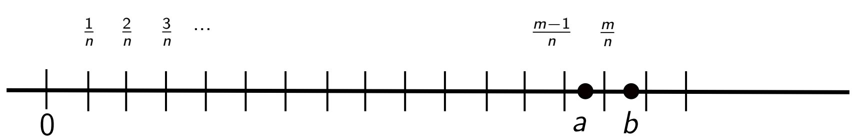
***Theorem:* 3.6.3: Density of $\mathbf{Q}$ in $\mathbf{R}$.**

For every two real numbers a and b with $a < b$, there exists a rational number r satisfying $a < r < b$.

Proof. A rational number is a quotient of integers, so we must produce $m \in \mathbf{Z}$ and $n \in \mathbf{N}$ so that

$$a < \frac{m}{n} < b \tag{3.1}$$

The first step is to choose the denominator n large enough so that consecutive increments of size $1/n$ are too close together to step over the interval (a, b).



Using the archimedean property (Theorem 3.6.1), we may pick $n \in \mathbf{N}$ large enough so that

$$\frac{1}{n} < b - a \tag{3.2}$$

Inequality (1) which we are trying to prove is equivalent to $na < m < nb$. With n already chosen, the idea now is to choose m to be the smallest integer greater than na. In other words, pick $m \in \mathbf{Z}$ so that

$$m - 1 \overset{(a)}{\leq} na \overset{(b)}{<} m \tag{3.3}$$

Now, inequality (b) immediately yields $a < \frac{m}{n}$, which is half of the battle. Keeping in mind that inequality (3.2) is equivalent to $a < b - 1/n$, we can use inequality (3.3) to write:

$$\begin{aligned} m &\leq na + 1 \\ &< n(b - \frac{1}{n}) + 1 \\ &= nb \end{aligned} \tag{3.4}$$

Because, $m < nb$ implies that $\frac{m}{n} < b$, we have $a < \frac{m}{n} < b$. □

Theorem 3.6.1 is paraphrased by saying that $\mathbf{Q}$ is *dense* in $\mathbf{R}$. Without working too hard, we can use this result to show that irrational numbers are also dense in $\mathbf{R}$.

3.6.2 The existence of square roots.

***Theorem:* 3.6.4: The existence of square roots.**

There exists a real number $\alpha \in \mathbf{R}$ satisfying $\alpha^2 = 2$.

Proof. Consider the set

$$T = \{t \in \mathbf{R} : t^2 < 2\} \tag{3.5}$$

and set $\alpha = \sup T$. We are going to prove that $\alpha^2 = 2$ by ruling out the possibilities $\alpha^2 < 2$ and $\alpha^2 > 2$. Keep in mind that there are two parts to the definition of $\sup T$, and they will both be important. The strategy is to demonstrate that $\alpha^2 < 2$ violates the fact that α is an upper bound for T, and $\alpha^2 > 2$ violates the fact that it is the least upper bound.

Let's first see what happens when $\alpha^2 < 2$. We choose β such that,

$$\beta = \alpha - \frac{\alpha^2 - 2}{\alpha + 2}$$

Then, if $\alpha^2 < 2$, $\beta > \alpha$. Moreover:

$$\begin{aligned}\beta &= \frac{\alpha(\alpha+2) - (\alpha^2 - 2)}{\alpha+2} \\ &= \frac{\alpha^2 + 2\alpha - \alpha^2 + 2}{\alpha+2} \\ &= \frac{2\alpha+2}{\alpha+2}\end{aligned}$$

We find that:

$$\begin{aligned}\beta^2 &= \frac{4(\alpha+1)^2}{(\alpha+2)^2} \\ \beta^2 - 2 &= \frac{4(\alpha+1)^2}{(\alpha+2)^2} - 2 \\ &= \frac{4(\alpha+1)^2 - 2(\alpha+2)^2}{(\alpha+2)^2} \\ &= \frac{4(\alpha^2+2\alpha+1) - 2(\alpha^2+4\alpha+4)}{(\alpha+2)^2} \\ &= \frac{4\alpha^2 + 8\alpha + 4 - 2\alpha^2 - 8\alpha - 8}{(\alpha+2)^2} \\ &= \frac{2\alpha^2 - 4}{(\alpha+2)^2} \\ &= \frac{2(\alpha^2-2)}{(\alpha+2)^2}\end{aligned}$$

If $\alpha^2 - 2 < 0$, we must have $\beta^2 - 2 < 0$. So, $\beta^2 < 2$. Thus, $\beta \in T$. Thus, there exists $\beta \in T$, such that $\beta > \alpha$. So, α is not an upper bound for T. This is a contradiction.

On the other hand, if $\alpha^2 > 2$, that is $\alpha^2 - 2 > 0$, we must have $\beta^2 - 2 > 0$, so $\beta^2 > 2$ and $\beta < \alpha$. Therefore, α cannot be the least upper bound of S. This is a contradiction.

Thus, the only possibility is $\alpha^2 = 2$. □

A small modification of this proof can be made to show that $\sqrt{x}$ exists for any $x \geq 0$.

Problem 3.9. (a) Prove that $\sqrt{3}$ is irrational. Does a similar argument work to show that $\sqrt{6}$ is irrational?

(b) Where does this proof break down if we try to use it to prove $\sqrt{4}$ is irrational?

(c) Prove that $\sqrt{12}$ is irrational.

Solution.

(a) By contradiction, let's assume that there exists a positive rational number p/q such that

$$\frac{p}{q} = \sqrt{3}$$

where p, q are positive integers.

$$\left(\frac{p}{q}\right)^2 = 3$$
$$p^2 = 3q^2$$

Observe that, $q^2 < p^2$, implying $q < p$. So, p^2 is a multiple of 3. Consequently, p is a multiple of 3. p must be a multiple of 3, for a) if $p = 3k+1$, $p^2 = 9k^2+6k+1 = 3(3k^2+2k)+1$ and b) if $p = 3k+2$, then $p^2 = (3k+2)^2 = 9k^2+12k+4 = 9k^2+12k+3+1 = 3(3k^2+4k+1)+1$. So, the only choice is $p = 3k$. Thus,

$$(3k)^2 = 3q^2$$
$$9k^2 = 3q^2$$
$$q^2 = 3k^2$$

Let $p' = q, q' = k$. Thus, we have passed from one solution (p, q) to another solution (p', q') of the same equation with a smaller value of p. We can repeat the procedure again and again, obtaining a sequence $(p', q'), (p'', q''), (p''', q'''), ...$ of solutions to $p^2 = 3q^2$ each with a smaller value of p than the previous one, and each one consisting of positive integers p, q. But, this contradicts the principle of infinite descent. This contradiction, shows that our initial assumption was wrong. We could not have had a rational number p/q, such that $\frac{p}{q} = \sqrt{3}$.

The same ideas can be used to prove that $\sqrt{6}$ is irrational.

(b) Suppose there is a positive rational number $r = \frac{p}{q}$ such that,

$$\frac{p}{q} = \sqrt{4}$$

and p, q have no common factors. Then,

$$p^2 = 4q^2$$

If p^2 is a multiple of 4, p must be a multiple of 2. Hence, $p = 2k$ and so $p^2 = 4k^2$.

$$\begin{aligned} 4k^2 &= 4q^2 \\ q^2 &= k^2 \\ q &= k \end{aligned}$$

Thus, (p, q) is the unique solution to the equation $p^2 = 4q^2$ and this fraction cannot be reduced further. Our assumption is correct, $\sqrt{4}$ is rational.

(c) Let x be a positive rational such that $x = \frac{p}{q}$ such that

$$x = \frac{p}{q} = \sqrt{12}$$

Then $q < p$ and,

$$\begin{aligned} \left(\frac{p}{q}\right)^2 &= 12 \\ p^2 &= 12q^2 \end{aligned}$$

p^2 is a multiple of 12. This means that p must be a multiple of 6. Let $p = 6k$. Therefore, $p^2 = 36k^2$.

$$\begin{aligned} 36k^2 &= 12q^2 \\ 4q^2 &= 12k^2 \\ (2q)^2 &= 12k^2 \\ \implies (p')^2 &= 12(q')^2 \end{aligned}$$

Here, $(p')^2 = 4q^2 < 12q^2 = p^2$. Thus, we passed from one solution (p, q) to another solution (p', q') of the same equation with a smaller value of p. We can repeat this procedure again and again obtaining a sequence $(p', q'), (p'', q''), (p''', q'''), ...$ of solutions to $p^2 = 2q^2$ each one with a smaller value of p than the previous one and each one consisting of positive integers. Again, this contradicts the principle of infinite descent.

Problem 3.10. Show that there is no rational number r, such that $2^r = 3$.

Proof. Let $r = \frac{p}{q}$ be a rational number such that $2^{\frac{p}{q}} = 3$. Then,

$$2^p = 3^q$$

This implies that 2^p is a integer multiple of 3. Therefore, 2 is an integer multiple of 3. But, this is a contradiction. This closes the proof. $\square$

Problem 3.11. Let $A \subseteq \mathbf{R}$ be nonempty and bounded above and let $s \in \mathbf{R}$ have the property that for all $n \in \mathbf{N}$, $s + \frac{1}{n}$ is an upper bound for A and $s - \frac{1}{n}$ is not an upper bound for A. Show that $s = \sup A$.

Proof. Let $s = \sup A$.

By the Archimedean property of reals, $\mathbf{N}$ sits inside $\mathbf{R}$. The statement for Archimedean property of real numbers says that:

For every real number $\epsilon > 0$, there exists a natural number n, such that $\frac{1}{n} < \epsilon$.

We proceed by contradiction.

(1) Claim: s is an upper bound for A.

Suppose that the above proposition is false. There exists $a \in A$ such that $a > s$. Then, $a - s > 0$. We are told that, $s + \frac{1}{n}$ is an upper bound for A for all natural numbers n. So, $a < s + \frac{1}{n}$. Consequently, $a - s < \frac{1}{n}$ for all $n \in \mathbf{N}$. But this is a contradiction. $(\exists n \in \mathbf{N})$ such that $\frac{1}{n} < a - s$.

(2) Claim: s is the least upper bound for A.

Suppose that the above proposition is false. Then, there exists an upper bound u for A, such that $u < s$. $u - s < 0 \implies s - u > 0$. We are told that $s - \frac{1}{n}$ is not an upper bound for A for all natural numbers $n \in \mathbf{N}$. Therefore, $s - \frac{1}{n} < u$ for all $n \in \mathbf{N}$. Consequently, $s - u < \frac{1}{n}$. This is a contradiction.

Therefore, $s = \sup A$ must be the least upper bound for A.

$\square$

Problem 3.12. Prove that $\bigcap_{n=1}^{\infty}(0, \frac{1}{n}) = \phi$. Notice that this demonstrates that the intervals in the nested interval property must be closed for the conclusion of the theorem to hold.

Proof. Let's assume that the intersection $\bigcup_{n=0}^{\infty}$ will be non-empty. Then there must be a real number $x \in \left(0, \frac{1}{n}\right)$ for all natural numbers $n \in \mathbf{N}$. Since, $\sup I_n = \frac{1}{n}$, we have :

$$(\forall n \in \mathbf{N})(\exists x > 0) \text{ such that } x < \frac{1}{n}$$

But this contradicts the Archimedean property of reals, which states that:

$$(\exists n \in \mathbf{N})(\forall x > 0) \text{ such that } \frac{1}{n} < x$$

Hence, our initial assumption is wrong. The countably infinite intersection is empty. $\square$

Problem 3.13. Let $a < b$ be real numbers and consider the set $T = \mathbf{Q} \cap [a, b]$. Show $\sup T = b$.

Proof. (1) Claim: b is an upper bound.

Let $t \in T$ be an arbitrary element in T. Then, $t \in \mathbf{Q}$ and $t \in [a, b]$. So, $t \leq b$ for all t. Consequently, b is an upper bound.

(2) Claim: b is the least upper bound.

Let $\epsilon > 0$ be an arbitrary small positive but fixed real number. $b - \epsilon < b$. Since, $\mathbf{Q}$ is dense in $\mathbf{R}$, we are guaranteed to find a rational number r smaller than b, i.e. $b - \epsilon < r < b$. Moreover, we can always find another rational number $s \in T$ such that, $r < s < b$. Hence, any rational r smaller than b is not an upper bound for T. Consequently, b is the least upper bound. $\square$

3.7 Cardinality.

At the moment, we think of $\mathbf{R}$ as consisting of rational and irrational numbers, continuously packed together along the real line. We have seen that both $\mathbf{Q}$ and $\mathbf{I}$ (the set of irrationals) are dense in $\mathbf{R}$, meaning that in every interval (a, b), there exist rational and irrational numbers alike. Mentally, there is a temptation to think of $\mathbf{Q}$ and $\mathbf{I}$ as being intricately mixed together in equal proportions, but this turns out not to be the case. In a way that George Cantor (1845-1918) made precise, the irrational numbers far outnumber the rational numbers in making up the real line.

3.7.1 1-1 Correspondence.

The term *cardinality* in mathematics is used to refer to a size of a set.

Definition 3.7. A function $f : X \to Y$ is said to be one-to-one (1-1) or injective, if and only if:

$$(x_1 \neq x_2) \implies f(x_1) \neq f(x_2) \tag{3.6}$$

Distinct inputs give distinct outputs.

Definition 3.8. A function $f : X \to Y$ is said to be onto or surjective, if and only if:

$$(\forall y \in Y)(\exists x \in X) \text{ such that } f(x) = y \tag{3.7}$$

The co-domain of the function equals the range of the function.

A function f that is both one-to-one and onto provides us with exactly what we mean by $1-1$ correspondence between two sets. The property of being 1-1 means that no two elements of X correspond to the same element of Y and the property of being onto ensures that every element of Y corresponds to something in A.

Definition 3.9. The set A *has the same cardinality* as B if there exists $f : A \to B$ that is 1-1 and onto. In this case, we write $A \sim B$.

Example 3.7.1. (i) If we let $E = \{2, 4, 6, ...\}$ be the set of even natural numbers, then we can show that $\mathbf{N} \sim E$. To see why, let $f : \mathbf{N} \to E$ be given by $f(n) = 2n$.

$$\begin{array}{lccccccc}
\mathbf{N}: & 1 & 2 & 3 & 4 & 5 & \cdots & n \\
 & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow \\
E: & 2 & 4 & 6 & 8 & 10 & & 2n
\end{array}$$

It is certainly true that E is a proper subset of $\mathbf{N}$, and for this reason it may seem logical to say that E is "smaller" set than $\mathbf{N}$. This is one way to look at it, but it represents a point of view that is heavily biased from an overexposure to finite sets. The definition of cardinality is quite specific, and from this point of view E and $\mathbf{N}$ are equivalent.

(ii) To make this point again, note that although $\mathbf{N}$ is contained in $\mathbf{Z}$ as a proper subset, we can show that $\mathbf{N} \sim \mathbf{Z}$. This time let

$$f(n) = \begin{cases} (n-1)/2 & \text{if } n \text{ is odd} \\ -n/2 & \text{if } n \text{ is even} \end{cases}$$

If we compute $f(n)$ for the first few natural numbers, we get $f(1) = (1-1)/2 = 0$, $f(2) = -2/2 = -1$, $f(3) = (3-1)/2 = 1$, $f(4) = -2, f(5) = 2, ...$ Let's verify that $f(n)$ is a bijection.

Claim: $f(n)$ is injective.

Case I. n_1, n_2 are odd.

$$
\begin{aligned}
f(n_1) &= f(n_2)\\
\frac{n_1 - 1}{2} &= \frac{n_2 - 1}{2}\\
n_1 - 1 &= n_2 - 1\\
n_1 = n_2
\end{aligned}
$$

Case II. n_1, n_2 are even.

$$
\begin{aligned}
f(n_1) &= f(n_2)\\
-\frac{n_1}{2} &= -\frac{n_2}{2}\\
n_1 &= n_2
\end{aligned}
$$

Hence, $f(n)$ is one-to-one or injective.

Claim: $f(n)$ is surjective. Let m be an arbitrary element belong to the set $\mathbf{Z}$. If we let $f(n) = m$, we find that:

Case I:

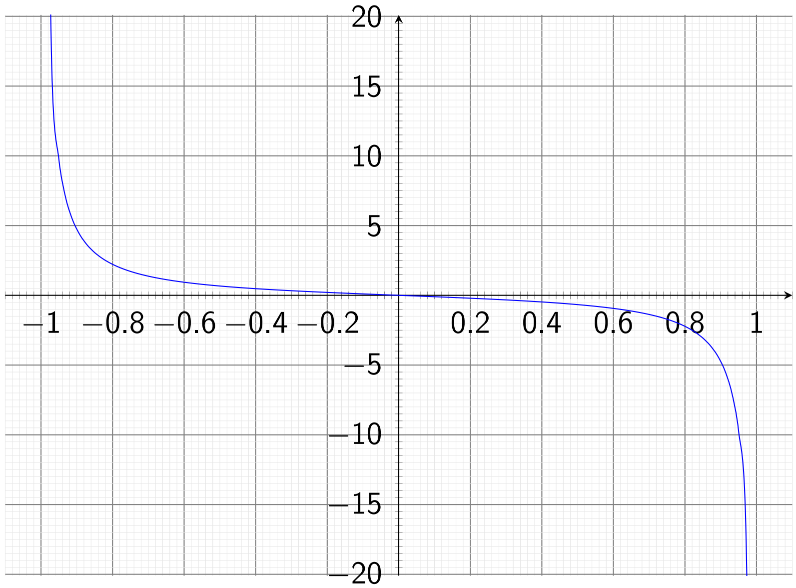
$$
\begin{aligned}
f(n) &= m\\
\frac{n-1}{2} &= m\\
n - 1 &= 2m\\
n &= 2m + 1
\end{aligned}
$$

Case II:

$$
\begin{aligned}
f(n) &= m\\
-\frac{n}{2} &= m\\
n &= -2m
\end{aligned}
$$

Thus, we can define the inverse mapping $g : \mathbf{Z} \to \mathbf{N}$ as

$$
g(m) := \begin{cases} 2m + 1 & \text{if } m \geq 0\\ -2m & \text{if } m < 0 \end{cases}
$$



$(-1,1) \sim \mathbf{R}$ using $f(x) = x/(x^2-1)$

Clearly, $\forall m \in \mathbf{Z}$, $\exists n \in \mathbf{N}$, such that $f(n) = m$. Hence, $f(n)$ is a surjection. Since, $f(n)$ is both injective and surjective, there is a $1-1$ correspondence between the natural numbers and the integers.

Example 3.7.2. A little Calculus shows that the function $f(x) = x/(x^2-1)$ takes the interval $(-1,1)$ onto $\mathbf{R}$ in a $1-1$ fashion. Thus $(-1,1) \sim \mathbf{R}$. In fact, $(a,b) \sim \mathbf{R}$ for any interval (a,b).

We see that both E and $\mathbf{Z}$ are countable sets. Putting the set into $1-1$ correspondence with $\mathbf{N}$, in effect, means putting all of the elements into an infinitely long list or sequence. From example 3.7.1, we can see that this was quite easy to do for E and required only a modest bit of shuffling for the set $\mathbf{Z}$. A natural question arises as to whether *all* infinite sets are countable. Given some infinite set such as $\mathbf{Q}$ or $\mathbf{R}$, it might seem as though with enough cleverness, we should be able to fit all elements of our set into a single list (i.e. into a correspondence with $\mathbf{N}$). After all, this list is infinitely long, so there must be plenty of room. But alas, as Hardy remarks, "[The mathematician's] subject is the most curious of all - there is none in which truth plays such odd pranks."

***Theorem:* 3.7.1: Countable and uncountable sets.**

(i) The set $\mathbf{Q}$ is countable. (ii) The set $\mathbf{R}$ is uncountable.

Proof. (i) Set $A_1 = \{0\}$ and for each $n \geq 2$, let A_n be the set given by:

$$\begin{array}{lccccccccccc} \mathbf{N}: & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ \mathbf{Q}: & 0 & \frac{1}{1} & -\frac{1}{1} & \frac{1}{2} & -\frac{1}{2} & \frac{2}{1} & -\frac{2}{1} & \frac{1}{3} & -\frac{1}{3} & \frac{3}{1} & -\frac{3}{1} \end{array}$$

$$\underbrace{0}_{A_1} \quad \underbrace{\tfrac{1}{1} \quad -\tfrac{1}{1}}_{A_2} \quad \underbrace{\tfrac{1}{2} \quad -\tfrac{1}{2} \quad \tfrac{2}{1} \quad -\tfrac{2}{1}}_{A_3} \quad \underbrace{\tfrac{1}{3} \quad -\tfrac{1}{3} \quad \tfrac{3}{1} \quad -\tfrac{3}{1}}_{A_4}$$

$$A_n = \{\pm\frac{p}{q} : \text{ where } p, q \text{ are the lowest terms with } p + q = n\}$$

The first few of these sets look like:

$$\begin{aligned} A_1 &= \{0\} \\ A_2 &= \{\frac{1}{1}\} \\ A_3 &= \{\frac{1}{2}, \frac{-1}{2}, \frac{2}{1}, \frac{-2}{1}\} \\ A_4 &= \{\frac{1}{3}, \frac{-1}{3}, \frac{3}{1}, \frac{-3}{1}\} \\ A_5 &= \{\frac{1}{4}, \frac{-1}{4}, \frac{2}{3}, \frac{-2}{3}, \frac{3}{2}, \frac{-3}{2}, \frac{4}{1}, \frac{-4}{1}\} \end{aligned}$$

The crucial observation is that A_n is *finite* and every rational number appears once and only once in exactly one of these sets. Our $1-1$ correspondence with $\mathbf{N}$ is then achieved by consecutively listing the elements in each A_n.

Admittedly, writing an explicit formula for this correspondence would be an awkward task, and attempting to do so is not the best use of time. What matters is that we see why every rational number appears in the correspondence exactly once. Given, say $\frac{22}{7}$, we have $\frac{22}{7} \in A_{29}$. Because the set of the elements in $A_1, A_2, \ldots, A_{29}$ is finite, we can be confident that $\frac{22}{7}$ eventually gets included in the sequence. The fact that this line of reasoning applies to any rational number p/q is our proof that the correspondence is onto. To verify that it is $1-1$, we observe that the sets A_n were constructed to be disjoint so that no rational number appears twice. This completes the proof of (i).

(ii) The second statement of this theorem is the truly unexpected part and it is proved by contradiction. Assume that there does exist a $1-1$, onto function $f : \mathbf{N} \to \mathbf{R}$. Again, what this suggests is that it is possible to enumerate the elements of $\mathbf{R}$. If we let $x_1 = f(1), x_2 = f(2)$, and so on, then our assumption that f is onto means that we can write:

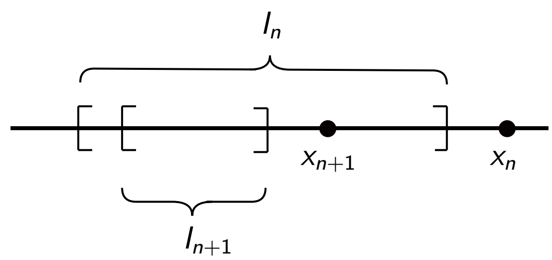
$$\mathbf{R} = \{x_1, x_2, x_3, x_4, \ldots\} \tag{1}$$

and be confident that every real number appears somewhere on the list. We will now use the Nested Interval property (Theorem 3.6) to produce a real number that is not there.

Let I_1 be a closed interval that *does not* contain x_1. Next, let I_2 be a closed interval, contained in I_1 that does not contain x_2. The existence of such an I_2 is easy to verify. Certainly, I_1 contains two smaller disjoint intervals and x_2 can be in only one of these. In general, given an interval I_n, construct I_{n+1} to satisfy

(i) $I_{n+1} \subseteq I_n$ and

(ii) $x_{n+1} \notin I_{n+1}$



We now consider the intersection $\bigcap_{n=1}^{\infty} I_n$. If x_{n_0} is some real number from the list in (1), then we have $x_{n_0} \notin I_{n_0}$, and it follows that :

$$x_{n_0} \notin \bigcap_{n=1}^{\infty} I_n$$

Now, we are assuming that the list in (1) contains every real number, and this leads to the conclusion that

$$\bigcap_{n=1}^{\infty} I_n = \phi$$

However, the Nested Interval Property (NIP) asserts that $\bigcap_{n=1}^{\infty} I_n \neq \phi$. By NIP, there is atleast one $x \in \bigcap_{n=1}^{\infty} I_n$ that, consequently, cannot be on the list in (1). This contradiction means that such an enumeration of **R** is impossible and we conclude that **R** is an *uncountable* set. $\square$

What exactly should we make of this discovery? It is an important exercise to show that any subset of a countable set must be either countable or finite. This should not be too surprising. If a set can be arranged into a single list, then deleting some elements from this list results in another(shorter and potentially) terminating list. This means that countable sets are the smallest type of infinite set. Anything smaller is either still countable or finite.

The force of theorem 3.7.1 is that the cardinality of $\mathbf{R}$ is, informally speaking, a larger type of infinity. The real numbers so outnumber the natural numbers that there is no way to map $\mathbf{N}$ onto $\mathbf{R}$. No matter how we attempt this, there are always real numbers to spare. The set $\mathbf{Q}$ on the other hand is countable. As far as infinite sets is concerned, this is as small as it gets. What does this imply about the set $\mathbf{I}$ of irrational numbers? By imitating the demonstration that $\mathbf{N} \sim \mathbf{Z}$, we can prove that the union of two countable sets must be countable. Because, $\mathbf{R} = \mathbf{Q} \cup \mathbf{I}$, it follows that $\mathbf{I}$ cannot be countable because otherwise $\mathbf{R}$ would be. The inescapable conclusion is that, despite the fact that, we have encountered so few of them, the irrational numbers form a far greater subset of $\mathbf{R}$ than $\mathbf{Q}$.

The properties of countable sets described in this discussion are useful for a few exercises in upcoming chapters. For easier reference, we state them as some final propositions and outline their proofs in the exercises that follow.

Theorem: 3.7.2: Subset of a countable set.

If $A \subseteq B$ and B is countable, then A is either countable or A is finite.

Proof. Assume that B is countable. Then the elements of B can be written in the form of a list:

$$B = \{b_1, b_2, b_3, b_4, \ldots\}$$

There exists an $f : \mathbf{N} \to B$ which is $1-1$ and onto. Let $A \subseteq B$ be an infinite subset of B. We must show that this A is countable.

Let $n_1 = \min I_1$ where $I_1 = \{n \in \mathbf{N} : f(n) \in A\}$. As a start to the definition of $g : \mathbf{N} \to A$, set $f(1) = f(n_1)$. We can inductively continue this process to produce a $1-1$ function g from $\mathbf{N}$ onto A.

$$\begin{aligned}
g(1) &= f(n_1) && \text{where } n_1 = \min I_1 \\
g(2) &= f(n_2) && \text{where } I_2 = I_1 \setminus \{n_1\},\ n_2 = \min I_2 \\
g(3) &= f(n_3) && \text{where } I_3 = I_2 \setminus \{n_2\},\ n_3 = \min I_3 \\
&\vdots \\
g(k+1) &= f(n_{k+1}) && \text{where } I_{k+1} = I_k \setminus \{n_k\},\ n_{k+1} = \min I_{k+1}
\end{aligned}$$

Moreover, we can show that $g : \mathbf{N} \to A$ is $1-1$. If $g(r) = g(s) \implies f(n_r) = f(n_s) \implies n_r = n_s \implies r = s$. Also, for all members $a \in A$, there exists a natural number $n \in \mathbf{N}$, such that $g(n) = a$. So, g is surjective. $\square$

Problem 3.14. Review the proof of theorem 3.7.1, part (ii) showing that $\mathbf{R}$ is uncountable, and then find the flaw in the erroenous proof that $\mathbf{Q}$ is uncountable:

Assume, for contradiction that $\mathbf{Q}$ is countable. Thus, we can write $\mathbf{Q} = \{r_1, r_2, r_3, ...\}$ and as before, construct a nested sequence of closed intervals with $r_n \notin I_n$. Our construction implies $\bigcap_{n=1}^{\infty} I_n = \phi$, while NIP implies that $\bigcap_{n=1}^{\infty} I_n \neq \phi$. This contradiction implies that $\mathbf{Q}$ must therefore be uncountable.

Proof. As these are intervals in $\mathbf{Q}$, NIP does not apply. □

Theorem: 3.7.3

(i) If $A_1, A_2, \ldots, A_m$ are each countable sets, then the union $A_1 \cup A_2 \cup \ldots A_m$ is countable.
(ii) If A_n is a countable set for each $n \in \mathbf{N}$, then $\bigcup_{n=1}^{\infty} A_n$ is countable.

Proof. (a) First, we prove □

Chapter 4

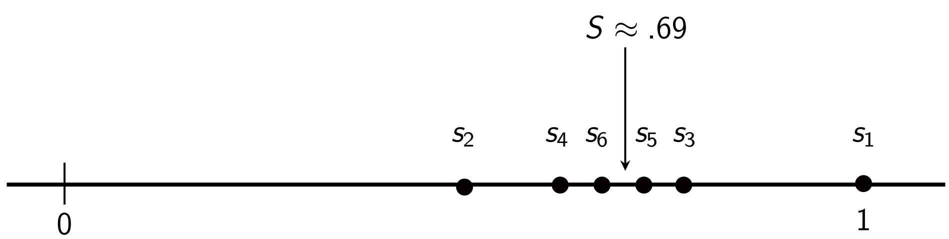
Sequences and Series.

4.1 Discussion: Rearrangement of Infinite series.

Consider the infinite series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \ldots \tag{4.1}$$

If we naively begin adding from the left hand side, we get a sequence of what are called *partial sums*. In other words, let s_n equal the sum of the first n terms of the series, so that $s_1 = 1, s_2 = \frac{1}{2}, s_3 = \frac{5}{6}, s_4 = \frac{7}{12}$, and so on. One immediate observation is that successive sums oscillate in a progressively narrower space. The odd sums decrease $(s_1 > s_3 > s_5 > s_7 > \ldots)$ while the even sums increase $(s_2 < s_4 < s_6 < s_8 < \ldots)$.



$$s_2 < s_4 < s_6 < \ldots < S < \ldots < s_5 < s_3 < s_1$$

It seems reasonable - and we will soon prove - that the sequence (s_n) eventually hones in on a value, call it S, where the odd and even partial sums **meet**. At this moment, we cannot compute S precisely, but we know that S falls somewhere between $5/6$ and $7/12$. Summing a few hundred terms reveals that $S \approx 0.69$. Whatever its value, there is an overwhelming temptation to write

$$S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \ldots \tag{1}$$

meaning, perhaps that if we could indeed add up all infinitely many of these numbers, then the sum would equal S. A more familiar example of an equation of this type might be:

$$2 = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \ldots$$

the only difference being that in the second equation we have a more recognizable value for the sum.

The crux of the matter is, the symbols $+$, $-$ and $=$ in the preceding equations are deceptively familiar notions being used in an unfamiliar way. The crucial question is whether or not the properties that are well understood for finite sums remain valid when applied to infinite objects such as equation (4.1). The answer as we are about to witness, is somewhat ambiguous.

Treating equation (4.1) in a standard algebraic way, let's multiply throughout by $1/2$ and add it back to equation (4.1):

$$\begin{array}{lllllll}
\frac{1}{2}S = & \frac{1}{2} & -\frac{1}{4} & +\frac{1}{6} & -\frac{1}{8} & +\frac{1}{10} & -\frac{1}{12} + \ldots \\
+S = 1 & -\frac{1}{2} + \frac{1}{3} & -\frac{1}{4} + \frac{1}{5} & -\frac{1}{6} + \frac{1}{7} & -\frac{1}{8} + \frac{1}{9} & -\frac{1}{10} + \frac{1}{11} & -\frac{1}{12} + \ldots \\
\hline
\frac{3}{2}S = 1 & +\frac{1}{3} & -\frac{1}{2} + \frac{1}{5} & +\frac{1}{7} & -\frac{1}{4} + \frac{1}{9} & +\frac{1}{11} & -\frac{1}{6} + \ldots
\end{array} \tag{2}$$

Now, look carefully at the result. The sum in equation (2) consists precisely of the same terms as those in the original equation (1), only in a different order. Specifically, the series in (2)is a rearrangement of the series in (1), where we list the first two positive terms $(1 + \frac{1}{3})$ followed by the first negative term $(-\frac{1}{2})$, followed by the next two positive terms $(\frac{1}{5} + \frac{1}{7})$ and then the next negative term $(-\frac{1}{4})$. Continuing this, it is apparent that every term that every term in (2) appears in (1) and vice versa. The rub comes when we realise that equation (2) asserts that the sum of these rearranged, but otherwise unaltered numbers is equal to $3/2$ its original value.Indeed, adding a few hundred terms of the equation (2) produces the partial sums in the neighborhood of 1.03. Addition, in this infinite setting, is not commutative!

Let's look at a similar rearrangement of the series

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

This series is geometric with first term 1 and common ratio $-\frac{1}{2}$. Using the formula $\frac{1}{1-r}$ for the sum of a geometric series, we get

$$1-\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\frac{1}{16}-\frac{1}{32}+\frac{1}{64}-\frac{1}{128}+\frac{1}{256}+\ldots=\frac{1}{1-\left(-\frac{1}{2}\right)}=\frac{2}{3}$$

This time, some computational experimentation with the two positives, one negative rearrangement:

$$1+\frac{1}{4}-\frac{1}{2}+\frac{1}{16}+\frac{1}{64}-\frac{1}{8}+\frac{1}{256}+\frac{1}{1024}-\frac{1}{32}\ldots$$

which yields partial sums quite close to $2/3$. The sum of the first 30 terms, for instance, equals 0.66667. Infinite addition is commutative in some instances but not in others.

Far from being a charming theoretical oddity of the infinite series, this phenomenon can be the source of great consternation in many applied situations. How, for instance, should a double summation over two index variables be defined? Let's say we are given a grid of real numbers $\{a_{ij} : i,j \in \mathbf{N}\}$, where $a_{ij} = 1/2^{j-i}$ if $j > i$, $a_{ij} = -1$ if $j = i$ and $a_{ij} = 0$ if $j < i$.

$$\begin{bmatrix} -1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} & \ldots \\ 0 & -1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \ldots \\ 0 & 0 & -1 & \frac{1}{2} & \frac{1}{4} & \ldots \\ 0 & 0 & 0 & -1 & \frac{1}{2} & \ldots \\ 0 & 0 & 0 & 0 & 1 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

We would like to attach a mathematical meaning to the summation

$$\sum_{i,j=1}^{\infty} a_{ij}$$

whereby we intend to include every term in the preceding array in the total. One natural idea is to temporarily fix i and sum across each row. A moment's reflection (and a fact about geometric series) shows that each row sums to 0. Summing the sums of the rows, we get

$$\sum_{i,j=1}^{\infty} a_{ij} = \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} a_{ij}\right) = \sum_{i=1}^{\infty}(0) = 0$$

We could have as easily have decided to fix j and sum down each column first. In this case we have

$$\sum_{i,j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty} a_{ij}\right) = \sum_{j=1}^{\infty}\left(\frac{-1}{2^{n-1}}\right) = -2$$

Changing the order of the summation changes the value of the sum! One common way that double sums arise (although not this particular one) is from the multiplication of two series. There is a natural desire to write

$$\left(\sum a_i\right)\left(\sum b_j\right) = \sum_{i,j} a_i b_j$$

except that the expression on the right hand side makes no sense at the moment.

It is the pathologies that give rise to the need for rigour. A satisfying resolution to the questions will require that we be absolutely precise about what we mean as we manipulate these infinite objects. It may seem that progress is slow at first, but that is because we do not want to fall in the trap of letting the biases of our intuition corrupt our arguments. Rigorous proofs are meant to be a check on intuition, and in the end we will see that they vastly improve our mental picture of the mathematical infinite.

As a final example, consider something intuitively fundamental as the associative property of addition applied to the series $\sum_{n=1}^{\infty}(-1)^n$. Grouping the terms one way gives :

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

whereas grouping in another yields

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

Manipulations that are legitimate in finite settings do not always extend to infinite settings. Deciding when they do and why they do not is one of the central themes of analysis.

4.2 The limit of a Sequence.

An understanding of infinite series depends heavily on a clear understanding of the theory of sequences. In fact, most of the concepts in analysis can be reduced to statements about the behavior of sequences. Thus, we will spend a significant amount of time investigating sequences, before taking on infinite series.

Definition 4.1. A sequence is a function whose domain is $\mathbf{N}$.

This formal definition leads immediately to the familiar depiction of a sequence as an ordered list of real numbers. Given a function $f : \mathbf{N} \to \mathbf{R}$, $f(n)$ is just the nth term on the list. The notation for sequences reinforces this familiar understanding.

Example 4.2.1. Each of the following are common ways to describe a sequence.

(i) $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots)$,

(ii) $\left(\frac{1+n}{n}\right)_{n=1}^{\infty} = \left(\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \ldots\right)$

(iii) (a_n), where $a_n = 2^n$ for each $n \in \mathbf{N}$

(iv) (a_n), where $x_1 = 2$ and $x_{n+1} = \frac{x_n+1}{2}$

On occasion, it will be more common to index a sequence beginning with $n = 0$ or $n = n_0$ for some natural number n_0 different from 1. These minor variations should cause no confusion. What is essential is that a sequence be an infinite list of real numbers. What happens at the beginning of such a list is of little importance in most cases. The business of analysis is concerned with the behavior of **the infinite tail** of a given sequence.

Definition 4.2. Convergence of a sequence. A sequence (a_n) converges to the real number a if, for every arbitrary small positive number $\epsilon > 0$, there exists $N \in \mathbf{N}$ such that whenever $n \geq N$, it follows that $|a_n - a| < \epsilon$.

To indicate that the sequence (a_n) converges to a, we usually either write either $\lim a_n = a$ or $(a_n) \to a$. The notation $\lim_{n\to\infty} a_n = a$ is also standard.

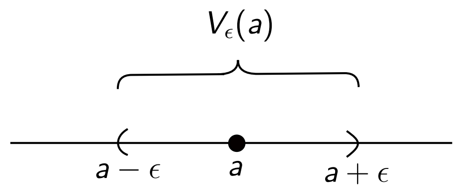
In an effort to decipher this definition, it helps first to consider the ending phrase $|a_n - a| < \epsilon$ and think about the points that satisfy an inequality of this type.

Definition 4.3. Given a real number $a \in \mathbf{R}$ and a positive number $\epsilon > 0$, the set

$$V_\epsilon(a) = \{x \in \mathbf{R} : |x - a| < \epsilon\}$$

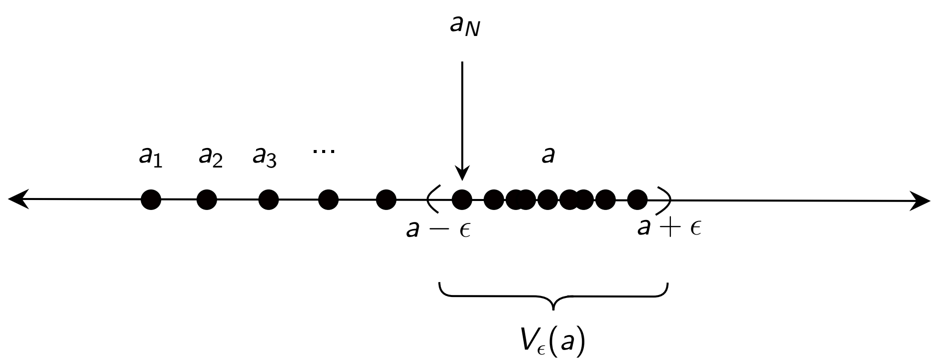
is called the ϵ-neighbourhood of a.

Notice that $V_\epsilon(a)$ consists of all those points whose distance from a is less than ϵ. Said another way, $V_\epsilon(a)$ is an interval, centered at a, with radius ϵ.



Recasting the definition of convergence in terms of ϵ-neighbourhoods gives a more geometric impression of what is being described.

Definition 4.4. Convergence of a sequence : Topological version. A sequence (a_n) converges to a, if given any ϵ-neighbourhood $V_\epsilon(a)$ of a, no matter how small, there exists a point in the sequence N, after which all of the terms are in $V_\epsilon(a)$. In other words, every ϵ-neighbourhood contains all but a finite number of terms of (a_n).



Definition 4.2 and 4.4 say precisely the same thing; the natural number N in the original version of the definition is the point where the sequence (a_n) enters $V_\epsilon(a)$, never to leave. It should be apparent that *the value of* N *depends upon the choice of* ϵ. The smaller the $\epsilon-$neighbourhood, the larger N may have to be.

Example 4.2.2. Consider the sequence (a_n), where $a_n = \frac{1}{\sqrt{n}}$. Our intuitive understanding of limits points confidently to the conclusion that :

$$\lim\left(\frac{1}{\sqrt{n}}\right) = 0$$

Before trying to prove this not too impressive fact, let's explore the relationship between ϵ and N in the definition of convergence. For the moment, take ϵ to be $1/10$. This sort of defines the *target zone* for the terms in the sequence. By claiming that the limit of (a_n) is 0, we are saying that the terms in this sequence eventually get **arbitrarily** close to 0. How close? What do we mean by *eventually*? We have set $\epsilon = 1/10$ as our standard for closeness, which leads to the $\epsilon-$neighbourhood $(-1/10, 1/10)$ centered around the limit 0. How far out into the sequence must we look before the terms fall into this interval? The 100th term $a_{100} = 1/10$ puts us right on the boundary, and a little thought reveals that

$$\text{if } n > 100, \text{ then } a_n \in \left(-\frac{1}{10}, \frac{1}{10}\right)$$

Thus, for $\epsilon = 1/10$, we choose $N = 101$ (or anything larger) as our response.

Now, our choice $\epsilon = 1/10$ was rather whimsical, and we can do this again, letting $\epsilon = 1/50$. In this case, our target neighbourhood shrinks to $(-1/50, 1/50)$, and it is apparent that we must travel further our into the sequence before a_n falls into this interval. How far? Essentially, we require that

$$\frac{1}{\sqrt{n}} < \frac{1}{50} \text{ which occurs as long as } n > 50^2 = 2500$$

Thus, $N = 2501$ is a suitable response to the challenge of $\epsilon = 1/50$.

It may seem as though this duel could continue forever, with different ϵ challenges being handed to us one after another, each one requiring a suitable value of N in response. In a sense, this is correct, except that the game is effectively over the instant we recognize a *rule* for how to choose N given an *arbitrary* $\epsilon > 0$. For this problem, the desired algorithm is implicit in the algebra to be carried out to compute the previous response of $N = 2501$. Whatever ϵ happens to be, we want

$$\frac{1}{\sqrt{N}} < \epsilon \text{ which is equivalent to insisting that } n > \frac{1}{\epsilon^2}$$

With this osbervation, we are ready to write the formal argument.

We claim that

$$\lim \left(\frac{1}{\sqrt{n}}\right) = 0$$

Proof. Let $\epsilon > 0$ be an arbitrary positive number. Choose a natural number N satisfying:

$$N > \frac{1}{\epsilon^2}$$

We now verify that this choice of N has the desired property. Let $n \geq N$. Then,

$$n > \frac{1}{\epsilon^2} \text{ implies } \frac{1}{\sqrt{n}} < \epsilon$$
$$\text{and hence } |a_n - 0| < |\epsilon - 0| = \epsilon$$

□

Example 4.2.3. Show

$$\lim \left(\frac{n+1}{n}\right) = 1$$

As mentioned earlier before beginning to attempt a formal proof, we first need to do some preliminary scratch work. In the first example, by experiment by assigning specific values to ϵ (and it is not a bad idea to do this again), but let us again skip straight to the algebraic punch line. The last line of our proof should be that for suitably large values of n,

$$|\frac{n+1}{n} - 1| < \epsilon$$

Because

$$|\frac{n+1}{n} - 1| = \frac{1}{n}$$

this equivalent to the inequality $1/n < \epsilon$ or $n > \frac{1}{\epsilon}$. Thus, choosing N to be an integer greater than $1/\epsilon$ will suffice. With the work of the proof done, all that remains is a formal writeup.

Proof. Let $\epsilon > 0$ be arbitrary. Choose $n \in \mathbf{N}$, so that $N > \frac{1}{\epsilon}$. To verify that this choice of N is appropriate, let $n \in \mathbf{N}$ satisfy $n \geq N$. Then, $n \geq N$ implies $n > \frac{1}{\epsilon}$, which is the same as saying $1/n < \epsilon$. Finally, this means that

$$|\frac{n+1}{n} - 1| < \epsilon$$

□

It is instructive to see in what goes wrong in our previous example if we try to prove that our existence converges to some limit other than 1.

Theorem: 4.2.1: Uniqueness of limits.

The limit of a sequence, when it exists must be unique.

Proof. □

4.2.1 Divergence.

Significant insight into the role of quantifiers in the definition of convergence can be gained by studying an example of a sequence that does not have a limit.

Example 4.2.4. Consider the sequence

$$\left(1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \frac{1}{5}, -\frac{1}{5}, \frac{1}{5}, -\frac{1}{5}, \frac{1}{5}, -\frac{1}{5}, \dots\right)$$

How can we argue that this sequence does not converge to zero? Looking at the first few terms, it seems the initial evidence actually supports such a conclusion. Given a challenge of $\epsilon = 1/2$, a little reflection reveals that after $N = 3$ all the terms fall into the neighbourhood $(-1/2, 1/2)$. We could also handle $\epsilon = 1/4$. (The smallest possible N in this case is 5).

But the definition of convergence says that $(\forall \epsilon > 0)$, and it should be apparent there is no response to a choice of $\epsilon = 1/10$, for instance. This leads us to an important observation about the logical negation of the definition of convergence of a sequence. To prove that a particular number x is *not* the limit of a sequence (x_n), we must produce a single value of ϵ for which no $N \in \mathbf{N}$ works. More generally speaking, the negation of a statement that begins with $(\forall x, P(x))$ is $(\exists x, \neg P(x))$. For instance, how could we disprove the spurious claim that "At every college in the United States, there is a student, who is atleast seven feet tall"? If there is atleast one college, such that no students are more than seven feet tall!

We have argued that the preceding sequence does not converge to 0. Let's argue against the claim that it converges to $1/5$. Choose $\epsilon = 1/10$ produces the neighbourhood $(1/10, 3/10)$. Although the sequence continually revisits the neighbourhood, there is no point at which it enters and never leaves as the definition requires. Thus, no N exists for $\epsilon = 1/10$, so the sequence does not converge to $1/5$.

Of course this sequence does not converge to any other real number, and it would be more satisfying to simply say that this sequence does not converge.

Definition 4.5. A sequence that does not converge is said to diverge.

Although it is not too difficult, we will postpone arguing divergence in general until we develop a more economical divergence criterion in a later section.

Problem 4.1. Verify using the definition of the convergence of a sequence, that the following sequences converge to the proposed limit.

(a) $\lim \frac{2n+1}{5n+4} = \frac{2}{5}$

(b) $\lim \frac{2n^2}{n^3+3} = 0$

(c) $\lim \frac{\sin(n^2)}{\sqrt[3]{n}} = 0$

Solution.

(a) The last line in our proof should be :

$$
\begin{aligned}
|\frac{2n+1}{5n+1} - \frac{2}{5}| &< \epsilon \\
\implies |\frac{10n+5-10n-2}{5(5n+1)}| &< \epsilon \\
\frac{3}{5(5n+1)} &< \epsilon \\
\frac{5(5n+1)}{3} &> \frac{1}{\epsilon} \\
5n+1 &> \frac{3}{5\epsilon} \\
5n &> \frac{3}{5\epsilon} - 1 \\
n &> \frac{3}{25\epsilon} - \frac{1}{5}
\end{aligned}
$$

We therefore choose $N \in \mathbf{N}$ to satisfy:

$$N > \frac{3}{25\epsilon} - \frac{1}{5}$$

Then, for all $n \geq N$, we have

$$|\frac{2n+1}{5n+1} - \frac{2}{5}| < \epsilon$$

Thus,

$$\lim \frac{2n+1}{5n+1} = \frac{2}{5}$$

(b) Let's choose $N > 2/\epsilon$.

Then, for all $n \geq N$,

$$n > \frac{2}{\epsilon}$$
$$\frac{n}{2} > \frac{1}{\epsilon} \implies \frac{2}{n} < \epsilon \implies \frac{2n^2}{n^3} < \epsilon$$
$$\text{Now, } \frac{2n^2}{n^3+3} < \frac{2n^2}{n^3} < \epsilon$$
$$\text{Therefore, } |\frac{2n^2}{n^3+3} - 0| < \epsilon$$
$$|a_n - 0| < \epsilon$$
$$\lim a_n = 0$$

(c) Let's choose $N > \frac{1}{\epsilon^3}$.

Then, for all $n \geq N$,

$$n > \frac{1}{\epsilon^3}$$
$$\frac{n}{2} > \frac{1}{\epsilon} \implies \frac{2}{n} < \epsilon \implies \frac{2n^2}{n^3} < \epsilon$$
$$\text{Now, } \frac{2n^2}{n^3+3} < \frac{2n^2}{n^3} < \epsilon$$
$$\text{Therefore, } |\frac{2n^2}{n^3+3} - 0| < \epsilon$$
$$|a_n - 0| < \epsilon$$
$$\lim a_n = 0$$

Problem 4.2. Describe what we would have to demonstrate in order to disprove each of the following statements.

(a) At every college in the United States, there is a student who is atleast 7 feet tall.

(b) For all colleges in the United States, there exists a professor who gives every student a grade of either A or B

(c) There exists a college in the United States where every student is atleast six feet tall.

Solution.

To disprove each of the statements, the negation of the statement must hold true.

(a) There is atleast one college in the US, where all students are less than 7 feet tall.

(b) There is atleast one college in the US, where all professors give atleast one student a grade which is neither A nor B.

(c) For all colleges in the United States, there is atleast one student that is not taller than six feet or more.

Problem 4.3. Give an example of each or state that the request is impossible. For any that are impossible, give a compelling argument for why that is the case.

(a) A sequence with an infinite number of ones that does not converge to one.

(b) A sequence with an infinite number of ones that converges to a limit not equal to one.

(c) A divergent sequence such that for every $n \in \mathbf{N}$, it is possible to find n consecutive ones somewhere in the sequence.

Solution.

(a) A sequence with an infinite number of ones may not always converge to one. Consider $a_n = (-1)^n$.

$$1, -1, 1, -1, 1, -1, 1, -1, \dots,$$

(b) This impossible.

(c) Consider the sequence (a_k) where $a_k = -1$ if $k = n(n+1)/2$ for all $n \in \mathbf{N}$ otherwise $a_k = 1$.

$$-1, 1, -1, 1, 1, -1, 1, 1, 1, -1, 1, 1, 1, 1, -1, 1, 1, 1, 1, 1, 1, -1 \dots,$$

Problem 4.4. Let $[[x]]$ be the greatest integer less than or equal to x. For example, $[[\pi]] = 3$ and $[[3]] = 3$. For each sequence, find $\lim a_n$, and verify it with the definition of convergence.

(a) $a_n = [[5/n]]$,

(b) $a_n = [[(12 + 4n)/3n]]$

Reflecting on these examples, comment on the statement following definition 4.2 that the smaller the ϵ-neighbourhood, the larger N may have to be.

Solution.

(a)

$$\begin{aligned}(a_n) &= [[5/1]], [[5/2]], [[5/3]], [[5/4]], [[5/5]], [[5/6]], [[5/7]], \dots \\ &= 5, 2, 1, 1, 1, 0, 0, 0, 0, \dots\end{aligned}$$

Thus, $\lim a_n = 0$.

We see that given any $\epsilon > 0$, for all $n \geq 6$, $|a_n - 0| < \epsilon$.

(b)

$$
\begin{aligned}
(a_n) &= [[16/3]], [[20/6]], [[24/9]], [[28/12]], [[32/15]], [[36/18]], [[40/21]], ... \\
&= 3, 3, 2, 2, 2, 2, 1, 1, ...
\end{aligned}
$$

$\lim a_n = 1$

We see that given any $\epsilon > 0$, for all $n > \frac{4}{\epsilon}$, we have:

$$
\begin{aligned}
n &> \frac{4}{\epsilon} \\
\frac{4}{n} &< \epsilon \\
\frac{4}{n} + \frac{4}{3} - \frac{4}{3} &< \epsilon \\
\frac{12+4n}{4n} - \frac{4}{3} &< \epsilon \\
|\frac{12+4n}{3n} - \frac{4}{3}| &< \epsilon \\
\lim \frac{12+4n}{3n} &= \frac{4}{3} \\
\lim \left[\left[\frac{12+4n}{3n}\right]\right] &= \left[\left[\frac{4}{3}\right]\right] \\
\lim \left[\left[\frac{12+4n}{3n}\right]\right] &= 1
\end{aligned}
$$

Problem 4.5. Prove that the limit of a sequence, when it exists, must be unique. To get started, assume $(a_n) \to a$ and also that the sequence $(a_n) \to b$. Now argue that $a = b$.

Proof. Assume that $\lim a_n = a$ and $\lim a_n = b$ and $a \neq b$.

Let $\epsilon_1 = \epsilon/2 > 0$ be any arbitrary small positive real number. For $\epsilon_1 > 0$, there exists $N(\epsilon_1)$ such that if $n \geq N(\epsilon_1)$, we have:

$$|a_n - a| < \frac{\epsilon}{2}$$

Similarly, let $\epsilon_2 = \epsilon/2 > 0$ be an arbitrary small positive real number. For $\epsilon_2 > 0$, there exists $N(\epsilon_2)$ such that if $n \geq N(\epsilon_2)$, we have:

$$|a_n - b| < \frac{\epsilon}{2}$$

Consider $|a - b|$. For $n \geq \max\{N(\epsilon_1), N(\epsilon_2)\}$,

$$\begin{aligned}|a-b| &= |a-a_n+a_n-b| \\ &< |a-a_n|+|a_n-b| \\ &< \frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon\end{aligned}$$

The choice of ϵ was arbitrary, so this holds for all $\epsilon > 0$. Thus, $\lim a - b = 0$ or $a = b$. This contradicts our initial assumption and so if $\lim a_n = a$, the limit a is unique. ☐

Problem 4.6. Here are two useful definitions:

(i) A sequence (a_n) is *eventually* in a set $A \subseteq \mathbf{R}$ if there exists an $N \in \mathbf{N}$ such that the $a_n \in A$ for all $n \geq N$.

(ii) A sequence (a_n) is *frequently* in a set $A \subseteq \mathbf{R}$ if for every $N \in \mathbf{N}$, there exists an $n \geq N$, such that $a_n \in A$.

(a) Is the sequence $(-1)^n$ eventually or frequently in the set $\{1\}$?

(b) Which definition is stronger? Does frequently imply eventually or eventually imply frequently?

(c) Give an alternate rephrasing of definition 4.4 using either frequently or eventually. Which is the term we want?

(d) Suppose an infinite number of terms of a sequence (x_n) are equal to 2. Is (x_n) necessarily eventually in the interval $(1.9, 2.1)$? Is it frequently in $(1.9, 2.1)$?

Solution.

(a) The sequence $(-1)^n$ frequently visits the set $\{1\}$.

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

(b) The definition of a sequence (a_n) being eventually in a set $A \subseteq \mathbf{R}$ is stronger. And if a sequence (a_n) is eventually in the set A, it is *frequently* in the set A.

(c) A sequence (a_n) converges to a, if given any $\epsilon-$neighbourhood $V_\epsilon(a)$ of a, the sequence (a_n) is eventually in $V_\epsilon(a)$.

(d) Consider the sequence $(a_n) = (-1)^n \cdot 2$. The first few terms of the sequence are

$$2, -2, 2, -2, 2, -2, ...$$

It has infinite number of two's. It is not eventually in $(1.9, 2.1)$, but the sequence is frequently in $(1.9, 2.1)$.

4.3 Algebraic and Order Limit theorems.

The real purpose of creating a rigorous definition for the convergence of a sequence is *not* to have a tool to verify the computational statements such as $\lim 2n/(n+2) = 2$. Historically, a definition of the limit like definition 4.2 came 150 years after the founders of calculus began working with intuitive notions of convergence. The point of having a logically tight description of convergence is so that we can confidently *prove statements about convergent sequences in general*. We are ultimately trying to resolve arguments about what is and what is not true regarding the behaviour of limits with respect to the mathematical manipulations we intend to inflict on them.

As a first example, let us prove that convergent sequences are bounded. The term bounded has a rather familiar connotation but, like everything else, we need to be explicit about what it means in this context.

Definition 4.6. Bounded sequence. A sequence (x_n) is bounded if there exists a number $M > 0$ such that $|x_n| \leq M$ for all $n \in \mathbf{N}$.

Geometrically, this means that we can find an interval $[-M, M]$ that contains every term in the sequence (x_n).

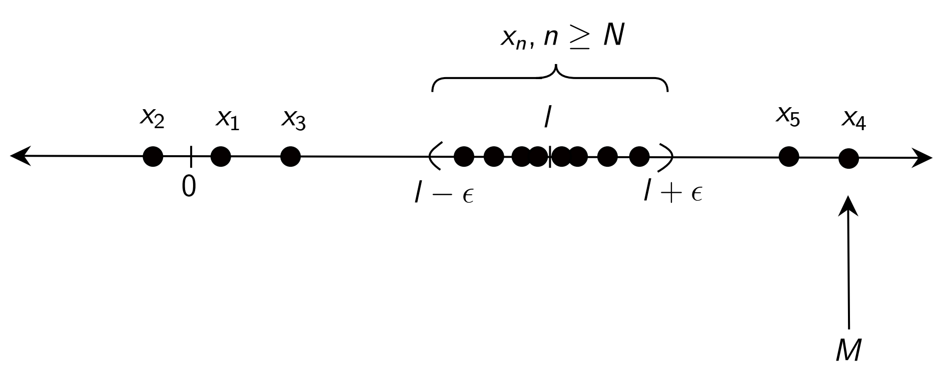
***Theorem:* 4.3.1:**

very convergent sequence is bounded.

Proof. Assume (x_n) converges to a limit l. This means that given a particular value of ϵ, say $\epsilon = 1$, we know that there exists $N \in \mathbf{N}$ such that if $n \geq N$, then x_n is in the interval $(l-1, l+1)$. Not knowing whether l is positive or negative, we can certainly conclude from the picture that :

$$|x_n - l| < 1$$
$$|x_n| < |l| + 1$$

for all $n \geq N$.



We still need to worry (slightly) about the terms in the sequence that come before the Nth term. Because there are only a finite number of these, we let

$$M = \max\{|x_1|, |x_2|, |x_3|, \dots, |x_{N-1}|, |l| + 1\}$$

It follows that $|x_n| \leq M$ for all $n \in \mathbf{N}$ as desired. □

This chapter began with a demonstration of how applying familiar algebraic properties (commutativity of addition) to infinite objects (series) can lead to paradoxical results. These examples are meant to instill in us a sense of caution and justify the extreme care we are taking in drawing our conclusions. The following theorems illustrate that sequences behave extremely well with respect to the operations of addition, subtraction, multiplication, division and order.

***Theorem:* 4.3.2: Algebraic Limit Theorem**

Let $\lim a_n = a$, and $\lim b_n = b$. Then,
(i) $\lim(ca_n) = ca$, for all $c \in \mathbf{R}$
(ii) $\lim(a_n + b_n) = a + b$
(iii) $\lim(a_n b_n) = ab$
(iv) $\lim(a_n/b_n) = a/b$, provided $b \neq 0$

Proof. (i) Consider the case where $c \neq 0$. We want to show that the sequence (ca_n) converges to ca, so the structure of the proof follows the template we described earlier. First, let ϵ be an arbitrary positive but fixed real number. Our goal is to find some point in the sequence (ca_n) after which we have :

$$|ca_n - ca| < \epsilon$$
$$|c||a_n - a| < \epsilon$$

We are given that $(a_n) \to a$, so we know that we can make $|a_n - a|$ as small as we like. In particular, we choose an N such that

$$|a_n - a| < \frac{\epsilon}{|c|}$$

whenever $n \geq N$.

To see that this N indeed works, we observe that, for all $n \geq N$,

$$|c|\,|a_n - a| < |c|\,\frac{\epsilon}{|c|}$$
$$|ca_n - ca| < \epsilon$$

The case $c = 0$ reduces to the constant sequence $(0, 0, 0, ...)$, that converges to zero, which can be easily verified.

Before continuing with parts (ii),(iii) and (iv), we should point out that the proof of (i), while somewhat, short is extremely atypical for a convergence proof. Before embarking on a formal argument, it is a good idea to take an inventory of what we *want* to make less than ϵ, and what we are *given* can be made small for suitable choices of n. For the previous proof, we wanted to make $|ca_n - ca| < \epsilon$ and we were given $|a_n - a| <$ *anything we like* (for large values of n). Notice that in (i), and all of the ensuing arguments, the strategy each time is to bound the quantity we want to be less than ϵ, which in each case is

$$|\text{(terms of the sequence)} - \text{(proposed limit)}|$$

(ii) To prove this statement, we need to argue that the quantity

$$|(a_n + b_n) - (a + b)|$$

can be made less than an arbitrary ϵ using the assumptions that $|a_n - a|$ and $|b_n - b|$ can be made as small as we like or large n. The first step is to use the triangle inequality to say that

$$|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \leq |a_n - a| + |b_n - b|$$

Again we let $\epsilon > 0$ be arbitrary. The technique this time is to divide the ϵ between the two expressions on the right hand side in the preceding inequality. Using the hypothesis, that $(a_n) \to a$, we know that there exists N_1, such that

$$|a_n - a| < \frac{\epsilon}{2} \text{ whenever } n \geq N_1$$

Likewise, the assumption that $(b_n) \to b$ means that we can choose N_2 so that

$$|b_n - b| < \frac{\epsilon}{2} \text{ whenever } n \geq N_2$$

The question now arises as to which of N_1 or N_2 should we take to be our choice of N. By choosing $N = \max\{N_1, N_2\}$, we ensure that if $n \geq N$, then $n \geq N_1$ and $n \geq N_2$. This allows us to conclude that

$$\begin{aligned} |(a_n + b_n) - (a + b)| &\leq |a_n - a| + |b_n - b| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

for all $n \geq N$, as desired.

(iii) To show that $(a_n b_n) \to ab$, we begin by observing that

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - ab_n + ab_n - ab| \\ &\leq |a_n b_n - ab_n| + |ab_n - ab| \\ &= |b_n|\,|a_n - a| + |a|\,|b_n - b| \end{aligned}$$

In the initial step, we subtracted and then added ab_n, which created an opportunity to use the triangle inequality. Essentially, we have broken up the distance from $a_n b_n$ to ab with a midway point and are using the sum of the two distances to overestimate the original distance. This clever trick will become a familiar technique in arguments to come.

Letting $\epsilon > 0$ be arbitrary, we again proceed with the strategy of making each piece in the preceding inequality less than $\epsilon/2$. For the piece on the right hand side $(|a|\,|b_n - b|)$, if we choose N_1 so that,

$$n \geq N_1 \text{ implies that } |b_n - b| < \frac{1}{|a|}\frac{\epsilon}{2}$$

Getting the term on the left-hand side $(|b_n|\,|a_n - a|$ to be less than $\epsilon/2$ is complicated by the fact that we have a variable quantity $|b_n|$ to contend with as opposed to the constant $|a|$ we encountered in the right-hand term. The idea is to replace $|b_n|$ with a worst-case estimate. Using the fact, that convergent sequences are bounded, we know that there exists a bound $M > 0$ satisfying $|b_n| < M$ for all $n \in \mathbf{N}$. Now, we can choose N_2 so that:

$$|a_n - a| < \frac{1}{M}\frac{\epsilon}{2} \text{ whenever } n \geq N_2$$

To finish the argument, pick $N = \max\{N_1, N_2\}$ and observe that if $n \geq N$, then,

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - ab_n + ab_n - ab| \\ &\leq |a_n b_n - ab_n| + |ab_n - ab| \\ &= |b_n|\,|a_n - a| + |a|\,|b_n - b| \\ &< M \cdot \frac{1}{M} \cdot \frac{\epsilon}{2} + |a| \cdot \frac{1}{|a|} \cdot \frac{\epsilon}{2} \\ &= \epsilon \end{aligned}$$

(iv) This final statement will follow from (iii) if we can prove that

$(b_n) \to b$ implies that $\left(\frac{1}{b_n}\right) \to \frac{1}{b}$ whenever $b \neq 0$. We begin by observing that

$$\begin{aligned} \left|\frac{1}{b_n} - \frac{1}{b}\right| &= \left|\frac{b_n - b}{b_n b}\right| \\ &= \frac{1}{|b_n|} \cdot \frac{1}{|b|} \cdot |b_n - b| \end{aligned}$$

Because $(b_n) \to b$, we can make the numerator as small as we like by choosing n large. The problem comes in that we need a worst-case estimate on the size of $1/|b_n|\,|b|$. Because the b_n terms are in the denominator, we are no longer interested in an upper bound on b_n, but rather in an inequality of the form $|b_n| \geq \delta > 0$. This will then lead to a bound on the size of $1/|b_b|\,|b|$.

The trick is to look far enough out into the sequence (b_n), so that the terms are closer to b than they are to 0. Consider the particular value $\epsilon_0 = \frac{|b|}{2}$. Because, $(b_n) \to b$, there exists N_1 such that $|b_n - b| < \frac{|b|}{2}$ for all $n \geq N_1$. This implies that :

$$|b_n - b| < \frac{|b|}{2}$$

In the below figure, the distance $|b_n - b|$ is smaller than $|b|/2$ for all terms $n \geq N$. Clearly, then $|b_n| > |b|/2$

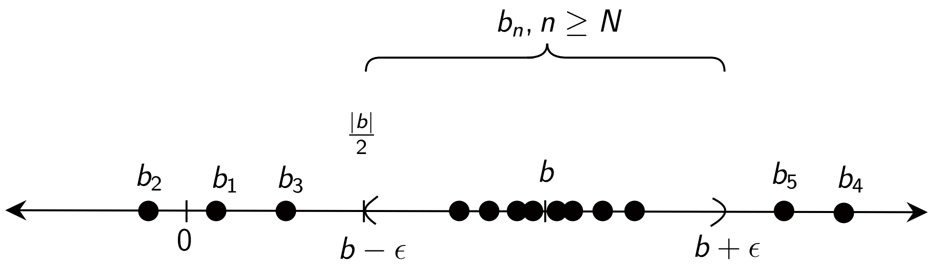


Figure. $|b_n| > \frac{|b|}{2}$

Next, choose N_2 so that $n \geq N_2$ implies that

$$|b_n - b| < \frac{|b|^2}{2}\epsilon$$

Finally, if we let $N = \max\{N_1, N_2\}$, then $n \geq N$ implies that

$$\begin{aligned}\left|\frac{1}{b_n} - \frac{1}{b}\right| &= \left|\frac{b_n - b}{b_n b}\right| \\ &= \frac{1}{|b_n|}\frac{1}{|b|}\,|b_n - b| \\ &< \frac{2}{|b|}\cdot\frac{1}{|b|}\cdot\frac{|b|^2}{2}\epsilon = \epsilon\end{aligned}$$

□

4.3.1 Limits and order.

Although there are few dangers to avoid, the algebraic limit theorem verifies that the relationship between algebraic combinations of sequences and the limiting process is as trouble-free as we could hope for. Limits can be computed from the individual component sequences provided that each component limit exists. This limiting process is also well behaved with respect to the order operation.

***Theorem:* 4.3.3: Order Limit Theorem.**

Assume $\lim a_n = a$ and $\lim b_n = b$.

(i) If $a_n \geq 0$ for all $n \in \mathbf{N}$, then $a \geq 0$

(ii) If $a_n \leq b_n$ for all $n \in \mathbf{N}$, then $a \leq b$.

(iii) If there exists $c \in \mathbf{R}$ for which $c \leq b_n$ for all $n \in \mathbf{N}$, then $c \leq b$. Similarly, if $a_n \leq c$ for all $n \in \mathbf{N}$, then $a \leq c$.

Proof. (i) We will prove this by contradiction; thus let's assume $a < 0$. The idea is to produce a term in the sequence (a_n) that is also less than zero. To do this, we consider the particular value $\epsilon = |a|$. The definition of convergence guarantees that we can find an N such that $|a_n - a| < |a|$ for all $n \geq N$. In particular, this would mean that $|a_N - a| < |a|$, which implies $a_N < 0$.

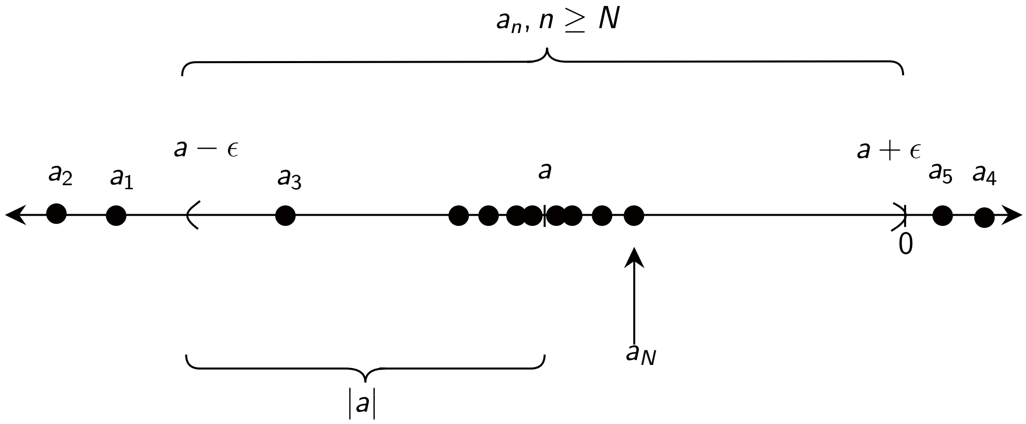


Figure. $|a_N - a| < |a|$ implies that $a_N < 0$

This contradicts our hypotheses that $a_N \geq 0$. We therefore conclude that $a \geq 0$.

(ii) The algebraic limit theorem ensures that sequence $(b_n - a_n)$ converges to $(b - a)$. Because $(b_n - a_n) \geq 0$, we can apply part (i) to get that $b - a \geq 0$.

(iii) Take $(a_n) = c$ (or $b_n = c$) for all $n \in \mathbf{N}$, and apply part (ii). □

A word about the idea of *tails* is in order. Loosely speaking, limits and their properties do not depend at all on what happens at the beginning of the sequence but are strictly determined by what happens when n gets large. Changing the value of the first ten or - ten thousand - terms in particular has no effect on the limit. Theorem 4.3.1 part(i) for instance assumes that $a_n \geq 0$ for all $n \in \mathbf{N}$. However, the hypotheses could be weakened by assuming only that there exists some point N_1 where $a_n \geq 0$ for all $n \geq N_1$. The theorem remains true, and in fact the same proof is valid with the provision that when N is chosen it be atleast as large as N_1.

In the language of analysis, when a property such as non-negativity is not necessarily possessed by some finite number of initial terms but is possessed by all terms in the sequence after some point N, we say that the sequence eventually has this property. Theorem 4.3.1 can be restated, "Convergent sequences that are eventually non-negative converge to non-negative limits". Parts (ii) and (iii) will have similar modifications as will many upcoming results.

Problem 4.7. Let $x_n \geq 0$ for all $n \in \mathbf{N}$.

(a) If $(x_n) \to 0$, show that $(\sqrt{x_n}) \to 0$.

(b) If $(x_n) \to x$, show that $(\sqrt{x_n}) \to \sqrt{x}$.

Proof. (a) We are given that $(x_n) \to 0$, so we can make $|x_n - x|$ as small as we like. In particular, we can choose an N such that $|x_n - 0| < \epsilon^2$ for all $n \geq N$.

Observe that $\left|\sqrt{x_n}\right| = \sqrt{|x_n|}$ since the sequence is non-negative; $x_n \geq 0$ for all $n \in \mathbf{N}$.

So,

$$\begin{aligned} |\sqrt{x_n} - 0| &= \sqrt{|x_n|} \\ &< \sqrt{\epsilon^2} = \epsilon \end{aligned}$$

for all $n \geq N$.

Thus, $(\sqrt{x_n}) \to 0$.

(ii) We are given that $(x_n) \to x$, so we can make $|x_n - x|$ as small as we like.

Observe that,

$$\begin{aligned} \left|\sqrt{x_n} - \sqrt{x}\right| &= \left|\sqrt{x_n} - \sqrt{x} \cdot \frac{\sqrt{x_n} + \sqrt{x}}{\sqrt{x_n} + \sqrt{x}}\right| \\ &= \frac{|x_n - x|}{\left|\sqrt{x_n} + \sqrt{x}\right|} \end{aligned}$$

Since, (x_n) is a non-negative sequence, $x_n \geq 0$ for all $n \in \mathbf{N}$. It follows that $\sqrt{x_n} \geq 0$, $\sqrt{x_n} + \sqrt{x} \geq \sqrt{x}$. So, we have a lower bound for the denominator.

$$\begin{aligned} \left|\sqrt{x_n} - \sqrt{x}\right| &= \frac{|x_n - x|}{\left|\sqrt{x_n} + \sqrt{x}\right|} \\ &\leq \frac{|x_n - x|}{|x|} \end{aligned}$$

We choose N such that, if $n \geq N$,

$$|x_n - x| < |x|\,\epsilon$$

Let's prove that this choice of N indeed works. If $n \geq N$,

$$\begin{aligned} \left|\sqrt{x_n} - \sqrt{x}\right| &\leq \frac{|x_n - x|}{|x|} \\ &< \frac{1}{|x|} \cdot |x| \cdot \epsilon = \epsilon \end{aligned}$$

Thus, $(\sqrt{x_n}) \to \sqrt{x}$. □

Problem 4.8. Using only the definition 4.2 of the convergence of a sequence, prove that if $(x_n) \to 2$, then,

(a) $\left(\frac{2x_n - 1}{3}\right) \to 1$

(b) $\left(\frac{1}{x_n}\right) \to \frac{1}{2}$

(For this exercise, the algebraic limit theorem is off-limits so to speak.)

Proof. (a) We are given that, $(x_n) \to 2$, so we can make the distance $|x_n - 2|$ as small as we like. Observe that,

$$\begin{aligned} \left| \frac{2x_n - 1}{3} - 1 \right| &= \left| \frac{2x_n - 1 - 3}{3} \right| \\ &= \frac{1}{3}|2x_n - 4| = \frac{2}{3}|x_n - 2| \end{aligned}$$

Let us choose an N such that,

$$|x_n - 2| < \frac{3}{2}\epsilon$$

whenever $n \geq N$.

To show that, this choice of N indeed works, formally, we can prove that:

$$\begin{aligned} \left| \frac{2x_n - 1}{3} - 1 \right| &= \frac{2}{3}|x_n - 2| \\ &< \frac{2}{3} \cdot \frac{3}{2} \cdot \epsilon = \epsilon \end{aligned}$$

Hence, $\left|\frac{2x_n - 1}{3}\right| \to 1$.

(b) We are given that $(x_n) \to 2$, so we can make the distance $|x_n - 2|$ as small as we like.

Observe that,

$$\begin{aligned} \left| \frac{1}{x_n} - \frac{1}{2} \right| &= \left| \frac{x_n - 2}{2x_n} \right| \\ &= \frac{|x_n - 2|}{|x_n| \, 2} \end{aligned}$$

Now, there exists N_1 such that $|x_n - 2| < 1$. If the distance between x_N and 2 is less than 1, than $|x_N|$ must be numerically greater than 1. Thus, $\frac{1}{|x_N|} < 1$.

Also, there exists N_2 such that $|x_n - 2| < 2\epsilon$, for all $n \geq N$.

Let us choose an $N = \max\{N_1, N_2\}$. To show that this choice of N indeed works, we prove that,

$$\begin{aligned} \left|\frac{1}{x_n} - \frac{1}{2}\right| &= \frac{|x_n - 2|}{|x_n|\, 2} \\ &< \frac{1}{2} \cdot 2\epsilon = \epsilon \end{aligned}$$

Hence, $\left(\frac{1}{x_n}\right) \to \frac{1}{2}$. □

Problem 4.9. Squeeze Theorem. Show that if $x_n \leq y_n \leq z_n$ for all $n \in \mathbf{N}$, and if $\lim x_n = \lim z_n = l$, then $\lim y_n = l$ as well.

Proof. The Order limit theorem 4.3.1 says that, if $x_n \leq y_n$, then $\lim x_n \leq \lim y_n$. Applying this result to the given proposition, we have:

$$\begin{aligned} & x_n \leq y_n \leq z_n \\ \implies & \lim x_n \leq \lim y_n \leq \lim z_n \end{aligned}$$

And $\lim x_n = \lim z_n = l$. So, $l \leq \lim y_n \leq l$. Consequently, $\lim y_n = l$. □

Problem 4.10. Let $(a_n) \to 0$, and use the algebraic limit theorem to compute each of the following limits (assuming the fractions are always defined):

(a) $\lim \left(\frac{1+2a_n}{1+3a_n-4a_n^2}\right)$

(b) $\lim \left(\frac{(a_n+2)^2-4}{a_n}\right)$

(c) $\lim \left(\frac{\frac{2}{a_n}+3}{\frac{1}{a_n}+5}\right)$

Solution.

(a) The sequence $(1+2a_n) \to 1$. And the sequence $(1+3a_n-4a_n^2) \to 1$. Hence, $\lim \left(\frac{1+2a_n}{1+3a_n-4a_n^2}\right) = 1$.

(b) This can be simplified as,

$$\begin{aligned}\frac{(a_n+2)^2-4}{a_n} &= \frac{a_n^2+4a_n+4-4}{a_n}\\ &= \frac{a_n^2+4a_n}{a_n}\\ &= a_n+4\end{aligned}$$

Hence, this sequence approaches the value 4 when we pass to the limits.

(c) This sequence can be written as

$$\left(\frac{2+3a_n}{1+5a_n}\right)$$

Hence,

$$\lim\left(\frac{2+3a_n}{1+5a_n}\right) = 2$$

Problem 4.11. Let (x_n) and (y_n) be given and define (z_n) to be the shuffled sequence $(x_1, y_1, x_2, y_2, x_3, y_3, \dots, x_n, y_n, \dots)$. Prove that (z_n) is convergent if and only (x_n) and (y_n) are both convergent with $\lim x_n = \lim y_n$.

Proof. $\to$ direction.

Suppose that $(z_n) = x_1, y_1, x_2, y_2, x_3, y_3, \dots, x_n, y_n, \dots$ is a convergent sequence. By the theorem on the uniqueness of limits, a convergent sequence has a unique limit. Assume that $\lim(z_n) = z$.

By definition, as (z_n) is convergent, we can make the distance $|z_n - z|$ as small as we like.

Pick an arbitrary $\epsilon > 0$. Then, there exists an N such that,

$$|z_n - z| < \epsilon$$

for all $n \geq N$.

In particular, if $z_N = x_n$ and $z_{n+1} = y_n$, then we have:

$$\begin{aligned}|x_n - z| &< \epsilon\\ |y_n - z| &< \epsilon\end{aligned}$$

for all $n \geq N$. Hence, $(x_n) \to z$ and $(y_n) \to z$.

$\leftarrow$ direction.

If $(x_n) \to z$, there exists an N_1, such that

$$|x_n - z| < \epsilon$$

whenever $n \geq N_1$.

If $(y_n) \to z$, there exists an N_2, such that

$$|y_n - z| < \epsilon$$

whenever $n \geq N_2$.

Let us choose $N := \max\{N_1, N_2\}$. Then, for all $n \geq N$,

$$|z_n - z| < \epsilon$$

Thus, $(z_n) \to z$. □

Problem 4.12. Consider the sequence given by $b_n = n - \sqrt{n^2 + 2n}$. Taking $(1/n) \to 0$ as given, and using both the algebraic limit theorem and the result in problem 4.7, show that $\lim b_n$ exists and find the value of the limit.

Proof.

$$\begin{aligned} b_n &= n^2 - \sqrt{n^2 + 2n} \\ &= n^2 - \sqrt{n^2 + 2n} \times \frac{n + \sqrt{n^2 + 2n}}{n + \sqrt{n^2 + 2n}} \\ &= \frac{n^2 - (n^2 + 2n)}{n + \sqrt{n^2 + 2n}} \\ &= \frac{-2n}{n + \sqrt{n^2 + 2n}} \\ &= \frac{-2}{1 + \sqrt{1 + \frac{2}{n}}} \end{aligned}$$

As, $\left(\frac{1}{n}\right) \to 0$, the sequence $(b_n) \to -1$. □

Problem 4.13. Give an example of each of the following, or state that such a request is impossible by referencing proper theorem(s):

(a) sequences (x_n) and (y_n), which both diverge, but whose sum $(x_n + y_n)$ converges;

(b) sequences (x_n) and (y_n), where (x_n) converges, (y_n) diverges, and $(x_n + y_n)$ converges;

(c) a convergent sequence (b_n) with $b_n \neq 0$ for all n such that $1/b_n$ diverges;

(d) an unbounded sequence (a_n) and a convergent sequence (b_n) with $(a_n - b_n)$ bounded

(e) two sequences (a_n) and (b_n), where $(a_n b_n)$ and a_n converge but (b_n) does not.

Solution.

(a) Consider the sequence (x_n) given by $x_n = \sqrt{n+1}$ and the sequence (y_n) given by $y_n = -\sqrt{n}$. Both sequences diverge, but the sum $(x_n + y_n)$ converges to 0.

Also, consider the sequence (x_n) given by $x_n = n$ and the sequence (y_n) given by $y_n = -\sqrt{n^2+2n}$. Both sequences diverge, but the sum $(x_n + y_n)$ converges to -1.

Short proof.

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

Observe that,

$$
\begin{aligned}
\sqrt{n+1} - \sqrt{n} &= (\sqrt{n+1} - \sqrt{n}) \times \frac{\sqrt{n+1} + \sqrt{n}}{\sqrt{n+1} + \sqrt{n}} \\
&= \frac{1}{\sqrt{n+1} + \sqrt{n}} \\
&< \frac{1}{\sqrt{n} + \sqrt{n}} = \frac{1}{2\sqrt{n}}
\end{aligned}
$$

Pick $\epsilon > 0$. We can choose $N > \frac{1}{4\epsilon^2}$. To show that this choice of N indeed works, we prove that, that for all $n \geq N$,

$$
\begin{aligned}
\left|\sqrt{n+1} - \sqrt{n}\right| &< \frac{1}{2\sqrt{n}} \\
&< \frac{1}{2\sqrt{(1/4\epsilon^2)}} = \epsilon
\end{aligned}
$$

Thus, $(\sqrt{n+1} - \sqrt{n}) \to 0$.

Consider $b_n = n - \sqrt{n^2+2n}$

Observe that:

$$\begin{aligned}
n-\sqrt{n^2+2n}-(-1) &= [(n+1)-\sqrt{n^2+2n}] \\
&= [(n+1)-\sqrt{n^2+2n}] \times \frac{(n+1)+\sqrt{n^2+2n}}{(n+1)+\sqrt{n^2+2n}} \\
&= \frac{(n+1)^2-(n^2+2n)}{(n+1)+\sqrt{n^2+2n}} \\
&= \frac{1}{(n+1)+\sqrt{n^2+2n}} \\
&< \frac{1}{n+\sqrt{n^2}} = \frac{1}{2n}
\end{aligned}$$

Pick an arbitrary $\epsilon > 0$. We choose an $N > \frac{1}{2\epsilon}$. To show that choice of N indeed works, we find that:

$$\begin{aligned}
\left|n-\sqrt{n^2+2n}-(-1)\right| &< \frac{1}{2n} \\
&< \frac{1}{2}\cdot(2\epsilon) = \epsilon
\end{aligned}$$

Thus, $(n-\sqrt{n^2+2n}) \to -1$.

(b) This request is impossible. Alternatively, we believe that if (x_n) converges and (x_n+y_n) converges, then (y_n) converges. Let us prove this fact rigorously. Suppose $\lim x_n = a$ and $\lim x_n + y_n = b$. We shall prove that $\lim y_n = b - a$.

Observe that,

$$\begin{aligned}
|y_n-(b-a)| &= |(x_n+y_n)-x_n-(b-a)| \\
&= |(x_n+y_n-b)-(x_n-a)| \\
&\leq |x_n+y_n-b|+|x_n-a|
\end{aligned}$$

Pick an $\epsilon > 0$. Since $(x_n) \to a$, we can make the distance $|x-a|$ as small as we like. There exists an N_1 such that

$$|x_n - a| < \frac{\epsilon}{2}$$

for all $n \geq N_1$.

Since $(x_n + y_n) \to b$, we can make the distance $|x_n + y_n - b|$ as small as we like. There exists an N_2 such that,

$$|x_n + y_n - b| < \frac{\epsilon}{2}$$

Let $N = \max\{N_1, N_2\}$. To show that this N indeed works, we find that:

$$\begin{aligned} |y_n - (b - a)| &\leq |x_n + y_n - b| + |x_n - a| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

(c) Consider the sequence (b_n) given by $b_n = \frac{1}{n}$. Then, (b_n) is a convergent sequence but $1/b_n$ is divergent.

Consider $(b_n) = \frac{1}{n}$. Pick an arbitrary $\epsilon > 0$. We can choose $N > \frac{1}{\epsilon}$. To show that this choice of N indeed works, we find that:

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

for all $n \geq N$. Consequently, $(1/n) \to 0$.

(d) This request is impossible.

(e) Consider the sequence (a_n) given by $a_n = \frac{1}{n}$ and $(a_n b_n)$ given by $a_n b_n = \frac{\sin n}{n}$. Thus, $(a_n b_n)$ and (a_n) converges, but (b_n) does not. Let us prove $\frac{\sin n}{n}$ converges to 0.

Observe that,

$$\begin{aligned} \left|\frac{\sin n}{n}\right| &= \frac{|\sin n|}{|n|} \\ &\leq \frac{1}{n} \end{aligned}$$

Let $\epsilon > 0$ be an arbitary small but fixed positive real number. We choose an $N > \frac{1}{\epsilon}$. To prove that this choice of N indeed works, we have,

$$\begin{aligned} \left|\frac{\sin n}{n}\right| &\leq \frac{1}{n} \\ &< \frac{1}{(1/\epsilon)} = \epsilon \end{aligned}$$

Consequently, $\frac{\sin n}{n} \to 0$.

Problem 4.14. Let $(x_n) \to x$ and let $p(x)$ be a polynomial.

(a) Show $p(x_n) \to p(x)$.

(b) Find an example of a function $f(x)$ and a convergent sequence $(x_n) \to x$ where the sequence $f(x_n)$ converges, but not to $f(x)$.

Proof. (a) We are given that $(x_n) \to x$, so we can make the distance $|x_n - x|$ as small as we like. Suppose $p(x) = x^k$ where k is a non-negative integer.

Observe that,

$$\begin{aligned} |p(x_n) - p(x)| &= \left|x_n^k - x^k\right| \\ &= \left|(x_n - x)(x_n^{k-1} + x_n^{k-2} \cdot x + ... + x_n \cdot x^{k-2} + x^{k-1})\right| \end{aligned}$$

Also,

$$\left|(x_n^{k-1} + x_n^{k-2} \cdot x + ... + x_n \cdot x^{k-2} + x^{k-1})\right| \leq \left|x_n^{k-1}\right| + \left|x_n^{k-2} \cdot x\right| + ... + \left|x_n \cdot x^{k-2}\right| + \left|x^{k-1}\right|$$
$$\frac{1}{|x_n^{k-1}| + |x_n^{k-2} \cdot x| + ... + |x_n \cdot x^{k-2}| + |x^{k-1}|} \leq \frac{1}{|(x_n^{k-1} + x_n^{k-2} \cdot x + ... + x_n \cdot x^{k-2} + x^{k-1})|}$$

Now, there exists N_1 such that $|x_n - x| < \frac{|x|}{2}$ for $n \geq N_1$. If the distance between x_N and x is smaller than $\frac{|x|}{2}$, then clearly $|x_n| > \frac{|x|}{2}$.

Thus,

$$\begin{aligned} &\frac{1}{k \cdot \frac{\left|x^{k-1}\right|}{2^{k-1}}} < \frac{1}{|(x_n^{k-1} + x_n^{k-2} \cdot x + ... + x_n \cdot x^{k-2} + x^{k-1})|} \\ \implies &\frac{2^{k-1}}{k \cdot |x^{k-1}|} < \frac{1}{|(x_n^{k-1} + x_n^{k-2} \cdot x + ... + x_n \cdot x^{k-2} + x^{k-1})|} \end{aligned}$$

There exists N_2 such that $|x_n - x| < \epsilon \cdot \frac{2^{k-1}}{k \cdot \left|x^{k-1}\right|}$ for all $n \geq N_2$.

Let $N = \max\{N_1, N_2\}$. To show that this N indeed works, we prove that:

$$
\begin{aligned}
|p(x_n) - p(x)| &= \left|(x_n - x)(x_n^{k-1} + x_n^{k-2} \cdot x + ... + x_n \cdot x^{k-2} + x^{k-1})\right| \\
&< \epsilon \cdot \frac{2^{k-1}}{k \cdot |x^{k-1}|} \cdot \left|(x_n^{k-1} + x_n^{k-2} \cdot x + ... + x_n \cdot x^{k-2} + x^{k-1})\right| \\
&< \epsilon \cdot \frac{1}{|(x_n^{k-1} + x_n^{k-2} \cdot x + ... + x_n \cdot x^{k-2} + x^{k-1})|} \cdot \left|(x_n^{k-1} + x_n^{k-2} \cdot x + ... + x_n \cdot x^{k-2} + x^{k-1})\right| \\
&= \epsilon
\end{aligned}
$$

for all $n \geq N$. Thus, $p(x_n) \to p(x)$.

By algebra of limits, any linear combination of the powers of the sequence (x_n) must be convergent and $\lim p(x_n) = p(x)$.

(b) Consider the following function with a jump discontinuity at $x = 0$:

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

Let (x_n) be a sequence, where $x_n = \frac{1}{n}$. The sequence (x_n) converges to 0. But, $f(x_n) = 1$ for all x_n, whereas $f(x) = f(0) = 0$. $\square$

Problem 4.15. (a) Let (a_n) be a bounded (not necessarily convergent) sequence, and assume $\lim b_n = 0$. Show that $\lim(a_n b_n) = 0$. Why are we not allowed to use the Algebraic limit theorem to prove this?

(b) Can we conclude anything about the convergence of $(a_n b_n)$ if we assume that (b_n) converges to some nonzero limit b?

(c) Use (a) to prove theorem 4.3 part (iii), for the case when $a = 0$.

Proof. (a) We are given that (a_n) is a bounded sequence. There exists a larrge number $M > 0$, such that

$$
|a_n| < M
$$

for all n.

Also, we have $\lim b_n = 0$. Pick a small arbitrary but fixed positive real number $\epsilon > 0$. There exists an N_ϵ such that,

$$
|b_n| < \frac{\epsilon}{M}
$$

for all $n \geq N_\epsilon$.

Observe that,

$$\begin{aligned} |a_n b_n - 0| &= |a_n| \cdot |b_n| \\ &< M \cdot \frac{\epsilon}{M} = \epsilon \end{aligned}$$

Hence, $|a_n b_n - 0| < \epsilon$ for all $n \geq N_\epsilon$.

Consequently, $\lim(a_n b_n) = 0$.

We are not allowed to use Algebraic limit theorem to prove this, because Algebraic limit theorem applies to two convergent sequences.

(b) Suppose (a_n) is a bounded sequence and (b_n) converges to some non-zero limit b. We can prove that $(a_n b_n)$ is also a bounded (not necessarily a convergent sequence).

Since, (a_n) is a bounded sequence, then there exists a large number $M > 0$, such that

$$|a_n| < M$$

Since, $(b_n) \to b$, we can make the distance $|b_n - b|$ as small as we like. Pick an arbitrary $\epsilon > 0$. There exists an $N_\epsilon \in \mathbf{N}$, such that

$$|b_n - b| < \epsilon$$

Clearly, if b_n eventually lies in the interval $(b-\epsilon, b+\epsilon)$, we have $|b_n| < |b| + \epsilon$ for all $n \geq N_\epsilon$.

Thus, we find that

$$\begin{aligned} |a_n b_n| &= |a_n| \cdot |b_n| \\ &< M \cdot (|b| + \epsilon) \end{aligned}$$

for all $n \geq N_\epsilon$. But, what about the terms of the sequence $(a_n b_n)$ preceding N_ϵ? There are only a finite number of them. If we choose $K = \max\{|a_1 b_1|, |a_2 b_2|, \ldots, M(|b| + \epsilon)\}$, then

$$|a_n b_n| < K$$

for all n.

In the special case that, (a_n) is also convergent, then it follows using the algebraic limit theorem, $(a_n b_n)$ also converge. $\square$

Problem 4.16. Consider the following list of conjectures. Provide a short proof for those that are true and a counterexample for any that are false.

(a) If $\lim(a_n - b_n) = 0$, then $\lim a_n = \lim b_n$.

(b) If $(b_n) \to b$, then $|b_n| \to |b|$.

(c) If $(a_n) \to a$ and $(b_n - a_n) \to 0$, then $(b_n) \to a$.

(d) If $(a_n) \to 0$ and $|b_n - b| \leq a_n$ for all n, then $(b_n) \to b$.

Proof. (a) Let $\lim a_n = a$, $\lim b_n = b$. Let's pick an arbitrary but fixed positive real $\epsilon > 0$. By definition, there exists $N_1 \in \mathbf{N}$, such that $|a_n - a| < \frac{\epsilon}{2}$ for all $n \geq N_1$. Also, there exists $N_2 \in \mathbf{N}$, such that $|b_n - b| < \frac{\epsilon}{2}$. Let $N = \max\{N_1, N_2\}$. We have,

$$\begin{aligned} |a_n - b_n - (a - b)| &\leq |a_n - a| + |b_n - b| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

for all $n \geq N$. Thus, $\lim(a_n - b_n) = \lim a_n - \lim b_n$. Thus, if $\lim(a_n - b_n) = 0$, we must have $\lim a_n = \lim b_n$.

(b) Since, $(b_n) \to b$, we can make $|b_n - b|$ as small as we like. Pick an arbitrary $\epsilon > 0$. There exists $N \in \mathbf{N}$, such that

$$|b_n - b| < \epsilon$$

for all $n \geq N$. Since, the terms of the sequence (b_n) are eventually in the set $(b - \epsilon, b + \epsilon)$, the terms $|b_n|$ will be within ϵ of $|b|$. That is,

$$||b_n| - |b|| < \epsilon$$

Hence, $|b_n| \to |b|$.

(c) Since, $\lim(b_n - a_n) + \lim(a_n) = \lim(b_n - a_n + a_n)$, it follows that $0 + a = \lim b_n$, that is $\lim b_n = a$.

(d) We are given that the sequence $(a_n) \to 0$. Thus, we can make $|a_n|$ as small as we like. Pick an arbitrary $\epsilon > 0$. There exists $N_\epsilon \in \mathbf{N}$, such that

$$|a_n| < \epsilon$$

The distance $|b_n - b|$ is always smaller than a_n for all $n \in \mathbf{N}$. Therefore, for $n \geq N_\epsilon$

$$|b_n - b| \leq a_n \leq |a_n| < \epsilon$$

Since, the initial choice of ϵ was arbitrary, by definition of convergence, it follows that $\lim b_n = b$. $\square$

Problem 4.17. Cesaro means. (a) Show that if (x_n) is a convergent sequence, then the sequence given by the averages

$$y_n = \frac{x_1 + x_2 + ... + x_n}{n}$$

also converges to the same limit.

(b) Give an example to show that it is possible for the sequence (y_n) of averages to converge even if (x_n) does not.

Proof. (a) If (x_n) is a convergent sequence, then it is also bounded. Therefore, $|x_n| < M$ for all $n \in \mathbf{N}$. Moreover, by the definition of convergence, we can make the distance $|x_n - x|$ as small as we like. Pick an arbitrary $\epsilon > 0$. There exists $N_1 \in \mathbf{N}$, such that,

$$|x_n - x| < \frac{\epsilon}{2}$$

for $n \geq N_1$. Therefore, if $n \geq N_1$

$$\begin{aligned}
|y_n - x| &= \left|\frac{x_1 + x_2 + ... + x_n}{n} - x\right| \\
&\leq \left|\sum_{i=1}^{N-1} \frac{x_i - x}{n}\right| + \left|\sum_{i=N}^{n} \frac{x_i - x}{n}\right| \\
&< \frac{(N-1)M}{n} + \frac{(N-1)\,|x|}{n} + \left(1 - \frac{N-1}{n}\right)\frac{\epsilon}{2} \\
&< \frac{(N-1)(M + |x|)}{n} + \frac{\epsilon}{2}
\end{aligned}$$

Further, as the sequence $\frac{C}{n}$ converges to 0, given any arbitrary $\epsilon > 0$, there exists $N_2 \in \mathbf{N}$, such that $\frac{C}{n} < \frac{\epsilon}{2}$ for $n \geq N_2$.

Let $N = \max\{N_1, N_2\}$. Then, for all $n \geq N$, we have:

$$|y_n - x| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

(b) □

Problem 4.18. A typical task in analysis is to decipher whether a property possessed by every term in a convergent sequence is necessarily inherited by the limit. Assume $(a_n) \to a$ and determine the validity of each claim. Assume $(a_n) \to a$ and determine the validity of each claim. Try to produce a counterexample of any that are false.

(a) If every a_n is an upper bound for a set B, then a is also an upper bound for B.

(b) If every (a_n) is the complement of the interval $(0, 1)$, then a is also in the complement of $(0, 1)$.

(c) If every (a_n) is rational, then a is rational.

Proof. (a) Let b be an arbitrary element of B. If every a_n is the upper bound for a set B, then $b \leq a_n$, for all $n \in \mathbf{N}$. By the order limit theorem, it follows that $b \leq \lim a_n$, so $b \leq a$. Thus, a is also an upper bound for B.

Let us rigorously prove this. Suppose that $a < \sup B$. Then, we can find a neighbourhood of radius ϵ around a, such that the entire neighbourhood is less than $\sup B$. This implies, there exists an a_n in this neighbourhood, that is less than $\sup B$. But, then a_n is not an upper bound for B, since it is smaller than the least upper bound.

(c) This claim is false. Consider the sequence $(a_n) = \{r \in \mathbf{Q} : r^2 < 2\}$. The supremum or limiting value of this increasing sequence is not a rational number. □

Definition 4.7. If (x_n) is a sequence of real numbers and if m is a given natural number, then the m-tail of the sequence is defined as,

$$(x_n)_{n=m}^{\infty} = (x_{m+n}) = x_{m+1}, x_{m+2}, x_{m+3}, \ldots$$

Theorem: 4.3.4: Behaviour of the m-tail of a sequence.

Let (x_n) be a sequence of real numbers and let $m \in \mathbf{N}$. Then the m-tail of the sequence $(x_n)_{n=m}^{\infty}$, $\forall n \in \mathbf{N}$, converges if and only if the sequence (x_n) converges. In this case,

$$\lim(x_n) = \lim(x_n)_{n=m}^{\infty}$$

Proof. $(\rightarrow)$ direction. If $q > m$, the qth term of the m-tail is the $(m+q)$th term of the original sequence. Pick an arbitrary $\epsilon > 0$. Since the m-tail of the sequence is convergent, there exists $N_m \in \mathbf{N}$, such that terms beginning with the N_mth term in the m-tail can be made as close as we like to x. That is,

$$|x_n - x| \leq \epsilon$$

for all $n \geq N_m$. But, the N_mth term in the m-tail is the $m + N_m$th term in the original sequence. Define $N := m + N_m$. Then, $(\forall \epsilon > 0)$, $\exists N$ such that $|x_n - x| < \epsilon$, whenever $n \geq N$. So, $(x_n) \rightarrow x$.

$(\leftarrow)$ direction. If $p > m$, the pth term of the sequence (x_n) is the $(p-m)th$ term of the original sequence. Pick an arbitrary $\epsilon > 0$. Since the sequence (x_n) is convergent, there exists $N < m \in \mathbf{N}$, such that terms beginning with the Nth term can be made as close as we like to x. That is,

$$|x_n - x| \leq \epsilon$$

for all $n \geq N$. Define $N_m := N - m$. Then, $(\forall \epsilon > 0)$, $\exists N_m$ such that the distance $|x_n - x| < \epsilon$ for the terms in m-tail, whenever $n \geq N_m$. So, $(x_n)_{n=m}^{\infty} \rightarrow x$. □

4.4 The Monotone Convergence Theorem.

We showed in theorem (4.3) that convergent sequences are bounded. The converse statement is certainly not true. It is not too difficult to produce a bounded sequence that does not converge. On the other hand, if a bounded sequence is *monotone*, then in fact, it does converge.

Definition 4.8. A sequence (a_n) is increasing if $a_n \leq a_{n+1}$ for all $n \in \mathbf{N}$ and decreasing if $a_n \geq a_{n+1}$ for all $n \in \mathbf{N}$. A sequence is monotone, if it is either increasing or decreasing.

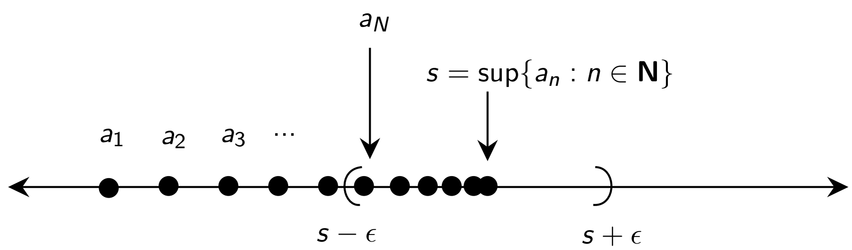
***Theorem:* 4.4.1: Monotone Convergence Theorem.**

If a sequence is monotone and bounded, then it converges.

Proof. Let (a_n) be monotone and bounded. To prove that (a_n) converges using the definition of convergence, we are going to need a candidate for the limit. Let's assume the sequence is increasing (the decreasing case is handled similarly), and we consider the set of points $\{a_n : n \in \mathbf{N}\}$. By assumption, this set is bounded, so we can let

$$s = \sup\{a_n : n \in \mathbf{N}\}$$

It is reasonable to claim that $\lim a_n = s$.



To prove this, let $\epsilon > 0$. Because, s is the least upper bound for the set $\{a_n : n \in \mathbf{N}\}$, $s - \epsilon$ is not an upper bound, so there exists a point in the sequence a_N, such that $s - \epsilon < a_N$. Now, the fact that (a_n) is increasing implies that if $n \geq N$,

$$s - \epsilon < a_N \leq a_n \leq s < s + \epsilon$$

which implies that $|a_n - s| < \epsilon$, as desired. □

The Monotone Convergence theorem is extremely useful for the study of infinite series, largely because it asserts the convergence of a sequence without explicit mention of the actual limit. This is a good moment to do some preliminary investigations, so it is time to formalize the relationship between sequences and series.

Definition 4.9. (Convergence of a series). Let (b_n) be a sequence. An infinite series is a formal expression of the form:

$$\sum_{n=1}^{\infty} b_n = b_1 + b_2 + b_3 + b_4 + b_5 + ...$$

We define the corresponding sequence of partial sums (s_m) by

$$s_m = b_1 + b_2 + b_3 + b_4 + b_5 + ... + b_m$$

and we say that the series $\sum_{n=1}^{\infty} b_n$ converges to B, if the sequence (s_m) converges to B. In this case, we write $\sum_{n=1}^{\infty} b_n = B$.

Example 4.4.1. Consider

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

Because the terms in the sum are all positive, the sequence of the partial sums given by

$$s_m = 1 + \frac{1}{4} + \frac{1}{9} + ... + \frac{1}{m^2}$$

is increasing. The question is whether or not, we can find some upper bound on (s_m). To this end, we observe that,

$$\begin{aligned}
s_m &= 1 + \frac{1}{2 \cdot 2} + \frac{1}{3 \cdot 3} + \frac{1}{4 \cdot 4} + ... + \frac{1}{m \cdot m} \\
&< 1 + \frac{1}{2 \cdot 1} + \frac{1}{3 \cdot 2} + \frac{1}{4 \cdot 3} + ... + \frac{1}{m(m-1)} \\
&= 1 + \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + ... + \left(\frac{1}{m-1} - \frac{1}{m}\right) \\
&= 1 + 1 - \frac{1}{m} \\
&< 2
\end{aligned}$$

Thus, 2 is an upper bound for the sequence of partial sums, so by the Monotone Convergence Theorem, $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges to some (for the moment) unknown limit less than 2. Finding the value of this limit is the subject of some further sections.

Example 4.4.2. (Harmonic Series). This time, consider the so called *harmonic series*

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

Because the terms in the sequence are all positive, again we have the sequence of partial sums

$$s_m = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + ... + \frac{1}{m}$$

increasing, that upon naive inspection appears as though it may be bounded. However, 2 is no longer an upper bound because,

$$s_4 = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) > 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) = 2$$

A similar calculation shows that $s_8 > \frac{5}{2}$,

$$\begin{aligned} s_8 &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \\ &> 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) = \frac{5}{2} \end{aligned}$$

and we can see in general that

$$\begin{aligned} s_{2^k} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \dots + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^{k-1}+1} + \dots + \frac{1}{2^k}\right) \\ &< 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \dots + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^k} + \dots + \frac{1}{2^k}\right) \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \dots + \frac{1}{2} \\ &= 1 + k\left(\frac{1}{2}\right) \end{aligned}$$

which is unbounded. Thus, despite the incredibly slow pace, the sequence of partial sums of $\sum_{n=1}^{\infty} 1/n$ eventually surpasses every number on the positive real line. Because, convergent sequences are bounded, the harmonic series diverges.

The previous example is a special case of a general argument that can be used to determine the convergence or divergence of a large class of infinite series.

***Theorem:* 4.4.2: Cauchy Condensation Test.**

Suppose (b_n) is decreasing and satisfies $b_n \geq 0$ for all $n \in \mathbf{N}$. Then, the series $\sum_{n=1}^{\infty} b_n$ converges if and only if the series

$$\sum_{n=0}^{\infty} 2^n b_{2^n} = b_1 + 2b_2 + 4b_4 + 8b_8 + 16b_{16} + \dots$$

converges.

Proof. First, assume that $\sum_{n=0}^{\infty} 2^n b_{2^n}$ converges. Since convergent sequences are bounded,theorem 4.3) guarantees that the partial sums

$$t_k = b_1 + 2b_2 + 4b_4 + \dots + 2^k b_{2^k}$$

are bounded; that is, there exists an $M > 0$ such that $t_k \leq M$ for all $k \in \mathbf{N}$. We want to prove that $\sum_{n=1}^{\infty} b_n$ converges. Because $b_n \geq 0$, we know that the partial sums are increasing, so we only need to show that

$$s_m = b_1 + b_2 + b_3 + ... + b_m$$

is bounded.

Fix m and let k be large enough to ensure $m \leq 2^{k+1} - 1$. Then, $s_m \leq s_{2^{k+1}-1}$ and

$$\begin{aligned} s_{2^{k+1}-1} &= b_1 + (b_2 + b_3) + (b_4 + b_5 + b_6 + b_7) + ... + (b_{2^k} + b_{2^k+1} + ... + b_{2^{k+1}-1}) \\ &\leq b_1 + (b_2 + b_2) + (b_4 + b_4 + b_4 + b_4) + ... + (b_{2^k} + b_{2^k} + ... + b_{2^k}) \\ &= b_1 + 2b_2 + 4b_4 + ... + 2^k b_{2^k} \\ &= t_k \end{aligned}$$

Thus, $s_m \leq t_k \leq M$ and the sequence (s_m) is bounded. By the Monotone Convergence Theorem, we can conclude that $\sum_{n=1}^{\infty} b_n$ converges. The proof that $\sum_{n=0}^{\infty} 2^n b_{2^n}$ diverges implies that $\sum_{n=1}^{\infty} b_n$ diverges is similar to the proof in the last example. □

Problem 4.19. (a) Prove that the sequence defined by $x_1 = 3$ and

$$x_{n+1} = \frac{1}{4 - x_n}$$

converges.

(b) Now that we know $\lim x_n$ exists, explain why $\lim x_{n+1}$ must also exist and equal the same value.

(c) Take the limit of each side of the recursive equation in part (a) to explicitly compute the $\lim x_n$.

Proof. (a) By direct computation, we find that $x_1 = 3$, $x_2 = 1$. Let us prove that (x_n) is a decreasing sequence, that is $x_{n+1} < x_n$ for all $n \in \mathbf{N}$. This is true for $n = 1$. By induction, let's assume that $x_{k+1} < x_k$. Therefore,

$$x_{k+1} < x_k \implies \frac{1}{4 - x_{k+1}} < \frac{1}{4 - x_k} \implies x_{k+2} < x_{k+1}$$

So, (x_n) is a monotonically decreasing sequence.

Moreover, we can show that (x_n) is bounded. We are interested to show that $x_n > 0$ for all $n \in \mathbf{N}$. This holds for $n = 1$. Assume that $x_k > 0$, then

$$x_{k+1} = \frac{1}{4 - x_k} > \frac{1}{4} > 0$$

Thus, the sequence (x_n) has a lower bound 0. By the Monotone Convergence Theorem, the sequence (x_n) converges.

(b) The sequence $(x_{n+1}) = x_2, x_3, x_4, ...$ also converges and has the same limiting value because the 1-tail of (x_n) is a (i) monotonically decreasing sequence (ii) has the same lower bound 0. It is the infinite tail of the sequence, that ultimately determines the convergence of a sequence.

(c) We have:

$$\begin{aligned}
\lim(x_n) &= \frac{1}{4 - \lim x_{n+1}} \\
L &= \frac{1}{4 - L} \\
4L - L^2 &= 1 \\
L^2 - 4L + 1 &= 0 \\
(L-2)^2 - 3 &= 0 \\
L &= 2 \pm \sqrt{3}
\end{aligned}$$

As $0 < L < 1$, we have $L = 2 - \sqrt{3}$. □

Problem 4.20. (a) Consider the recursively defined sequence $y_1 = 1$,

$$y_{n+1} = 3 - y_n$$

and set $y = \lim y_n$ Because (y_n) and (y_{n+1}) have the same limit, taking the limit across the recursive equation gives $y = 3 - y$. Solving for y, we conclude that $\lim y_n = \frac{2}{3}$. What is wrong with this argument?

(b) This time set $y_1 = 1$ and $y_{n+1} = 3 - \frac{1}{y_n}$. Can the strategy in (a) be applied to compute the limit of this sequence?

Proof. (a) The first few terms of the sequence are $y_1 = 1, y_2 = 2, y_3 = 1, y_4 = 2, y_5 = 1, ...$ The sequence (y_n) is not a convergent sequence. Hence, $\lim y_n$ does not exist. If a sequence is divergent, then it's m-tail would also be divergent. So, $\lim y_{n+1}$ does not exist.

(b) The first few terms of the sequence are $1, 2, \frac{5}{2}, \frac{13}{5}, ...$ We would like to prove that (y_n) is a monotonically increasing sequence. Clearly, $y_2 > y_1$, so this is true for $n = 1$. Assume that $y_{k+1} > y_k$. Thus, $3 - \frac{1}{y_{k+1}} > 3 - \frac{1}{y_k}$. Consequently, $y_{k+2} > y_{k+1}$. Whence, $y_{n+1} > y_n$ for all $n \in \mathbf{N}$.

Thus, (y_n) is a monotonically increasing sequence. All of the terms of the sequence (y_n) are real positive numbers. So, $\frac{1}{y_n} > 0$ for all natural numbers $n \in \mathbf{N}$.

Therefore, $y_{k+1} = 3 - \frac{1}{y_k} < 3$ for all $k \in \mathbf{N}$. So, the sequence (y_n) is bounded. Therefore, by the Monotone Convergence Theorem, the sequence (y_n) converges.

The strategy in (a) can now be applied to compute the limit of this sequence.

$$\begin{aligned}
\lim y_{n+1} &= 3 - \frac{1}{\lim y_n} \\
L &= 3 - \frac{1}{L} \\
L^2 &= 3L - 1 \\
L^2 - 3L + 1 &= 0 \\
L^2 - 2 \cdot L \cdot \frac{3}{2} + \left(\frac{3}{2}\right)^2 - \frac{5}{4} &= 0 \\
\left(L - \frac{3}{2}\right)^2 - \left(\frac{\sqrt{5}}{2}\right)^2 &= 0 \\
L &= \frac{1}{2}(3 \pm \sqrt{5})
\end{aligned}$$

Thus, $\lim(y_n) = \frac{1}{2}(3 + \sqrt{5})$ □

Problem 4.21. (a) Show that

$$\sqrt{2}, \sqrt{2 + \sqrt{2}}, \sqrt{2 + \sqrt{2 + \sqrt{2}}}, \ldots$$

converges and find the limit.

(b) Does the sequence

$$\sqrt{2}, \sqrt{2\sqrt{2}}, \sqrt{2\sqrt{2\sqrt{2}}}, \ldots$$

converge? If so, find the limit.

Proof. (a) The given sequence is defined by the recurrence relation

$$x_{n+1} = \sqrt{2 + x_n}$$

The first few terms of the sequence are $\sqrt{2}, \sqrt{2 + \sqrt{2}}, \ldots$. Let us prove that (x_n) is a monotonically increasing sequence. We induct on n. For $n = 1$, $x_2 > \sqrt{2}$, so $x_2 > x_1$. Assume that, $x_{k+1} > x_k$. Then, $2 + x_{k+1} > 2 + x_k$. It follows that, $\sqrt{2 + x_{k+1}} > 2 + x_k$. Therefore, we have $x_{k+2} > x_{k+1}$. By principle of mathematical induction, $x_{n+1} > x_n$ for all $n \in \mathbf{N}$.

We would also like to prove that (x_n) is bounded. Clearly, $x_1 \le 2$. Assume that $x_k \le 2$. Then, $2 + x_k \le 4$, so that $\sqrt{2 + x_k} < \sqrt{4} = 2$. Therefore, $x_{k+1} < 2$. By principle of mathematical induction, $x_n < 2$ for all $n \in \mathbf{N}$.

By the Monotone Convergence Theorem, the sequence (x_n) converges. Let $\lim x_n = x$. Taking limits on both sides, we have,

$$\begin{aligned} x^2 &= 2 + x \\ x^2 - x - 2 &= 0 \\ (x-2)(x+1) &= 0 \end{aligned}$$

Thus, $x = 2$ or $x = -1$. Since, $x_n > 0$ for all $n \in \mathbf{N}$, the limit value $x > 0$, and thus we have $x = 2$.

(b) The sequence is defined by the recurrence relation:

$$x_{n+1} = \sqrt{2x_n}$$

The first few terms of the sequence are $\sqrt{2}, \sqrt{2\sqrt{2}}, \ldots$. We would like to prove that (x_n) is a monotonically increasing sequence. For $n = 1$, $x_2 > x_1$. Assume that $x_{k+1} > x_k$. Then $2x_{k+1} > 2x_k$, from which it follows that $\sqrt{2x_{k+1}} > \sqrt{2x_k}$. Consequently, $x_{k+2} > x_{k+1}$. By the principle of mathematical induction, $x_{k+1} > x_k$ for all $k \in \mathbf{N}$. (x_n) is an increasing sequence.

We would like to find an upper bound for the sequence (x_n). Again, $x_1 < 2$. Assume that $x_k < 2$. Then, $x_{k+1} = \sqrt{2x_k} < \sqrt{2 \cdot 2} = 2$. By principle of mathematical induction, $x_k < 2$ for all $k \in \mathbf{N}$. 2 is an upper bound for (x_n).

By the monotone convergence theorem, the sequence (x_n) is convergent. Therefore, $\lim x_{n+1} = \lim x_n = x$. We have,

$$\begin{aligned} x^2 &= 2x \\ x(x-2) &= 0 \end{aligned}$$

So, $x = 0$ or $x = 2$. Since, $x > 0$, the sequence (x_n) converges to 2. □

Problem 4.22. Let $y_1 = \sqrt{p}$, where $p > 0$ and $y_{n+1} := \sqrt{p + y_n}$ for all $n \in \mathbf{N}$. Show that (y_n) converges and find the limit.

Proof. The first few terms of the sequence are $\sqrt{p}, \sqrt{p + \sqrt{p}}, \sqrt{p + \sqrt{p + \sqrt{p}}}, \ldots$. We would like to prove that (y_n) is a monotonically increasing sequence. Clearly, $y_2 > \sqrt{p} = y_1$. Assume that $y_{k+1} > y_k$. Then, $p + y_{k+1} > p + y_k$. So, $\sqrt{p + y_{k+1}} > \sqrt{p + y_k}$. Thus, it follows that $y_{k+2} > y_{k+1}$. By principle of mathematical induction, $y_{k+1} > y_k$ for all $k \in \mathbf{N}$.

We are interested to find an upper bound for the sequence (y_n). We can show that $1 + 2\sqrt{p}$ is an upper bound for the sequence (y_n). For $n = 1$, we have $x_1 = \sqrt{p} < 1 + 2\sqrt{p}$. Assume that $y_k < 1 + 2\sqrt{p}$. Then, $y_{k+1} = \sqrt{p + y_k} < \sqrt{p + 1 + 2\sqrt{p}} = \sqrt{(1 + \sqrt{p})^2} = 1 + \sqrt{p} < 1 + 2\sqrt{p}$. Thus, $y_{k+1} < 1 + 2\sqrt{p}$. Consequently, $1 + 2\sqrt{p}$ is an upper bound for the sequence (y_n).

By the monotone convergence theorem, the sequence (y_n) is convergent. Let $\lim y_n = \lim y_{n+1} = y$. Taking limits on both sides of the equation, we have:

$$\begin{aligned} y &= \sqrt{p+y} \\ y^2 &= p+y \\ y^2 - y - p &= 0 \\ \left(y - \frac{1}{2}\right)^2 - \left(p + \frac{1}{4}\right) &= 0 \\ \left(y - \frac{1}{2}\right)^2 - \left(\sqrt{p + \frac{1}{4}}\right)^2 &= 0 \\ y &= \frac{1}{2} \pm \left(\sqrt{p + \frac{1}{4}}\right) \end{aligned}$$

Thus, the sequence (y_n) converges to $(1 + \sqrt{4p+1})/2$. □

Problem 4.23. Let $x_1 := a > 0$ and $x_{n+1} := x_n + \frac{1}{x_n}$ for $n \in \mathbf{N}$. Determine whether (x_n) converges or diverges.

Proof. The first few terms of the sequence (x_n) are $a, a + \frac{1}{a}, a + \frac{1}{a} + \frac{1}{a+\frac{1}{a}}, \ldots$. Assume that the sequence (x_n) converges to the limit $\lim x_n = x$. Taking limits on both sides, we have:

$$\begin{aligned} x &= x + \frac{1}{x} \\ \frac{1}{x} &= 0 \end{aligned}$$

This equation has no real solutions. Hence, the sequence (x_n) is divergent. □

Problem 4.24. (a) In subsection 3.6.1 we used the Axiom of Completeness(AoC) to prove the Archimedean property of real numbers $\mathbf{R}$ (Theorem 3.6.1). Show that the Monotone Convergence Theorem can also be used to prove the Archimedean property without making any use of AoC.

(b) Use the Monotone Convergence Theorem to supply a proof for the Nested Interval Property (Theorem 3.6) that doesn't make use of the AoC.

These two results would suggest that we could have used the Monotone Convergence Theorem in place of the AoC as our starting axiom for building a proper theory of the real numbers.

Proof. (a) The Archimedean property states that, $\mathbf{N}$ is an unbounded set that sits in $\mathbf{R}$. Given any real number x, there exists $n \in \mathbf{N}$ such that $x < n$.

By contradiction, let us assume that $\mathbf{N}$ is bounded subset of $\mathbf{R}$. Then, there exists an upper bound $x \in \mathbf{R}$, such that $n < x$ for all natural numbers $n \in \mathbf{N}$. The natural numbers are recursively defined using the sequence $x_{n+1} = 1 + x_n$. This is a monotonically increasing sequence. Hence, $\mathbf{N}$ is monotonic increasing and bounded, by monotone convergence theorem, the sequence (x_n) is convergent. Let $\lim_{n\to\infty} x_n = x$. Taking limits on both sides, we have:

$$\begin{aligned}\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} &= \lim_{n\to\infty}(1 + x_n) = 1 + x \\ x &= 1 + x \\ 0 &= 1\end{aligned}$$

This is a false statement and violates Peano's axioms. So, (x_n) is not a convergent sequence $\implies$ (x_n) is unbounded $\implies$ $\exists n \in \mathbf{N}$, such that $x < n$ for any given real x.

(b) The nested interval property states that the real number line $\mathbf{R}$ contains no gaps. I reproduce the statement of NIP for completeness.

For each $n \in \mathbf{N}$, assume that we are given a closed interval $I_n = [a_n, b_n] = \{x \in \mathbf{R} : a_n \leq x \leq b_n\}$. Assume also that each I_n contains I_{n+1}. Then, the resulting nested sequence of closed intervals

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq I_4 \ldots$$

have a non-empty intersection; that is

$$\bigcap_{n=1}^{\infty} I_n \neq \phi$$

Since, $I_n \supseteq I_{n+1}$, that is $[a_n, b_n] \supseteq [a_{n+1}, b_{n+1}]$, the sequence consisting of the left hand end-points $a_1, a_2, a_3, \ldots, a_n$; (a_n) is monotonically increasing. Moreover, the sequence (a_n) is bounded above by b_n. By the Monotone Convergence Theorem, the sequence is convergent and $\lim_n(a_n) = s = \sup\{a_n : n \in \mathbf{N}\}$.

Since, s is an upper bound for the sequence (a_n), $a_n \leq s$ for all $n \in \mathbf{N}$. Moreover, as s is the least upper bound $s \leq b_n$ for all $n \in \mathbf{N}$. Consequently, $a_n \leq s \leq b_n$ for all $n \in \mathbf{N}$. Thus, $s \in I_n$ for all $n \in \mathbf{N}$, and the countable intersection of these intervals is non-empty. □

Problem 4.25. (Calculating Square Roots). Let $x_1 = 2$ and define

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right)$$

(a) Show that x_n^2 is always greater than or equal to 2, and then use this to prove that $x_n - x_{n+1} \geq 0$. Conclude that $\lim x_n = \sqrt{2}$.

(b) Modify the sequence (x_n) so that it converges to $\sqrt{c}$.

Proof. (a) We can express the recurrence relation as a quadration equation in x_n. We have,

$$\begin{aligned} x_{n+1} &= \frac{x_n^2+2}{2x_n} \\ 2x_{n+1}x_n &= x_n^2+2 \\ x_n^2 - 2x_{n+1}x_n + 2 &= 0 \end{aligned}$$

This equation has real roots, if the discriminant is non-negative. So, $b^2 - 4ac \geq 0$. This yields,

$$\begin{aligned} 4x_{n+1}^2 - 4(2) &\geq 0 \\ x_{n+1}^2 - 2 &\geq 0 \\ x_{n+1}^2 &\geq 2 \end{aligned}$$

Thus, the sequence (x_n) has a lower bound. Moreover, we can show that (x_n) is a monotonically decreasing sequence.

$$\begin{aligned} x_n - x_{n+1} &= x_n - \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) \\ &= x_n - \frac{x_n}{2} - \frac{1}{x_n} \\ &= \frac{x_n}{2} - \frac{1}{x_n} \\ &= \frac{x_n^2 - 2}{2x_n} \\ &\geq 0 \end{aligned}$$

where the last inequality follows from the fact, that $x_n^2 \geq 2$.

By the Monotone Convergence Theorem, the sequence (x_n) is convergent. Let $\lim x_n = x$. We have,

$$\begin{aligned} \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} &= \frac{1}{2}\left(\lim_{n\to\infty} x_n + \lim_{n\to\infty} \frac{2}{x_n}\right) \\ x &= \frac{1}{2} \cdot \frac{x^2+2}{x} \\ 2x^2 &= x^2 + 2 \\ x^2 &= 2 \\ x &= \sqrt{2} \end{aligned}$$

(b) The sequence (x_n) can be modified as,

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{c}{x_n}\right)$$

so that it converges to $\sqrt{c}$. $\square$

Problem 4.26. (Arithmetic-Geometric Mean). (a) Explain why $\sqrt{xy} \leq \frac{x+y}{2}$ for any two positive real numbers x and y. The geometric mean is always less than or equal to the arithmetic mean.

(b) Now let $0 \leq x_1 \leq y_1$ and define

$$x_{n+1} = \sqrt{x_n y_n}$$

and

$$y_{n+1} = \frac{x_n + y_n}{2}$$

Show that $\lim x_n$ and $\lim y_n$ both exist and are equal.

Proof. (a) We know that, the square of real numbers is non-negative. Consider the expression $(\sqrt{x} - \sqrt{y})^2$. We have,

$$\begin{aligned}
(\sqrt{x} - \sqrt{y})^2 &\geq 0 \\
(\sqrt{x})^2 + (\sqrt{y})^2 - 2\sqrt{xy} &\geq 0 \\
x + y &\geq 2\sqrt{xy} \\
\sqrt{xy} &\leq \frac{x+y}{2}
\end{aligned}$$

(b) The first terms of the sequence (x_n) are

$$x_1, \sqrt{x_1 y_1}, \sqrt{x_2 y_2}, \sqrt{x_3 y_3}, \ldots$$

We are interested to show that this is a monotonically increasing sequence. For $n = 1$, $\sqrt{x_1 y_1} \geq \sqrt{x_1^2} = x_1$. So, $x_2 > x_1$. Assume that $x_{k+1} > x_k$. Our claim is that $x_{k+2} \geq x_{k+1}$. We have,

$$\begin{aligned}
x_{k+2} &= \sqrt{x_{k+1} y_{k+1}} \\
&= \sqrt{\left(\sqrt{x_k y_k}\right) \cdot y_{k+1}} \\
&= \sqrt{\left(\sqrt{x_k y_k}\right) \cdot \left(\frac{x_k + y_k}{2}\right)} \\
&\geq \sqrt{\left(\sqrt{x_k y_k}\right) \cdot \left(\sqrt{x_k y_k}\right)} = \sqrt{x_k y_k} = x_{k+1}
\end{aligned}$$

By principle of mathematical induction, $x_{n+1} > x_n$ for all $n \in \mathbf{N}$.

Likewise, we can show that the sequence (y_n) is a monotonically decreasing sequence. The first few terms of this sequence are,

$$y_1, \frac{x_1 + y_1}{2}, \frac{x_2 + y_2}{2}, \ldots, \frac{x_n + y_n}{2}, \ldots$$

For $n = 1$, we find that $y_2 = (1/2)(x_1+y_1) \leq (1/2)(y_1+y_1) = y_1$. Assume that $y_{k+1} \leq y_k$. Our claim is that $y_{k+2} \leq y_{k+1}$. Using the property that, the geometric mean of two real numbers is always less than or equal to their arithmetic mean; for any $k \in \mathbf{N}$, $\sqrt{x_k y_k} \leq \frac{x_k+y_k}{2}$, so $x_k \leq y_k$. Thus,

$$\begin{aligned} y_{k+2} &= \frac{x_{k+1} + y_{k+1}}{2} \\ &\leq \frac{y_{k+1} + y_{k+1}}{2} = y_{k+1} \end{aligned}$$

It follows that the sequence (y_n) is monotonically decreasing.

Define $X := \{x_n : n \in \mathbf{N}\}$ and $Y := \{y_n : n \in \mathbf{N}\}$. It is easy to see that, y_1 is an upper bound for X, x_1 is a lower bound for Y. For all $k \in \mathbf{N}$,

$$x_1 \leq x_2 \leq ... \leq x_{k-1} \leq x_k \leq y_k \leq y_{k-1} \leq ... \leq y_1$$

By the Monotone Convergence Theorem, the two sequences (x_n) and (y_n) are convergent. And $\lim_{n\to\infty} x_n = \sup X$, $\lim_{n\to\infty} y_n = \inf Y$. We can further deduce that $\sup X = \inf Y$.

Let $\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = x$ and $\lim_{n\to\infty} y_n = \lim_{n\to\infty} y_{n+1} = y$. Then, taking limits on both sides, we have

$$\begin{aligned} \lim_{n\to\infty} x_{n+1} &= \sqrt{\lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n} \\ x &= \sqrt{xy} \\ \sqrt{x}(\sqrt{x} - \sqrt{y}) &= 0 \end{aligned}$$

Since, the sequences are positive, the limits are also positive. So, $\sqrt{x} \neq 0$. Thus, $\sqrt{x} = \sqrt{y}$ or $x = y$. $\square$

Problem 4.27. (Limit Superior). Let (a_n) be a bounded sequence.

(a) Prove that the sequence defined by $y_n = \sup\{a_k : k \geq n\}$ converges.

(b) The *limit superior* of (a_n) or $\limsup a_n$ is defined by

$$\limsup a_n = \lim y_n$$

where (y_n) is the sequence from part (a) of this exercise. Provide a reasonable definition for $\liminf a_n$ and briefly explain why it always exists for any bounded sequence.

(c) Prove that $\liminf a_n \leq \limsup a_n$ for every bounded sequence, and given an example of a sequence where this inequality is strict.

(d) Show that $\liminf a_n = \limsup a_n$ if an only if $\lim a_n$ exists. In this case, all three share the same value.

Proof. (a) The sequence (a_n) is a bounded sequence. Hence, $|a_n| \leq M$ for all $n \in \mathbf{N}$.

Claim. The sequence (y_n) is a monotonically decreasing sequence. For $n = 1$, $y_1 >= y_2$ because $y_1 = \sup\{a_k : k \geq 1\} = \max\{a_1, y_2\} \geq y_2$. Assume that $y_k \geq y_{k+1}$. Then, $y_{k+1} = \sup\{a_k : k+1 \geq n\} = \max\{a_{k+1}, y_{k+2}\} \geq y_{k+2}$. Hence, $y_k \geq y_{k+1}$ for all $k \in \mathbf{N}$.

We can further show that (y_n) is a bounded sequence. Since, $|a_n| \leq M$ for all $n \in \mathbf{N}$, the supremum of any of the subsequences $\{a_k : k \geq n\}$ no matter what $k \in \mathbf{N}$, is always less than or equal to M. Thus, $|y_n| \leq M$ for all $n \in \mathbf{N}$. Therefore, (y_n) is a bounded sequence.

By the Monotone Convergence Theorem, the sequence (y_n) is convergent.

(b) From part (a), we can deduce that, the limit superior of a sequence is,

$$\limsup a_n = \lim_{n\to\infty} \left(\sup\{a_k : k \geq n\}\right) = \inf\{\sup\{a_k : k \geq n\}\}$$

whereas, the limit inferior of a sequence is,

$$\liminf a_n = \lim_{n\to\infty} \left(\inf\{a_k : k \geq n\}\right) = \sup\{\inf\{a_k : k \geq n\}\}$$

The above definition is easy to verify.

Define $z_n := \inf\{a_k : k \geq n\}$. Let us prove that (z_n) is a convergent sequence.

Claim. The sequence (z_n) is a monotonically increasing sequence. For $n = 1$, $z_1 \leq z_2$ because $z_1 = \inf\{a_k : k \geq 1\} = \min\{a_1, z_2\} \leq z_2$. Assume that $z_k \leq z_{k+1}$. We have, $z_{k+1} = \inf\{a_{k+1} : k+1 \geq n\} = \min\{a_{k+1}, z_{k+2}\} \leq z_{k+2}$. Hence, by principle of mathematical induction, $z_n \leq z_{n+1}$ for all $n \in \mathbf{N}$.

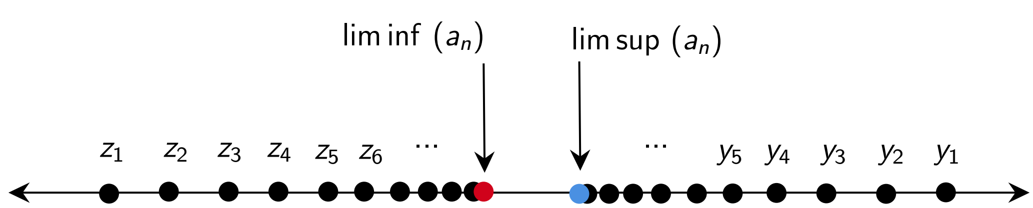
Claim. The sequence (z_n) is bounded below. Since, (a_n) is a bounded sequence, $a_n \geq -M$ for all $n \in \mathbf{N}$. So, given any $k \in \mathbf{N}$, the infimum of the subsequence $\{a_k : k \geq n\}$ will be greater than or equal to $-M$. So, $-M$ is a lower bound for the sequence (z_n).

By the Monotone Convergence Theorem, (z_n) is convergent sequence.

(c) (y_n) and (z_n) are convergent sequences; their limiting value is the infimum/supremum. So, we are justified in setting,

$$y = \inf\{y_n : n \in \mathbf{N}\}$$
$$z = \sup\{z_m : m \in \mathbf{N}\}$$

Observe that, the infimum of k-tail of the sequence is always less than or equal to the supremum of the tail. Hence, $z_k \leq y_k$ for all $k \in \mathbf{N}$.



The visual representation makes it clear that, every y_n is an upper bound for the sequence $\{z_m : m \in \mathbf{N}\}$. So, $\inf\{y_n : n \in \mathbf{N}\}$ is also an upper bound for the set Z. Thus, $\liminf a_n \leq \limsup a_n$.

Consider the sequence

$$a_n := \left(\frac{n+1}{n}\right) \sin\left(\frac{n\pi}{8}\right)$$

This sequence is bounded, but not convergent, and so a strict inequality holds.

(d) Assume that $\limsup a_n = \liminf a_n = l$. Pick an abritrary $\epsilon > 0$. There exists points N_1, N_2, such that,

if $n \geq N_1$

$$\sup\{a_k : k \geq n\} - l \leq \epsilon$$
$$a_k \leq \sup\{a_k : k \geq n\} \leq l + \epsilon$$

if $n \geq N_2$

$$-\epsilon \leq \inf\{a_k : k \geq n\} - l$$
$$l - \epsilon \leq \inf\{a_k : k \geq n\} \leq a_k$$

Therefore, if $n \geq \max\{N_1, N_2\}$, we have,

$$l - \epsilon \leq \inf\{a_k : k \geq n\} \leq a_k \leq \sup\{a_k : k \geq n\} \leq l + \epsilon$$
$$l - \epsilon \leq a_k \leq l + \epsilon$$

Thus, $\lim_{n\to\infty} a_n = l$. □

The below code snippet is a really nice visualization of the limit superior and limit inferior of a sequence.

```
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import style
import math

# The sequence x_n = (n+1)/n sin(n/8)
n = np.linspace(1,500,500)
x_n = np.multiply(np.divide(n+10,n),np.sin(n/8))

# The supremum of a subsequence
def sup(x_m,n):
    return max(x_m[int(n-1):])
```

```
# The infimum of a subsequence
def inf(x_m,n):
    return min(x_m[int(n-1):])

limsup = [sup(x_n,m) for m in n]
liminf = [inf(x_n,m) for m in n]

plt.style.use('fivethirtyeight')
plt.figure(figsize=(20,10))
plt.xlim(0,450)
plt.text(250,1.2,r'$\sup \{x_m: m > n\}$')
plt.text(250,-1.2,r'$\inf \{x_m: m > n\}$')
plt.plot(n,x_n,'o')
plt.plot(n,limsup)
plt.plot(n,liminf)
plt.show()
```

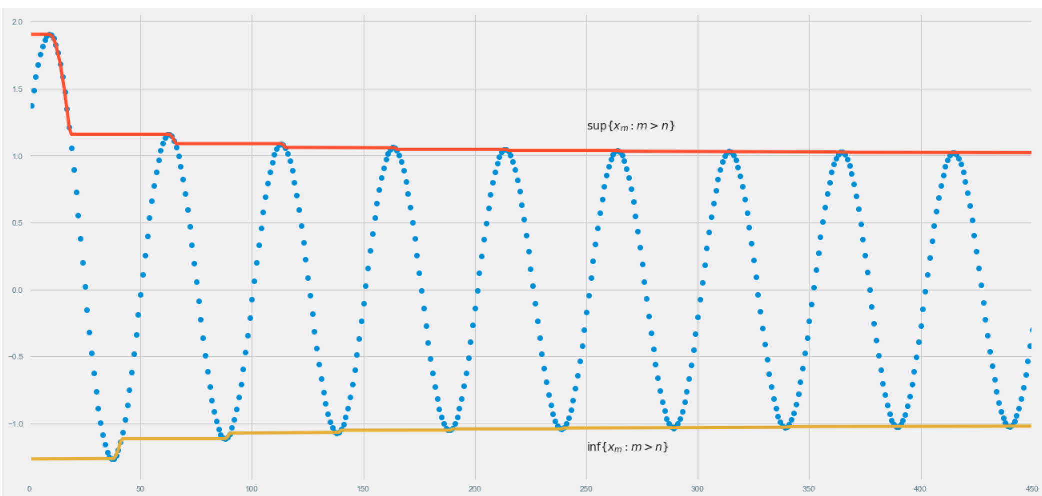


Fig. The limit superior and limit inferior of a bounded sequence (x_n)

Problem 4.28. For each series, find an explicit formula for the sequence of partial sums and determine if the series converges.

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

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

(c) $\sum_{n=1}^{\infty} \log\left(\frac{n+1}{n}\right)$

Proof. (a) The partial sum s_m is given by,

$$\begin{aligned} s_m &= \frac{1}{2} + \frac{1}{2^2} + ... + \frac{1}{2^m} \\ &= \frac{(1/2)(1-(1/2)^m)}{1-(1/2)} \\ &= 1 - \frac{1}{2^m} \end{aligned}$$

The sequence (s_m) is monotonically increasing because $s_{m+1} = 1 - \frac{1}{2^{m+1}} > 1 - \frac{1}{2^m} = s_m$. Moreover, the sequence (s_m) is bounded, since,

$$\begin{aligned} s_m &= 1 - \frac{1}{2^m} \\ &= \frac{2^m - 1}{2^m} \\ &< \frac{2^m}{2^m} = 1 \end{aligned}$$

By the Monotone Convergence Theorem, (s_m) is convergent and the sum of the infinite series is given by,

$$\begin{aligned} \lim_{m\to\infty} s_m &= 1 - \lim_{m\to\infty} \left(\frac{1}{2^m}\right) \\ &= 1 \end{aligned}$$

(b) The partial sum s_m is given by,

$$\begin{aligned} s_m &= \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + ... + \frac{1}{m(m+1)} \\ &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + ... + \left(\frac{1}{m} - \frac{1}{m+1}\right) \\ &= 1 - \frac{1}{m+1} \end{aligned}$$

The sequence (s_m) is monotonically increasing, since $s_{m+1} = 1 - 1/(m+1) > 1 - 1/m = s_m$. Moreover, the sequence (s_m) is bounded above by 1. Hence, by the Monotone Convergence Theorem, the sequence (x_m) is convergent. The sum of the infinite series is given by,

$$\begin{aligned} \lim_{m\to\infty} s_m &= 1 - \lim_{m\to\infty} \left(\frac{1}{2^m}\right) \\ &= 1 \end{aligned}$$

(c) The partial sum s_m is given by,

$$\begin{aligned} s_m &= \log\left(\frac{2}{1}\right) + \log\left(\frac{3}{2}\right) + \log\left(\frac{4}{3}\right) + \ldots + \log\left(\frac{m+1}{m}\right) \\ &= \log(m+1) \end{aligned}$$

The sequence (s_m) is unbounded, and hence the sum of the infinite series is divergent. □

Problem 4.29. Complete the proof of the theorem 4.4 by showing that if the series $\sum_{n=0}^{\infty} 2^n b_{2^n}$ diverges, then so does $\sum_{n=1}^{\infty} b_n$.

Proof. □

4.5 Subsequences and the Bolzano Weierstrass Theorem.

We have seen that the sequence of the partial sums (s_m) of the harmonic series does not converge by focusing our attention on a particular subsequence s_{2^k} of the original sequence. For the moment, we will put the topic of infinite series aside and more fully develop the important concept of subsequences.

Definition 4.10. Let (a_n) be a sequence of real numbers, and let $n_1 < n_2 < n_3 < n_4 < n_5 < \ldots$ be an increasing sequence of natural numbers. Then the sequence

$$(a_{n_1}, a_{n_2}, a_{n_3}, a_{n_4}, a_{n_5}, \ldots)$$

is called a subsequence of (a_n) and is denoted by (a_{n_k}), where $k \in \mathbf{N}$ indexes the subsequence.

Notice that the order of the terms in a subsequence is the same as that in the original sequence, and repetitions are not allowed. Thus, if

$$(a_n) = \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \ldots\right)$$

then

$$\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \ldots\right)$$

and

$$\left(\frac{1}{10}, \frac{1}{100}, \frac{1}{1000}, \frac{1}{10000}, \ldots\right)$$

are examples of legitimate subsequences, whereas

$$\left(\frac{1}{10}, \frac{1}{5}, \frac{1}{100}, \frac{1}{50}, \frac{1}{1000}, \frac{1}{500}, \ldots\right)$$

and

$$\left(1, 1, \frac{1}{3}, \frac{1}{3}, \frac{1}{5}, \frac{1}{5}, \ldots\right)$$

are not.

Theorem: 4.5.1: Subsequences of a convergent sequence.

Subsequences of a convergent sequence converge to the same limit as the original sequence.

Proof. Assume that $(a_n) \to a$ and let (a_{n_k}) be a subsequence. Given $\epsilon > 0$, there exists a positive number $N \in \mathbf{N}$ such that $|a_n - a| < \epsilon$, for all $n \geq N$. Because $n_k \geq k$ for all k, the same N will suffice for the subsequence; that is $|a_{n_k} - a| < \epsilon$ for $k \geq N$. □

This not too surprising result has several somewhat surprising applications. It is the key ingredient for understanding when infinite sums are associative. We can also use it in the following clever way to compute the values of some familiar limits.

Example 4.5.1. Let $0 < b < 1$. Because,

$$b > b^2 > b^3 > b^4 > \ldots > 0$$

the sequence (b^n) is decreasing and bounded below. The Monotone Convergence Theorem allows us to conclude that (b^n) converges to some l satisfying $b > l \geq 0$. To compute l, notice that (b^{2n}) is a subsequence, so $(b^{2n}) \to l$ by theorem 4.5. But, $b^{2n} = b^n \cdot b^n$, so by the algebraic limit theorem, $l^2 = l$ and thus $l(l-1) = 0$, so $l = 0$.

Without much trouble, we can generalize this example to conclude $(b^n) \to 0$ if and only if $-1 < b < 1$.

Example 4.5.2. (Divergence Criterion). The theorem 4.5 is also useful for providing economical proofs for divergence. In Example 4.2.4, we were quite sure that,

$$\left(1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \frac{1}{5}, -\frac{1}{5}, \frac{1}{5}, -\frac{1}{5}, \frac{1}{5}, -\frac{1}{5}, \frac{1}{5}, -\frac{1}{5}, \frac{1}{5}, -\frac{1}{5}, \ldots\right)$$

did not converge to any proposed limit. Notice that,

$$\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \ldots\right)$$

is a subsequence that converges to $1/5$. Also,

$$\left(-\frac{1}{5}, -\frac{1}{5}, -\frac{1}{5}, -\frac{1}{5}, -\frac{1}{5}, \ldots\right)$$

is a different subsequence of the original sequence that converges to $-1/5$. Because we have two subsequences converging to two different limits, we can rigorously conclude that the original sequence diverges.

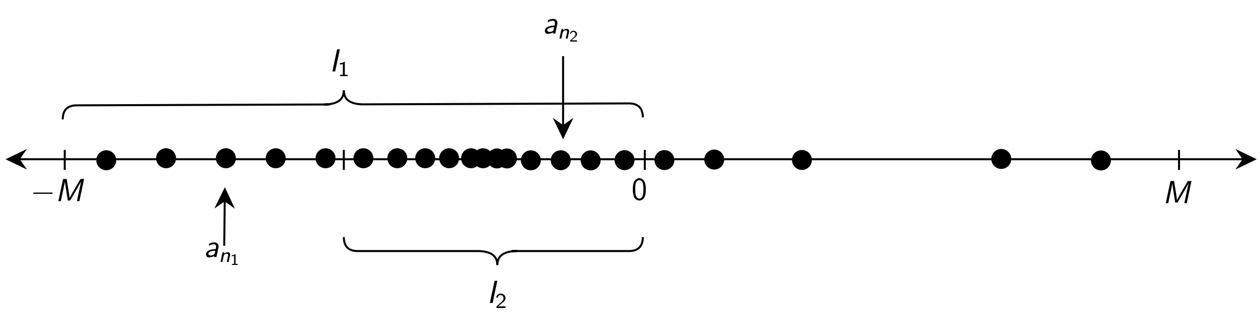
4.5.1 The Bolzano-Weierstrass Theorem.

In the previous example, it was rather easy to spot a convergent subsequence (or two) hiding in the original sequence. For bounded sequences, it turns out that it is always possible to find atleast one such convergent subsequence.

Theorem: 4.5.2: Bolzano-Weierstrass Theorem.

Every bounded sequence contains a convergent subsequence.

Proof. Let (a_n) be a bounded sequence, so there exists $M > 0$ satisfying $|a_n| < M$ for all $n \in \mathbf{N}$. Bisect the closed interval $[-M, M]$ into two closed intervals $[-M, 0]$ and $[0, M]$. (The midpoint is included in both halves.) Now, it must be that atleast one of these closed intervals contains an infinite number of the terms in the sequence (a_n). Select a half for which this is the case and label that interval as I_1. Then, let a_{n_1} be some term in the sequence (a_n) satisfying $a_{n_1} \in I_1$.



Next, we bisect I_1 into closed intervals of equal length, and let I_2 be a half that again contains an infinite number of terms of the original sequence. Because, there are an infinite number from (a_n) to choose from, we can select an $a_{n_2} \in I_2$. In general, we construct the closed interval I_k by taking a half of I_{k-1} containing an infinite number of terms of (a_n) and then select $n_k > n_{k-1} > \ldots > n_2 > n_1$, so that $a_{n_k} \in I_k$.

We want to argue that (a_{n_k}) is a convergent subsequence, but we need a candidate for the limit. The sets

$$I_1 \supseteq I_2 \supseteq I_3 \ldots$$

form a nested sequence of closed intervals, and by the nested interval property there exists atleast one point $x \in \mathbf{R}$, contained in every I_k. This provides us with the candidate that we were looking for. It just remains to show that $a_{n_k} \to x$.

Let $\epsilon > 0$. By construction, the length of I_k is $M(1/2)^{k-1}$ which converges to zero (This follows from example 4.5.1 and the Algebraic Limit Theorem). Choose N so that $k \geq N$ implies that the length of I_k is less than ϵ. Because, x and a_{n_k} are both in I_k, if follows that $|a_{n_k} - x| < \epsilon$. □

Problem 4.30. Give an example of each of the following, or argue that such a request is impossible.

(a) A sequence that has a subsequence that is bounded but contains no subsequence that converges.

(b) A sequence that does not contain 0 or 1 as a term but contains subsequences converging to each of these values.

(c) A sequence that contains subsequences converging to every point in the infinite set $\{1, 1/2, 1/3, 1/4, 1/5, ...\}$.

(d) A sequence that contains subsequences converging to every point in the infinite set $\{1, 1/2, 1/3, 1/4, 1/5, ...\}$ and no subsequences converging to points outside the set.

Proof. (a) This request is impossible. If a sequence (a_n) has a subsequence (a_{n_k}) that is bounded, then because (a_{n_k}) is bounded, it must have atleast one convergent subsequence by the Bolzano Weierstrass theorem.

(b) Consider the sequence (a_n) given by,

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{4}{5}, \frac{1}{6}, \frac{5}{6}, \ldots\right)$$

given by

$$a_m := \begin{cases} \frac{1}{m+1} & \text{if } n = 2m - 1 \\ 1 - \frac{1}{m+1} & \text{if } n = 2m \end{cases}$$

This sequence does not contain 0 or 1 as a term, but contains subsequences converging to each of these values.

(c) Suppose the sequence (a_n) is given by,

$$\left(1, 1, \frac{1}{2}, 1, \frac{1}{2}, \frac{1}{3}, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right)$$

In particular,

$$a_{\frac{m(m-1)}{2}+n} := \frac{1}{n}$$

where $0 < n \leq m$.

This sequence contains subsequences converging to each of the values in the set $\{1, 1/2, 1/3, 1/4, 1/5, ...\}$.

(d) This request is impossible. If we build a sequence of the form

$$\left(1, \frac{1}{2}, 1, \frac{1}{3}, 1, \frac{1}{4}, 1, \frac{1}{5}, ...\right)$$

it is easy to see that the subsequence $1/2, 1/3, 1/4, 1/5, ...$ converges to 0, which is a point outside the set. □

Problem 4.31. Decide whether the following propositions are true or false, providing a short justification for each conclusion.

(a) If every proper subsequence of (x_n) converges, then (x_n) converges as well.

(b) If (x_n) contains a divergent subsequence, then (x_n) diverges.

(c) If (x_n) is bounded and diverges, then there exist two subsequences of (x_n) that converge to different limits.

(d) If (x_n) is monotone and contains a convergent subsequence, then (x_n) converges.

Proof. (a) This proposition is true.

Let (x_n) be a sequence. Assume that all proper subsequences are convergent. Consider the subsequences (x_{2n}), (x_{2n+1}) and (x_{3n}). Suppose that these subsequences converge to the limits a, b and c. Then, the subsequence (x_{6n}) can be extracted from both (x_{2n}) and (x_{3n}) and it is convergent, so $a = c$. Moreover, (x_{6n+3}) can be extracted from both (x_{2n+1}) and (x_{3n}), and since it is convergent, by the uniqueness of limits, $b = c$.

(b) This proposition is true.

Justification. If a subsequence of (x_n) diverges, it implies that, given any $\epsilon > 0$, there are an infinite number of real numbers in the tail that do not lie in an ϵ-neighbourhood of a limit point. So, the original sequence is divergent.

(c) This proposition is true.

Justification. Assume that the sequence (x_n) is bounded, but divergent. By definition, there exists large number M, such that $|x_n| \leq M$ for all $n \in \mathbf{N}$. Consider the subsequences

$$\begin{aligned} (x_1, x_2, x_3, x_4, ...) \\ (x_2, x_3, x_4, ...) \\ (x_3, x_4, ...) \\ \vdots \end{aligned}$$

Each of these subsequences are bounded subsets of $\mathbf{R}$ and therefore by AoC, it is reasonable to talk about their supremum and infimum. Define the subsequences,

$$u_n := \sup\{x_k : k \geq n\}$$
$$d_n := \inf\{x_k : k \geq n\}$$

u_n is bounded and monotonically decreasing. d_n is bounded and monotonically increasing. By the Montone Convergence theorem, u_n and d_n are convergent.

(d) This proposition is true.

Justification. Assume that the subsequence (x_{n_k}) converges to x. Then, given an $\epsilon > 0$, there exists N such that x_{n_k} lies in the interval $(x - \epsilon, x + \epsilon)$ for all $n \geq N$. Because (x_n) is monotone, the values of the sequence (x_n) between x_{n_k} and $x_{n_{k+1}}$ also lie in $(x - \epsilon, x + \epsilon)$. Consequently, the same N can be used for the original sequence. □

Problem 4.32. (a) Prove that if an infinite series converges, then the associative property holds. Assume that $a_1 + a_2 + a_3 + a_4 + a_5 + ...$ converges to a limit L (that is the sequence of partial sums $(s_n) \to L$. Show that any regrouping of the terms

$$(a_1 + a_2 + ... + a_{n_1}) + (a_{n_1+1} + ... + a_{n_2}) + (a_{n_2+1} + ... + a_{n_3}) + ...$$

leads to a series that also converges to L.

(b) Compare this result to example discussed earlier, where infinite addition was shown not to be associative. Why doesn't our proof in (a) apply to this example?

Proof. (a) Assume that the infinite series

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

converges to a finite limit L. Thus, the sequence of the partial sums (s_n) converges to L.

Let $n_1 < n_2 < n_2 < ... < n_k$ be an increasing sequence of natural numbers. Thus,

$$(s_{n_1}, s_{n_2}, s_{n_3}, \dots, s_{n_k}, ...)$$

represents a subsequence (extraction) of (s_n), where,

$$s_{n_k} = (a_1 + a_2 + ... + a_{n_1}) + (a_{n_1+1} + ... + a_{n_2}) + (a_{n_2+1} + ... + a_{n_3}) + ... + (a_{n_{k-1}+1}, \dots, a_{n_k})$$

If a sequence (x_n) converges to the limit L, then all its subsequences converge to the same limiting value. Thus, the subsequence (s_{n_k}) converges to the same limit L, for any initial choice of $n_1, n_2, ...$ Thus, any regrouping of terms leads to a series that also converges to L.

(b) The series

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

is divergent. For a divergent series, all subsequences of the partial sums do not converge to the same limit value. Therefore, for a divergent series, infinite addition is not associative. □

Problem 4.33. The Bolzano-Weierstrass theorem is extremely important, and so is the strategy employed in the proof. To gain some more experience with this technique, assume that the nested interval property is true and use it provide a proof of the Axiom of Completeness. To prevent the argument from being circular, assume also that $(1/2^n) \to 0$. (Why precisely is this last argument needed to avoid circularity?)

Proof. Claim. Every non-empty subset of real numbers that is bounded above has a least-upper bound.

Let A be a subset of $\mathbf{R}$ bounded above by M. It could very well be the case that $A \cap M = \emptyset$. We start with an interval $I_1 = [a_1, b_1]$ such that $a_1 \in A$ and $b_1 = M$.

Bisect the interval I_1 into two closed intervals L_1 and R_1 of equal length. Now, it must be the case that atleast one of these closed intervals contains an upper bound for A. We need to make sure, that we select a sub-interval that contains point of A. We choose the half-interval and label it I_2 as follows,

$$I_2 = \begin{cases} R_1 & \text{if } A \cap R_1 \neq \emptyset \\ L_1 & \text{otherwise} \end{cases}$$

In general, we construct a closed interval $I_{k+1} = [a_{k+1}, b_{k+1}]$ by taking a half of $I_k = [a_k, b_k]$, always containing points of A and an upper bound for A.

$$I_{k+1} = \begin{cases} R_k & \text{if } A \cap R_k \neq \emptyset \\ L_k & \text{otherwise} \end{cases}$$

The set

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq \ldots$$

forms a nested sequence of intervals. By the nested interval property, there exists atleast one point $x \in \bigcap_{k\geq 1} I_k$. This provides us the candidate that we are looking for. It just remains to show that the sequence $x = \sup A$.

To show that $x = \sup A$, we first prove that x is an upper bound for A. Remember that $x \in \bigcap_{k\geq 1} I_k$. By construction, every b_k is an upper bound for A. If there were some $y \in A$, such that $y > x$, then there would be a k, such that $b_k < y$. That would imply that b_k is not an upper bound for A. This leads to a contradiction, so x must be an upper bound for A.

Next, we show that x is the least upper bound. If there were y, such that $y < x$, then there would be a $a_k \in I_k \cap A$, such that $a_k > y$. So, y cannot be an upper bound for A. Thus, x is the least upper bound for A. $\square$

Problem 4.34. Assume that (a_n) is a bounded sequence with the property that every convergent subsequence of (a_n) converges to the same limit $a \in \mathbf{R}$. Show that (a_n) must converge to a.

Proof. Assume that all convergent subsequences converge to the limit a. Then, if $\epsilon > 0$, for any subsequence $(a_{f(n)})$, there exists an $N_j \in \mathbf{N}$, such that, we can make the distance $\left|a_{f(n)} - a\right|$ as small as we like. That is,

$$\left|a_{f(n)} - a\right| < \epsilon$$

for all $n \geq N_j$.

Define $N = \max\{N_1, N_2, ...\}$. If $n \geq N$, then

$$|a_n - a| < \epsilon$$

So, the sequence (a_n) converges to the limiting value a. $\square$

Problem 4.35. Use a similar strategy to the one in example 4.5.1 to show that $\lim b^{1/n}$ exists for all $b \geq 0$ and find the value of the limit. (The results in exercise

Proof. Let $b_n = b^{\frac{1}{n}}$; $b \geq 0$. Then, because

$$b > b^{1/2} > b^{1/3} > ... > b^{1/n}$$

(b_n) is a monotonically decreasing sequence. Moreover, $b^{1/n} \geq 0$ for all $n \in \mathbf{N}$. So, (b_n) is bounded below.

By the Monotone Convergence Theorem, the sequence (b_n) converges. To compute l, notice that (b_{2n}) is a subsequence of (b_n), so $(b_{2n}) \to l$. But, $b^{\frac{1}{2n}} = (b^{\frac{1}{n}})^{1/2}$. Taking limits on both sides, we have:

$$\begin{aligned}\lim_{n\to\infty} b_{2n} &= \lim_{n\to\infty} b_n \\ \sqrt{l} &= l \\ \sqrt{l} &= 1\end{aligned}$$

Thus, $(b_n) \to 1$. $\square$

Problem 4.36. Extend the result proved in example 4.5.1 to the case $|b| < 1$; that is show that $\lim(b^n) = 0$ if and only if $-1 < b < 1$.

Proof. Suppose $-1 < b \leq 0$. For all $n \in \mathbf{N}$, we have $|b_n| \leq 1$. Thus, (b_n) is a bounded sequence.

Consider the two subsequences b_{2n} and b_{2n+1}. The subsequence consisting of all even powers of b satisfies:

$$b^0 > b^2 > b^4 > b^6 > ...$$

Thus, (b_{2n}) is a monotonically decreasing sequence. (b_{2n}) is bounded below by 0. By the Montonone convergence theorem, (b_{2n}) is convergent, and its limiting value is 0.

The subsequence consisting of all odd powers of b satisfies:

$$b^1 < b^3 < b^5 < b^7 < ...$$

Thus, (b_{2n+1}) is a monotonically increasing sequence. (b_{2n}) is bounded above by 0. By the Montonone convergence theorem, (b_{2n+1}) is convergent, and its limiting value is 0.

Since (b_n) is a bounded sequence with the property that the subsequences (b_{2n}) and (b_{2n+1}) converge to the same limit 0, the original sequence (b_n) also converges to 0. $\square$

Problem 4.37. Another way to prove the Bolzano-Weierstrass Theorem is to show that every sequence contains a monotone subsequence. A useful device in this endeavor is the notion of a *peak term*. Given a sequence (x_n), a particular term x_m is a peak term if no later term in the sequence exceeds it; i.e. if $x_m \geq x_n$ for all $n \geq m$.

(a) Find examples of sequences with zero, one and two peak terms. Find an example of a sequence with infinitely many peak terms that is not monotone.

(b) Show that every sequence contains a monotone sequence and explain how this furnishes a new proof of the Bolzano-Weierstrass Theorem.

4.6 The Cauchy Criterion.

The following definition bears a striking resemblance to the definition of convergence for a sequence.

Definition 4.11. A sequence (a_n) is called a *Cauchy sequence*, if, for every $\epsilon > 0$, there exists an $N \in \mathbf{N}$ such that whenever $m, n \geq N$ it follows that $|a_m - a_n| < \epsilon$.

To make the comparison easier, let's restate the definition of convergence.

Definition 4.12. A sequence (a_n) converges to a real number a if, for every $\epsilon > 0$, there exists an $N \in \mathbf{N}$, such that whenever $n \geq N$, it follows that $|a_n - a| < \epsilon$.

As we have discussed, the definition of convergence asserts that, given an arbitrary positive $\epsilon > 0$, it is possible to find a point in the sequence after which the terms of the sequence are all closer to the limit a than the given ϵ. On the other hand, a sequence is a Cauchy sequence if, for every ϵ, there is a point in the sequence after which the terms are closer *to each other* than a given ϵ. To spoil the surprise, we will argue in this section that in fact these two definitions are equivalent: Convergent sequences are Cauchy sequences, and Cauchy sequences converge. The significance of the definition of a Cauchy sequence is that there is no mention of a limit. This is somewhat like the situation with the Monotone Convergence Theorem in that we will have another way of proving that sequences converge without having any explicit knowledge of what the limit might be.

***Theorem:* 4.6.1: Convergent sequences are Cauchy sequences.**

Every convergent sequence is a Cauchy sequence.

Proof. Assume that (x_n) converges to x. To prove that (x_n) is Cauchy, we must find a point in the sequence after which we have $|x_n - x_m| < \epsilon$. This can be done using an application of the triangle inequality.

Given any $\epsilon > 0$, there exists $N \in \mathbf{N}$, such that for $m, n \geq N$ $|x_n - x| < \epsilon/2$ and $|x_m - x| < \epsilon/2$. Therefore, we have:

$$\begin{aligned} |x_n - x_m| &= |(x_n - x) - |x_m - x|| \\ &\leq |x_n - x| + |x_m - x| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

Thus, (x_n) is a Cauchy sequence. □

***Theorem:* 4.6.2: Cauchy sequences are bounded.**

Cauchy sequences are bounded.

Proof. Given $\epsilon = 1$, there exists an N such that $|x_m - x_n| < 1$ for all $m, n \geq N$. Thus, we must have that $|x_n| < |x_N| + 1$ for all $n \geq N$. It follows that

$$M = \max\{|x_1|, |x_2|, |x_3|, \ldots, |x_{N-1}, |x_N + 1||$$

is a bound for the sequence (x_n). □

***Theorem:* 4.6.3: Cauchy Criterion**

A sequence converges if and only if it is a *Cauchy sequence*.

Proof. ($\rightarrow$ direction) This direction is in theorem 4.6.

($\leftarrow$ direction) For this direction, we start with a Cauchy sequence (x_n). Theorem (4.6) guarantees that (x_n) is bounded, so we may use the Bolzano-Weierstrass theorem to produce a convergent subsequence (x_{n_k}). Set

$$x = \lim x_{n_k}$$

The idea is to show that the original sequence (x_n) converges to the same limit. Once again, we will use the triangle inequality argument. We know, that the terms in the subsequence are getting close to the limit x, and the assumption that the sequence (x_n) is Cauchy implies the terms in the tail of the sequence are close to each other. Thus, we want to make each of these distances less than half of the prescribed ϵ.

Let $\epsilon > 0$. Because (x_n) is Cauchy, there exists N_1 such that

$$|x_n - x_m| < \frac{\epsilon}{2}$$

whenever $m, n \geq N_1$. Now, we also know that $(x_{n_k}) \rightarrow x$, so there exists $N_2 \in \mathbf{N}$ such that,

$$|x_{n_k} - x| < \frac{\epsilon}{2}$$

Choose a term x_{n_K} in the subsequence (x_{n_k}), so that $n_K \geq N_2$.

Let $N = \max\{N_1, N_2\}$. To see that N has the desired property (for the original sequence (x_n)), observe that if $n \geq N$, then

$$\begin{aligned} |x_n - x| = |x_n - x_{n_K} + x_{n_K} - x| \\ &\leq |x_n - x_{n_K}| + |x_{n_K} - x| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

□

The Cauchy criterion is named after the French mathematician Augustin Louis Cauchy. Cauchy is a major figure in the history of many branches of Mathematics - number theory and the theory of finite groups, to name a few - but he is most widely recognised for his enormous contributions in analysis, especially Complex Analysis. He is deservedly credited with inventing the ϵ-based definitions of limits we use today, although it is probably better to view him as a pioneer of analysis in the sense that his work did not attain the level of refinement that modern mathematicians have to come to expect. The Cauchy criterion, for instance, was devised and used by Cauchy to study infinite series, but he never actually proved it in both directions. The fact that there gaps in Cauchy's work should not diminish his brilliance in any way. The issues of the day were both difficult and subtle, and Cauchy was far and away the most influential in laying the groundwork for modern standards of rigor. Karl Weierstrass played a major role in sharpening Cauchy's arguments. We will hear a good deal more from Weierstrass, most notably in further chapters when we take up uniform convergence. Bernard Bolzano was working in Prague and was writing and thinking about many of these same issues surrounding limits and continuity. Because his work was not widely available to the rest of the mathematical community, his historical reputation never achieved the distinction that his impressive accomplishments would seem to merit.

4.6.1 Completeness revisited.

In the first chapter, we established the Axiom Of Completeness (AoC) to be the assertion that nonempty sets bounded above has least upper bounds. We then used this axiom as the crucial step in the proof of the Nested Interval Property (NIP). In this chapter, AoC was the central step in the Monotone Convergence Theorem (MCT), and NIP was the key to proving the Bolzano-Weierstrass theorem (BW). Finally, we needed BW in our proof of the Cauchy Criterion (CC) for convergent sequences. The list of implications then looks like

$$AoC \implies \begin{cases} NIP \\ MCT \end{cases} \implies BW \implies CC$$

But, this one-directional list is not the whole story. Recall that in our original discussions about completeness, the fundamental problem was that the rational numbers contained *gaps*. The reason for moving from the rational numbers to the real numbers to do analysis is so that when we encounter a sequence that looks as if it is converging to some number - say $\sqrt{2}$ - then we can be assured that there is indeed a number there that we can call the limit. The assertion that *non-empty sets bounded above have least upper bounds* is simply one way to mathematically articulate our insistence that there be no holes in our ordered field, but it is not the only way. Instead, we could have taken MCT to be our defining axiom and used it to prove NIP and the existence of the least upper bounds.

How about NIP? Could this property serve as the starting point for a proper axiomatic treatment of the real numbers? Almost. In exercise 4.33, we showed that NIP implies AoC, but to prevent the argument from implicit use of AoC, we needed an extra assumption that is equivalent to the Archimedean property. This extra hypotheses is unavoidable. Whereas AoC and MCT can both be used to prove that $\mathbf{N}$ is not a bounded subset of $\mathbf{R}$, there is no way to prove this same fact from NIP. The upshot is that NIP is a perfectly reasonable candidate to use as the fundamental axiom of real numbers, provided that we also include the Archimedean property as a second unproven assumption.

In fact, if we assume the Archimedean Property holds, then AoC, NIP, MCT, BW and CC are equivalent in the sense that once we take any one of them to be true, it is possible to derive the other four. However, because we have an example of an ordered field that is not complete - namely, the set of rational numbers - we know it is impossible to prove any of them using only the field and the order properties. Just how we decide which should be the axiom and which then become theorems depends largely on preference and context, and in the end is not especially significant. What is important is that we understand all of these results as belonging to the same family, each asserting the completeness of $\mathbf{R}$ in its own particlar language.

Problem 4.38. Give an example of each of the following, or argue that such a request is impossible.

(a) A Cauchy sequence that is not monotone.

(b) A Cauchy sequence with an unbounded subsequence.

(c) A divergent monotone sequence with a Cauchy subsequence.

(d) An unbounded sequence containing a subsequence that is Cauchy.

Proof. (a) By CC(Cauchy Criterion),

$$\text{Cauchy sequences} \iff \text{Convergent sequences}$$

A convergent sequence could be monotone or oscillating. Consider the sequence (a_n) given by,

$$a_n = \frac{1}{n}\sin\left(\frac{n\pi}{4}\right)$$

is an example of a Cauchy sequence that is not monotone.

(b) This request is impossible. Cauchy sequences are bounded. All subsequences of a bounded sequence must be bounded.

(c) This request is impossible. By MCT, a monotone and bounded sequence is convergent. The contrapositive of this statement is; if a sequence is divergent, atleast one of the two possibilities must hold (1) the sequence is oscillating (2) the sequence is unbounded. As we are told, the sequence is monotone, it must be unbounded. An unbounded montone sequence a_n cannot have, for instance a bounded subsequence a_{2n} or $a_{f(n)}$.

The assertion that an unbounded monotone sequence cannot have a bounded subsequence needs proof.

Let (a_n) be unbounded montone; implying for any $M > 0$ there exists $N \in \mathbf{N}$, such that $|a_n| > M$ for all $n > N$. Let (a_{n_k}) be a subsequence. Assume that the subsequence is bounded by b.

Choose $M > b$, then because (a_n) is unbounded, there exists an N_b such that $|a_{N_b}| > b$. If $k \geq N_b$, then $|a_{n_k}| \geq a_k \geq a_{N_b} > b$ (given a_n is monotonic increasing). This is a contradiction. Our initial assumption is wrong.

(d) Consider a sequence (x_n) with the following subsequences: $x_{2n} = n$ and $x_{2n+1} = 0$. This is an unbounded sequence containing a Cauchy subsequence. In fact any interlacing of an unbounded sequence with a Cauchy sequence should give a counterexample. $\square$

Problem 4.39. If (x_n) and (y_n) are Cauchy sequences, then one easy way to prove that $(x_n + y_n)$ is Cauchy is to use the Cauchy criterion. By CC, (x_n) and (y_n) must be convergent, and the Algebraic Limit theorem then implies that $(x_n + y_n)$ is convergent and Cauchy.

(a) Give a direct argument that $(x_n + y_n)$ is a Cauchy sequence that does not use the Cauchy criterion or the Algebraic Limit Theorem.

(b) Do the same for the product $(x_n y_n)$

Proof. (a) Pick an $\epsilon > 0$. By definition of Cauchy sequences,

(1) There exists an $N_1 \in \mathbf{N}$, such that for $m, n \geq N_1$

$$|x_n - x_m| < \epsilon/2$$

(2) There exists an $N_2 \in \mathbf{N}$, such that for $m, n \geq N_1$

$$|y_n - y_m| < \epsilon/2$$

Define $N = \max\{N_1, N_2\}$. Then, for all $m, n \geq N$, we have:

$$\begin{aligned}
|(x_n + y_n) - (x_m + y_m)| &= |(x_n - x_m) + (y_n - y_m)| \\
&\leq |x_n - x_m| + |y_n - y_m| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

Consequently, $(x_n + y_n)$ is Cauchy.

(b) We are given that (x_n) and (y_n) are Cauchy sequences. Cauchy sequences are bounded. So, $|x_n| \leq M_1$ and $|y_n| \leq M_2$ for all $n \in \mathbf{N}$.

(1) Given an $\epsilon > 0$, there exists N_1, such that for all $m, n \geq N_1$,

$$|x_n - x_m| < \frac{\epsilon}{2M_2}$$

(2) Given an $\epsilon > 0$, there exists N_2, such that for all $m, n \geq N_2$,

$$|y_n - y_m| < \frac{\epsilon}{2M_1}$$

Define $N = \max\{N_1, N_2\}$. For $m, n \geq N$, we have:

$$\begin{aligned}
|x_n y_n - x_m y_m| &= |x_n y_n - x_m y_n + x_m y_n - x_m y_m| \\
&\leq |x_n y_n - x_m y_n| + |x_m y_n - x_m y_m| \\
&\leq |y_n|\,|x_n - x_m| + |x_m|\,|y_n - y_m| \\
&< M_2 \cdot \frac{\epsilon}{2M_2} + M_1 \cdot \frac{\epsilon}{2M_1} = \epsilon
\end{aligned}$$

This closes the proof. □

Problem 4.40. Let (a_n) and (b_n) be Cauchy sequences. Decide whether each of the following sequences is a Cauchy sequence, justifying each conclusion.

(a) $c_n = |a_n - b_n|$

(b) $c_n = (-1)^n a_n$

(c) $c_n = [[a_n]]$, where $[[x]]$ refers to the greatest integer less than or equal to x.

Proof. (a) The following inequalities hold for any reals $a, b \in \mathbf{R}$:

$$|a + b| \leq |a| + |b|$$

Replacing b by $-b$, we have:

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

Writing $a = a - b + b$, we have:

$$\begin{aligned}
|a| &= |a - b + b| \\
&\leq |a - b| + |b| \\
\iff |a| - |b| &\leq |a - b|
\end{aligned}$$

Writing $b = b - a + a$, we have:

$$\begin{aligned}
|b| &= |b - a + a| \\
&\leq |a - b| + |a| \\
\iff -(|a| - |b|) &\leq |a - b|
\end{aligned}$$

Consequently,

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

Now, the distance between any two terms of the sequence (c_n) is given by,

$$
\begin{aligned}
|c_n - c_m| &= ||a_n - b_n| - |a_m - b_m|| \\
&\leq |a_n - b_n - (a_m - b_m)| = |(a_n - a_m) - (b_n - b_m)| \\
&\leq |a_n - a_m| + |a_m - b_m|
\end{aligned}
$$

We desire each of these distances to be smaller than $\epsilon/2$.

Since (a_n) and (b_n) are Cauchy sequences, given any $\epsilon > 0$, there exists N_1, N_2 such that for all $m, n \geq N_1$, $|a_n - a_m| < \epsilon/2$ and for all $m, n \geq N_2$, $|b_n - b_m| < \epsilon/2$. Altogether, define $N = \max\{N_1, N_2\}$, if $m, n \geq N$, the terms in the above inequality are each smaller than $\epsilon/2$, which is what we required.

(b) This statement is false. For example, consider the constant sequence $a_n := 1$. This is Cauchy, but $(-1)^n a_n$ is not.

(c) This statement is false as well. Consider the sequence

$$\frac{(-1)^n}{n}$$

This sequence is Cauchy (convergent). But, $c_n = [[a_n]]$ given by

$$\left[\left[\frac{(-1)^n}{n}\right]\right] = \begin{cases} -1 & \text{if n is odd} \\ 0 & \text{if n is even} \end{cases}$$

is not a Cauchy(convergent) sequence. □

Problem 4.41. Consider the following (invented) definition. A sequence (s_n) is *pseudo-Cauchy* if, for all $\epsilon > 0$, there exists N such that if $n \geq N$, then $|s_{n+1} - s_n| < \epsilon$.

Decide which one of the following two propositions is actually true. Supply a proof for the valid statement and a counterexample for the other.

(a) Pseudo-Cauchy sequences are bounded.

(b) If (x_n) and (y_n) are pseudo-Cauchy, then $(x_n + y_n)$ is pseudo-Cauchy as well.

Proof. (a) Consider the sequence (s_n) defined by terms equal to the partial sums of the harmonic series $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + ...$,

$$s_m = \sum_{n=0}^{m} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + ... + \frac{1}{m}$$

We have,

$$|s_{m+1} - s_m| = \frac{1}{m+1}$$

Given any $\epsilon > 0$, there exists an $N \in \mathbf{N}$, such that for $m \geq N$,

$$|s_{m+1} - s_m| < \epsilon$$

So, (s_n) is pseudo-cauchy. But, we know that, the harmonic series is divergent, and hence not Cauchy. The harmonic series is also unbounded.

(b) This is true. If (x_n) and (y_n) are pseudo-Cauchy, then

(1) Given $\epsilon > 0$, there exists N_1 such that for all $n \geq N_1$,

$$|x_{n+1} - x_n| < \frac{\epsilon}{2}$$

(2) Given $\epsilon > 0$, there exists N_2 such that for all $n \geq N_2$,

$$|y_{n+1} - y_n| < \frac{\epsilon}{2}$$

Define $N = \max\{N_1, N_2\}$. Then for all $n \geq N$, we have:

$$\begin{aligned}
|(x_{n+1} + y_{n+1}) - (x_n + y_n)| &= |(x_{n+1} - x_n) + (y_{n+1} - y_n)| \\
&\leq |(x_{n+1} - x_n)| + |(y_{n+1} - y_n)| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

□

Problem 4.42. Let's call a sequence (a_n) quasi-increasing if for all $\epsilon > 0$ there exists an N such that whenever $n > m \geq N$ it follows that $a_n > a_m - \epsilon$.

(a) Give an example of a sequence that is quasi-increasing but not monotone or eventually monotone.

(b) Give an example of a quasi-increasing sequence that is divergent and not monotone or eventually monotone.

(c) Is there an analogue of the Monotone Convergence Theorem for quasi-increasing sequences? Give an example of a bounded, quasi-increasing sequence that doesn't converge, or prove that no such sequence exists.

Proof. □

4.7 Introduction to Infinite Series.

We will now give a brief introduction to the infinite series of real numbers. This is topic that will be discussed in more detail in a further chapter, but because of its importance, we will establish a few results here. These results will be seen to be immediate consequences of theorems we have met in this chapter.

Given an infinite $\sum_{k=1}^{\infty} a_k$, it is important to keep a clear distinction between

(i) the sequence of *terms*: $(a_1, a_2, a_3, ...)$

(ii) the sequence of the *partial sums* : $(s_1, s_2, s_3, ...)$, where $s_n = a_1 + a_2 + ... + a_n$.

The convergence of the series $\sum_{k=1}^{\infty}$ is defined in terms of the sequence (s_n). Specifically, the statement

$$\sum_{k=1}^{\infty} a_k = A$$

means that

$$\lim s_n = A$$

It is for this reason that we can immediately translate many of our results from the study of sequences into statements about the behavior of infinite series.

Theorem: 4.7.1: Algebraic Limit Theorem for series.

If $\sum_{k=1}^{\infty} a_k = A$ and $\sum_{k=1}^{\infty} b_k = B$, then

(i) $\sum_{k=1}^{\infty} ca_k = cA$ for all $c \in \mathbf{R}$ and
(ii) $\sum_{k=1}^{\infty}(a_k + b_k) = A + B$

Proof. In order to show that $\sum_{k=1}^{\infty} ca_k = cA$, we must argue that the sequence of partial sums

$$t_m = ca_1 + ca_2 + ca_3 + ... + ca_m$$

converges to cA. But, we are given that $\sum_{k=1}^{\infty} a_k$ converges to A, meaning that the partial sums

$$s_m = a_1 + a_2 + ... + a_m$$

converge to A. Because, $t_m = cs_m$, applying the Algebraic Limit Theorem for sequences yields $(t_m) \to cA$ as desired.

The proof of part (ii) is analogous. □

One way to summarize Theorem (4.7) (i) is to say that infinite addition still satisfies the distributive property. Part (ii) verifies that series can be added in the usual way. Missing from this theorem is any statement about the *product* of two infinite series. At the heart of this question is the issue of commutativity, which requires a more delicate analysis and so is postponed until later.

Theorem: 4.7.2: Cauchy criterion for Series.

The series $\sum_{k=1}^{\infty} a_k$ converges if and only if, given $\epsilon > 0$, there exists an $N \in \mathbf{N}$ such that whenever $n > m \geq N$ it follows that

$$|a_{m+1} + a_{m+2} + \ldots + a_n| < \epsilon$$

Proof. Observe that

$$|s_n - s_m| = |a_{m+1} + a_{m+2} + \ldots + a_n|$$

and apply the Cauchy criterion for sequences. □

The Cauchy criterion leads to economical proofs of several basic facts about series.

Theorem: 4.7.3: Implication of the convergence of an infinite series.

If the series $\sum_{k=1}^{\infty} a_k$ converges, then $(a_k) \to 0$.

Proof. Consider the special case $n = m + 1$ in the Cauchy criterion for Series. The series is Cauchy(convergent) if and only if, given any $\epsilon > 0$, there exists $N \in \mathbf{N}$, such that for all $m, n \geq N$,

$$|s_n - s_m| < \epsilon$$

Substituting $n = m + 1$, we have:

$$\begin{aligned} |s_{m+1} - s_m| &< \epsilon \\ |(a_1 + a_2 + \ldots + a_m + a_{m+1}) - (a_1 + a_2 + \ldots + a_m)| &< \epsilon \\ |a_{m+1}| &< \epsilon \end{aligned}$$

for all $m \geq N$. Thus, the infinite series approaches the limiting value 0. $(a_k) \to 0$. □

Every statement of this result should be accompanied with a reminder to look at the harmonic series to erase any misconception that the converse of the statement is true. Knowing that $(a_k) \to 0$ does not imply that the series converges.

Theorem: 4.7.4: Comparision Test.

Assume that (a_k) and (b_k) are sequences satisfying $0 \leq a_k \leq b_k$ for all $k \in \mathbf{N}$.
(i) If $\sum_{k=1}^{\infty} b_k$ converges, then $\sum_{k=1}^{\infty} a_k$ converges.
(ii) If $\sum_{k=1}^{\infty} a_k$ diverges, then $\sum_{k=1}^{\infty} b_k$ diverges.

Proof. Both statements follow immediately from the Cauchy criterion for Series and the observation that

$$|a_{m+1} + a_{m+2} + ... + a_n| \leq |b_{m+1} + b_{m+2} + ... + b_n|$$

Cauchy criterion for the convergence of an infinite series implies that, given any $\epsilon > 0$, there exists an $N \in \mathbf{N}$, such that, for all $n > m \geq N$, we have:

$$|b_{m+1} + b_{m+2} + ... + b_n| < \epsilon$$

Thus,

$$|a_{m+1} + a_{m+2} + ... + a_n| < \epsilon$$

for all $n > m \geq N$.

On the other hand, the negation of the definition of the Cauchy sequence is: there exists $\epsilon_0 > 0$, such that for all $H \in \mathbf{N}$, there is atleast one $n > H$ and there is atleast one $m > H$, such that

$$|a_{m+1} + a_{m+2} + ... + a_n| > \epsilon_0$$

Since $b_k \geq a_k$ for all $k \in \mathbf{N}$, we have:

$$|b_{m+1} + b_{m+2} + ... + b_n| \geq |a_{m+1} + a_{m+2} + ... + a_n| > \epsilon_0$$

□

This is a good point to remind ourselves again that the statements about the convergence of sequences and series are immune to changes in some finite number of initial terms. In the Comparision test, the requirement $0 \leq a_k \leq b_k$ does not really need to hold for all $k \in \mathbf{N}$, but just needs to be *eventually* true. A weaker, but sufficient, hypotheses would be to assume that there exists some point $M \in \mathbf{N}$, such that the inequality $a_k \leq b_k$ is true for all $k \geq M$.

The Comparision test is used to deduce the convergence or divergence of one series based on the behavior of another. Thus, for the test to be of any great use, we need a catalog of series we can us as measuring sticks.

Recall the cauchy condensation test for infinite series. Suppose (b_n) is decreasing and satisfies $b_n \geq 0$ for all $n \in \mathbf{N}$. Then, the infinite series $\sum_{n=1}^{\infty} b_n$ converges if and only if the series

$$\sum_{n=0}^{\infty} 2^n b_{2^n} = b_1 + 2b_2 + 4b_4 + 8b_8 + 16b_{16} + ...$$

converges.

An important class of infinite series is $\sum_{n=1}^{\infty} \frac{1}{n^p}$. We can use the Cauchy condensation test to prove that this series is convergent for $p > 1$. The first few terms of the sequence of partial sums (s_n) are

Proof.

$$\begin{aligned}
s_1 &= \frac{1}{1} \\
s_2 &= \frac{1}{1} + \frac{1}{2^p} \\
s_3 &= \frac{1}{1} + \frac{1}{2^p} + \frac{1}{3^p} \\
s_4 &= \frac{1}{1} + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} \\
s_{2^{k+1}-1} &= \frac{1}{1} + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p} + \ldots + \frac{1}{(2^k)^p} + \frac{1}{(2^k+1)^p} + \ldots + \frac{1}{(2^{k+1}-1)^p}
\end{aligned}$$

If $p > 1$, (s_n) is a monotonic decreasing sequence. Therefore, we can write:

$$\begin{aligned}
s_{2^{k+1}-1} &= \frac{1}{1} + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p} + \ldots + \frac{1}{(2^k)^p} + \frac{1}{(2^k+1)^p} + \ldots + \frac{1}{(2^{k+1}-1)^p} \\
&\leq \frac{1}{1} + \left(\frac{1}{2^p} + \frac{1}{2^p}\right) + \left(\frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p}\right) + \ldots \left(\frac{1}{(2^k)^p} + \ldots + \frac{1}{(2^k)^p}\right) \\
&= \frac{1}{1} + \frac{2}{2^p} + \frac{4}{4^p} + \ldots + \frac{2^k}{(2^k)^p} = b_1 + 2b_2 + 4b_4 + 8b_8 + 16b_{16} + \ldots + 2^k b_{2^k} \\
&= \frac{1}{1} + \frac{2^1}{(2^1)^p} + \frac{2^2}{(2^2)^p} + \ldots + \frac{2^k}{(2^k)^p} \\
&= \frac{1}{1} + \left(\frac{1}{2^{p-1}}\right) + \left(\frac{1}{2^{p-1}}\right)^2 + \ldots + \left(\frac{1}{2^{p-1}}\right)^k \\
&= \frac{1\left[1 - \left(\frac{1}{2^{p-1}}\right)^{k+1}\right]}{1 - \frac{1}{2^{p-1}}} \\
&< \frac{1}{1 - \frac{1}{2^{p-1}}}
\end{aligned}$$

Therefore, the sequence $\sum_{n=0}^{\infty} 2^n b_{2^n}$ is convergent, and by Cauchy condensation test hence the infinite series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is convergent for $p > 1$. □

Example 4.7.1. Geometric Series. A series is called *geometric* if it is of the form

$$\sum_{k=0}^{\infty} ar^k = a + ar + ar^2 + ar^3 + ...$$

If $r = 1$ and $a \neq 0$, the series diverges. For $r \neq 1$, the algebraic identity

$$(1-r)(1 + r + r^2 + r^3 ... + r^{m-1}) = 1 - r^m$$

enable us to write the partial sum

$$s_m = a + ar + ar^2 + ar^3 + ... + ar^{m-1} = \frac{a(1-r^m)}{1-r}$$

Now the Algebraic Limit Theorem for sequences and the fact that the sequence $x_n := b^n$, where $|b| < 1$, as $n \to \infty$, in the limit, approaches 0, we have:

$$\begin{aligned}
\lim_{m\to\infty} s_m &= \lim_{m\to\infty} \frac{a(1-r^m)}{1-r} \\
&= \frac{\lim_{m\to\infty} a(1-r^m)}{\lim_{m\to\infty}(1-r)} \\
&= \frac{a - a\lim_{m\to\infty}(r^m)}{1-r} \\
&= \frac{a}{1-r}
\end{aligned}$$

Although the Comparison test requires that the terms of the series be positive, it is often used in conjunction with the next theorem to handle series that contains some negative terms.

***Theorem:* 4.7.5: Absolute Convergence Test.**

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

Proof. This proof makes use of both the necessity (the "if" direction) and the sufficiency (the "only if" direction) of the Cauchy Criterion for Series. Because, $\sum_{n=1}^{\infty} |a_n|$ converges, we know that, given an $\epsilon > 0$, there exists an $N \in \mathbf{N}$ such that

$$|a_{m+1}| + |a_{m+2}| + ... + |a_n| < \epsilon$$

for all $n > m \geq N$. By the triangle inequality,

$$|a_{m+1} + a_{m+2} + ... + a_n| \leq |a_{m+1}| + |a_{m+2}| + ... + |a_n| < \epsilon$$

so the sufficiency of the Cauchy criteria guarantees that $\sum_{n=1}^{\infty} a_n$ also converges. $\square$

The converse of this theorem is false. In the opening discussion of this chapter, we considered the *alternating harmonic series*

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

Taking the absolute values of the terms, gives us the harmonic series $\sum_{n=1}^{\infty}(1/n)$, which we have seen diverges. However, it is not too difficult to prove that with the alternating negative signs the series indeed converges. This is a special case of the Alternating Series Test.

***Theorem:* 4.7.6: Alternating Series Test.**

Let (a_n) be a sequence satisfying
(i) $a_1 \geq a_2 \geq a_3 \ldots \geq a_n \geq a_{n+1} \geq \ldots$ and
(ii) $(a_n) \to 0$
Then, the alternating series

Chapter 5

Basic Topology of $\mathbf{R}$

5.1 Cantor's Theorem.

Cantor's work into the theory of infinite sets extends far beyond far beyond the conclusions of the theorem 3.7.1. Although initially resisted, his creative and relentless assault in this area eventually produced a revolution in theory and paradigm shift in the way mathematicians came to understand the infinite.

5.1.1 Cantor's Diagonalization method.

Cantor published his discovery that $\mathbf{R}$ is uncountable in 1874. Although it has some modern polish on it, the argument presented

5.2 Discussion: The Cantor Set.

What follows is a fascinating mathematical construction, due to Greg Cantor, which is extremely useful for extending the horizons of our intuition about the nature of subsets of the real line. Cantor's name has already appeared in the first chapter in our discussion of uncountable sets. Indeed Cantor's proof, that $\mathbf{R}$ is uncountable occupied another spot on the short list of most significant contributions towards understanding the mathematical infinite. In the words of the mathematician David Hilbert, "No one shall expel us from the paradise that Cantor has created for us."

Let C_0 be the closed interval $[0,1]$, and define C_1 to be the set that results when the open middle third is removed, that is,

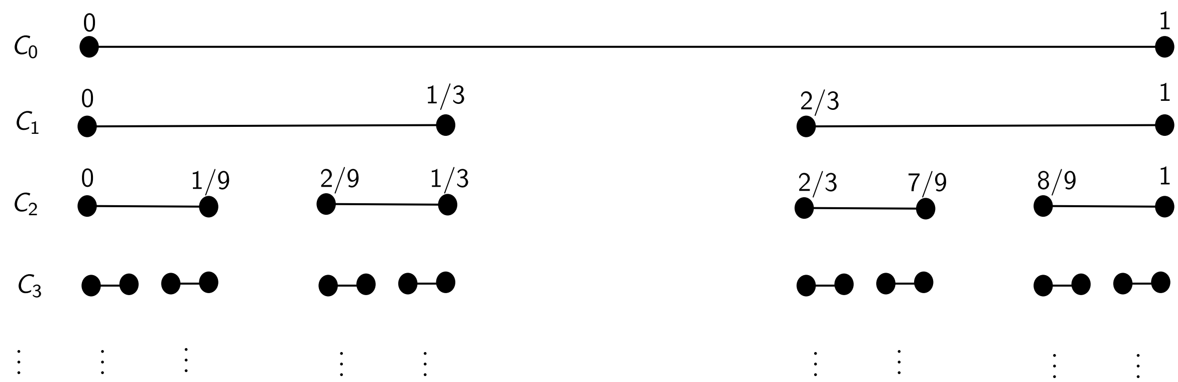
$$C_1 = C_0 - \left(\frac{1}{3}, \frac{2}{3}\right) = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

Now, construct C_2 in a similar way by removing the open middle third of each of the two components of C_1 :

$$C_2 = \left(\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{3}{9}\right]\right) \cup \left(\left[\frac{6}{9}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]\right)$$

If we continue this process inductively, then for each $n = 0, 1, 2, ...$, we get a set C_n consisting of 2^n closed intervals each having length $\left(\frac{1}{3}\right)^n$. Finally, we define the *Cantor set* C to be the intersection

$$C = \bigcap_{n=0}^{\infty} C_n$$



Defining the Cantor Set;

It may be useful to understand C as the remainder of the interval $[0, 1]$ after the iterative process of removing the middle one-thirds is taken to infinity:

$$C = [0, 1] \quad \left[\left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right) \cup ...\right]$$

There is some initial doubt whether anything remains at all, but notice that because we are always removing open middle thirds, then for every $n \in \mathbf{N}$, $0 \in C_n$ and hence $0 \in C$. The same argument shows that $1 \in C$. In fact, if y endpoint of some closed interval of some particular set C_n, then it is also an end-point of one of the intervals of C_{n+1}. Because, at each stage, endpoints are never removed, it follows that $y \in C_n$ for all n. Thus, C atleast contains the endpoints of all of the intervals that make up each of the sets C_n.

Is there anything else? Is C countable? Does C contain any intervals? Any irrational numbers? These are difficult questions at the moment. All of the endpoints mentioned earlier are rational numbers (they have the form $m/3^n$), which means that if it is true that C consists of only these endpoints, then C would be a subset of $\mathbf{Q}$ and hence countable. We shall see about this. There is some strong evidence that not much is left in C if we consider the total length of the intervals removed. To form C_1, an open interval of length $1/3$ was taken out. In the second step, we removed two intervals of length $1/9$, and to construct C_n, we removed 2^{n-1} middle thirds of length $1/3^n$. There is some logic, then, to defining the length of C to be 1 minus the total

$$\frac{1}{3} + 2 \cdot \frac{1}{9} + 4 \cdot \frac{1}{27} + ... + 2^{n-1} \cdot \left(\frac{1}{3}\right)^n + ... = \frac{\frac{1}{3}}{1 - \frac{2}{3}} = 1$$

The Cantor set has *zero length.*

To this point, the information we have collected suggests a mental picture of C as a relatively small, thin set. For these reasons, the set C is often referred to as *Cantor dust.* But there are some very strong counterarguments that imply a very different picture. First, C is actually uncountable, with cardinality equal to the cardinality of $\mathbf{R}$. One slightly intuitive but convincing way to see this is to create a $1-1$ correspondence between C and sequences of the form $(a_n)_{n=1}^{\infty}$ where $a_n = 0$ or $a_n = 1$. For each $c \in C$, set $a_n = 0$ or 1. For each $c \in C$, set $a_1 = 0$ if c falls in the left-hand component of C_1 and set $a_1 = 1$ if c falls in the right hand component. Having established where in C_1 the point c is located, there are now two possible components of C_2 that might contain c. This time, we set $a_2 = 0$ or 1 depending on whether c falls in the left half or right half of these two components of C_2. Continuing in this way, we come to see that every element $c \in C$ yields a sequence $(a_1, a_2, a_3, ...)$ of zeroes and ones that acts as a set of directions for how to locate c within C. Likewise, every such sequence corresponds to a point in the Cantor set.

Bibliography

[1] Terrence Tao. *Analysis I*, 3rd Edition. Springer, 2007.

[2] Stephen Abbott. *Understanding Analysis*, 2nd Edition. Springer Undergraduate texts in Mathematics, 2015.