Real Analysis

Long Le

Winter 2018

Contents

Preface

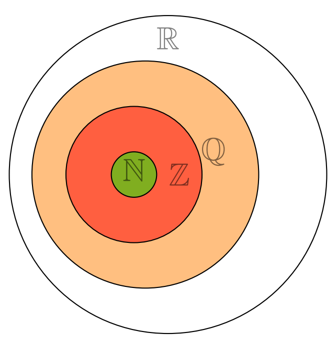
This is my note during the lonely winter break of 2018 in my dorm North Apartment, Amherst, UMass. During this season, I read *Analysis* by *Terence Tao* to prepare for graduate classes Math 697AM Math Application and Modeling, and Math 697U Stochastic Processes that I will take in Spring 2019. The book was based on Math 131AH taught at UCLA by Tao himself (course website). SPOILER ALERT: Tao's *Analysis* is a superb book.

In the first four chapters, we will build our basic number systems including the natural numbers, integers, rationals, and real numbers respectively. As we will see, everything about the number systems that we cherish are built from the ground up – starting from 5 axioms of Peano to construct the natural numbers.

Induction is a powerful tool to prove many properties about the natural numbers. This is understandable because the natural numbers itself are defined recursively (if $n \in \mathbb{N}$, then $n{+}{+} \in \mathbb{N}$). Fundamental operations are also defined recursively. Additions are recursive incrementation. Multiplication is recursive addition, and exponentiation is recursive multiplication. "Fancier" operations like subtraction, division and negation are built later when we have more generalized number system like integers and the rational numbers.

Properties about operations like associativity and commutativity are also proven. The integers are built from the Cartesian product of the natural numbers, using equivalence classes. Equivalence classes are also the way that we will build the rationals and real numbers later on.

For each new number system, we will define the equivalence relation that cut the set into equivalence classes. We then define the usual operations and prove properties about them. For appropriate number systems, these properties are understood systematically through axioms of rings or fields. We will also briefly show that equivalence class is indeed equivalent in the light of operations. *Tao* calls this *substitution axiom*: if $x = x'$, then $x \cdot y = x' \cdot y$ where $\cdot$ is some binary operation like addition. We also show the embedding of the previous number system inside this current system through *group homomorphism*. Lastly, we deal with problems relating to ordering and trichotomy. Properties about partial order and total order will be proven. Trichotomy is related to the connex property of total order. For a simple case of integers, trichotomy states that a number $a \in \mathbb{Z}$ is either positive, negative or 0.



Chapter 1

Natural Numbers

1.1 Peano's Axioms

This is an axiomatic treatment of the natural numbers. We will construct the natural numbers that we know and love from scratch, based on a minimal and powerful set of axioms. First to remind the readers, the natural numbers that we know are

$$0,\ 1,\ 2,\ ...$$

To define this type of numbers, we will need two things: an initial starting point (0) and an increment operator (++). By incrementing, we mean increasing a number by 1 so $1 = 0++, 2 = 1++, ...$ For now, we can think of ++ as an arbitrary symbol that when applied to a natural number would produce another natural number.

$1, 2, 3, ...$ are of course just convenient notation. 1 is the "increment of 0". 2 is the the "increment of 1" or the "increment of "the increment of 0" " and so on. Everything we discuss here still hold if we use a different notations like Roman numerals i.e. $I, II, III, ...$

Now we will define two foundational axioms

Axiom 1 (Initial Value)**.** *0 is a natural number.*

Axiom 2 (Recursive increment operator)**.** *If n is a natural number, then $n++$ is also a natural number.*

However, these two axioms are not sufficient. We might have a weird natural number system that "loops back" like $0, 1, \mathbf{2}, 3, \mathbf{2}, 4, ...$ In fact, in computer system, overflows will cause the 2's complement representation to loop back like that. In mathematics, we do not allow this behavior. So we will introduce two additional axioms to avoid repetition in our number system.

Axiom 3 (Bounded at one end)**.** *0 has no predecessor i.e. there does not exist natural number n such that $0 = n++$.*

Axiom 4 (Unique Predecessor)**.** *If n and m are two different natural numbers, then $n++$ and $m++$ are different. Equivalently, if $n++$ and $m++$ are the same, then n and m must be the same.*

In other words, axiom 4 does **not** allow this type of situation.

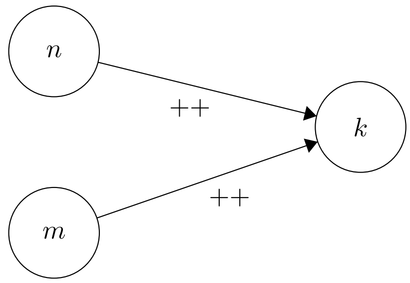


Figure 1.1: **Not** unique predecessors

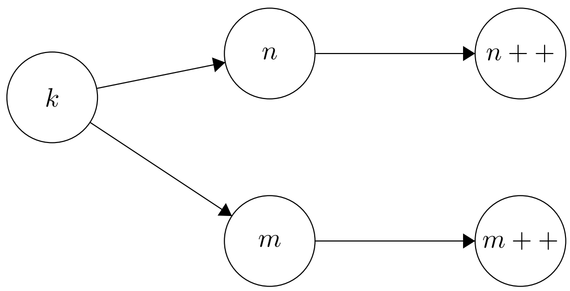


Figure 1.2: **Unique** predecessor but **not** unique successor

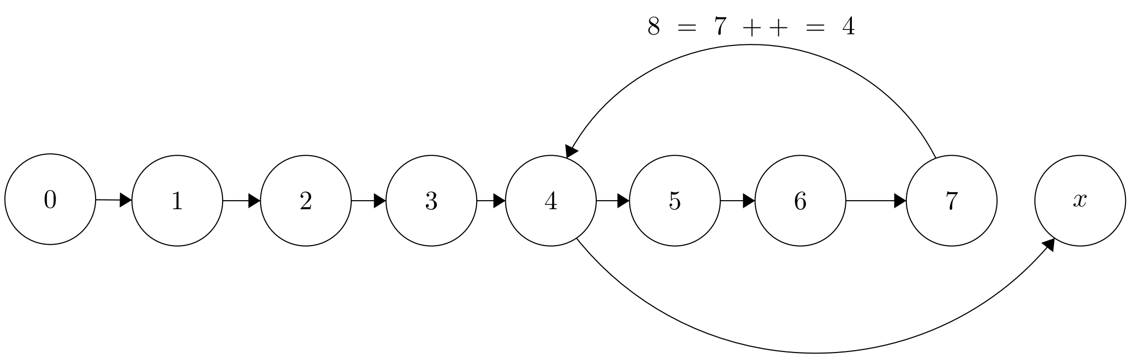
Axiom 3 basically re-affirms axiom 1 that 0 is the initial starting point. In the natural number system, there is no number that comes before 0. From this axiom, we know that 0 is unique. Axiom 4 basically provides a mechanism that we can use to tell if two given natural numbers are equal to each other.

Example 1: Uniqueness of 0
0 is not equal to 2.

Proof. $2 = 1++$, but 0 does not have a predecessor. Thus, 2 and 0 must be different. This argument can be easily extended to say that 0 is different from $1, 3, 4, ...$ and so on. $\square$

We will soon see that axiom 3 in conjunction with axiom 4 will avoid any repetition in our natural number system sequence.

Example 2: Uniqueness of any natural number
8 is not equal to 4.



Proof. For the sake of contradiction, say we have a sequence with two 4 like

0, 1, 2, 3, 4, 5, 6, 7, 4, x, y, z, k, h , ...

By the unique predecessor axiom 4, we know that $7 = 3$ because $4 = 7{+}{+} = 3{+}{+}$. Similarly, we know that $6 = 2$, $5 = 1$ and $4 = 0$. The last statement violates example 1, so we know that 4 can not be repeated. This essentially says proves that $8 := 7{+}{+}$ is different from 4. □

Figure 1.2 is a little bit concerning since we know that each natural number should also have a unique successor. For this reason, the ++ increment operator can sometimes be defined as a successor **function** $S(n) = n{+}{+}$. And of course, a function should map a unique x to a unique y. Axiom 4 asserts that $S(n)$ is injective.

This is what the natural numbers should look like

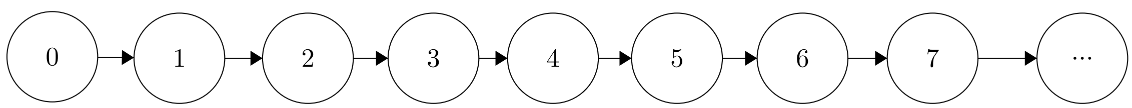


Figure 1.3: Natural numbers

However, our 4 axioms so far are not sufficient to define the natural numbers that we know. We can have a weird natural number system as in the picture that still obeys all the previous axioms.

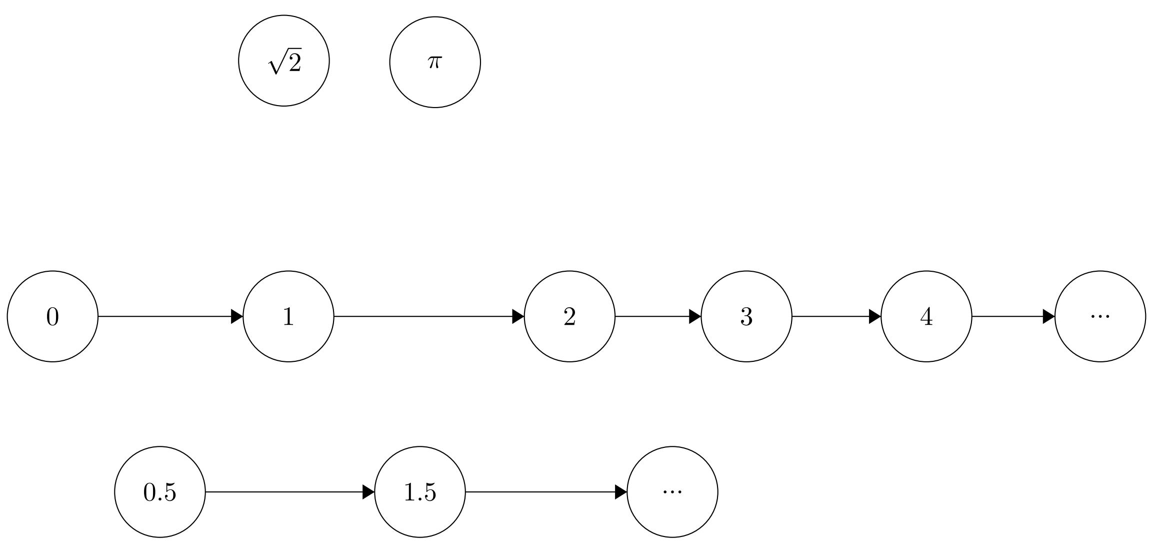


Figure 1.4: Pathological natural numbers

The problem with this system is that not all "natural numbers" other than 0 i.e. $\sqrt{2}, \pi, 0.5$ are preceded by a natural number. In other words, we should require the successor function $S : \mathbb{N} \to \mathbb{N} \setminus \{0\}$ to be surjective in addition to being injective in axiom 4. If S is a bijection, then there exists the predecessor function S^{-1} that takes $S^{-1}(n) = n{-}{-}$. This would require $\sqrt{2}, \pi, 0.5$ to have predecessors. However, even this would not prevent a situation like

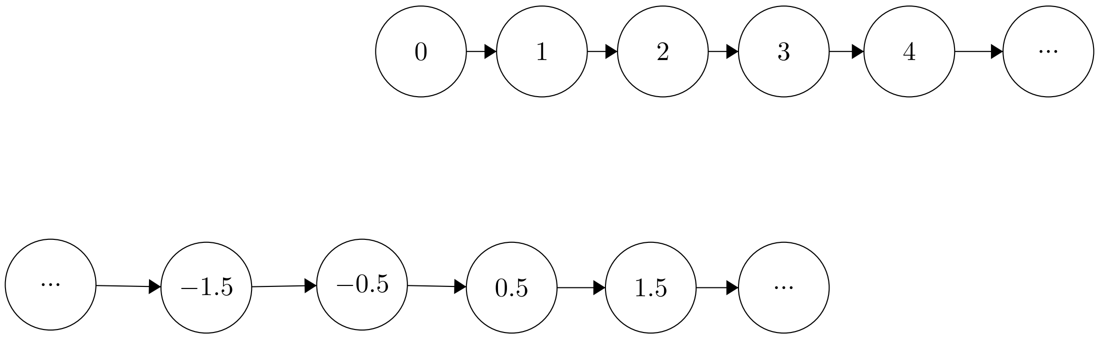


Figure 1.5: "Natural numbers" extended in both directions

Note that the second chain extends infinitely towards both left and right directions and is totally disconnected from the first chain, which truly is the natural numbers. In fact, we need to assume a very strong axiom that will serve as a powerful tool for many proofs later on.

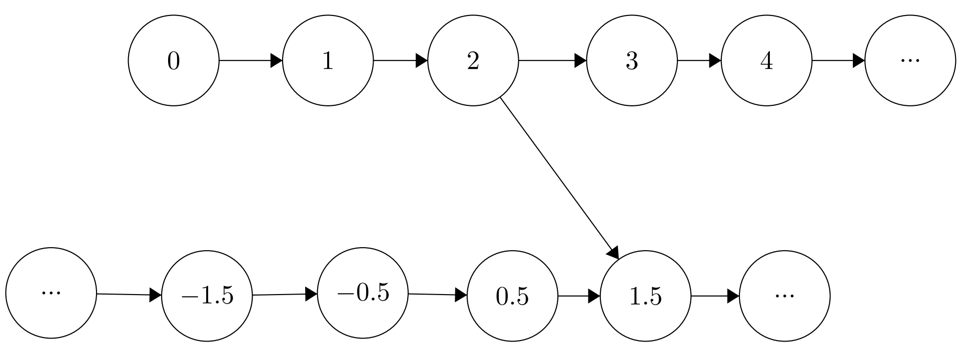
Axiom 5 (Induction)**.** *Define a predicate as a Boolean-valued function $f : X \to \{true, false\}$ for some set X. A unary predicate that takes in only one single value (instead of a vector for example) is called property P. The axiom states that if*

- *$P(0)$ is true and*
- *for every natural number n, $P(n)$ is true implies that $P(n++)$ is true*

then P is true for every natural number.

Note that Figure 1.4 can be eliminated by setting $P(n) =$ "n is 0 or n has a predecessor" i.e. for every natural numbers $n \neq 0$ there exists $m \in \mathbb{N}$ such that $n = m++$.

We can prevent Figure 1.5 by setting $p(n) =$ " n is 0 or there is a path (in a graphical sense) from n to 0. Obviously, if there is a path from n to 0, then $n++$ can follows the edge from $n++$ to n and will find a path back to 0. By induction axiom 5, there must be a path from n to 0 for every natural number n. This means that the second chain must be connected to the first chain at some point.



This would violate unique predecessor axiom 4. This is a non-rigorous argument since we have not defined what is a path. To sum up, we have defined 5 Peano's axioms

- Axiom 1 (starting point): $0 \in \mathbb{N}$.
- Axiom 2 (increment): If $n \in \mathbb{N}$, then $n++ = S(n) \in \mathbb{N}$.
- Axiom 3 (bounded): 0 has no predecessor i.e. there does **not** exist $m \in N$ such that $m++ = 0$.
- Axiom 4 (unique predecessor): $S(n)$ is injective. $m = n$ if and only if $m++ = n++$.
- Axiom 5 (Induction) For unary predicate (property) P, if $P(0)$ is true and $P(n) \implies P(n++)$, then $P(n)$ is true for all $n \in N$.

Axiom 4 defines what it means to say two natural numbers are equal. Before proceeding, we will lay down a standard definition of equality that would apply to all number sets including natural numbers, integers, rationals and real numbers.

Definition 1.1.1 (Equality)**.** *Equality relation $P(x, y)$ is a binary predicate between two objects x and y satisfies three properties:*
Reflexive: $x = x$ i.e. $P(x, x)$ is true.
Symmetric: If $x = y$ then $y = x$ i.e. $P(x, y) \implies P(y, x)$.
Transitive: If $x = y$ and $y = z$, then $x = z$ i.e. $(P(x, y) \wedge P(y, z)) \implies P(x, z)$
Finally, for the case of natural numbers, we also assume that if $x = y$ and $x \in \mathbb{N}$ then $y \in \mathbb{N}$.

We can verify that the equality definition of natural numbers in axiom 4 satisfies three properties above.

1.2 Addition

We will define addition recursively. But before we are able to do it, we must prove that recursive definition makes sense. First, we remind the readers of the recursive definition in the case of the Fibonacci Sequence: $a_0, a_1, a_2, ...$

$$a_0 = a_1 := 1$$

$$a_n = a_{n-1} + a_{n-2}$$

We can see that the recursive definition starts with some initial value. Then the element a_n is defined as some some function f of the previous values.

In this section, we are only concerned with one simple type of recursive definition. That is

Definition 1.2.1 (Recursive definition)**.** *The sequence $a_0, a_1, a_2, ...$ are defined recursively as follows*

$$a_0 = c \in \mathbb{N}$$

$$a_{n+1} = f_n(a_n)$$

for f_n is some function $\mathbb{N} \to \mathbb{N}$.

ote that the functions $f_0, f_1, f_2, ...$ can be the same or different. The instance that these functions are different is something like

$$a_{n+1} = f_n(a_n) = (a_n)^n$$

So $f_0(x) = 1, f_1(x) = x, f_2(x) = x^2, ...$ are all different functions.

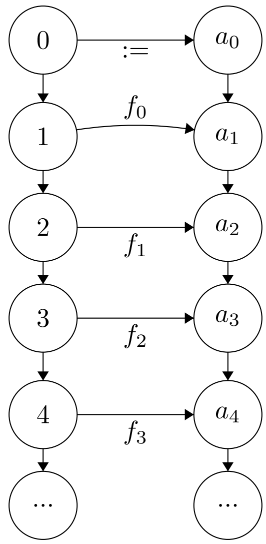


Figure 1.6: Recursive definition

We can see why definition 1.2.1 will not make sense for some weird number system that loops back.

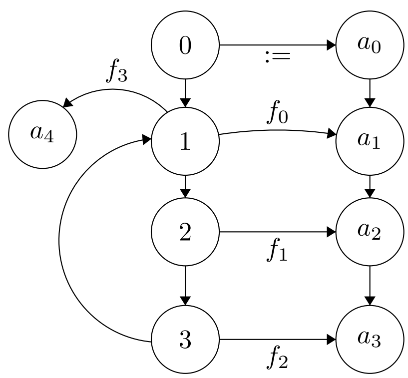


Figure 1.7: Rogue recursive definition

In this situation, we see that a_1 is effectively defined twice: once as $a_1 = f_0(a_0)$ and as $a_1 = f_3(a_3)$. This is because 1 is the successor of both 0 and 3 in this weird system, leading to duplicate definitions.

We will use induction to prove the validity of recursive definition 1.2.1. Induct on n. When $n = 0$, we can see that a_0 is defined exactly once because 0 has no predecessor (axiom 3). Now, if we assume that a_n is defined once, then $a_{n+1} = f_n(a_n)$ is also defined once. This is because the predecessor of $n{+}{+}$ is the unique n by the unique predecessor axiom 4.

We are now ready to define addition.

Definition 1.2.2 (Addition of two natural numbers)**.** *$m+n$ is defined as the value of a_n in the sequence $A = a_0, a_1, a_2, ..., a_{n-1}, a_n$. The sequence A is defined recursively*

$$a_0 = m$$

$$a_n = f_{n-1}(a_{n-1}) = (a_{n-1}) + +$$

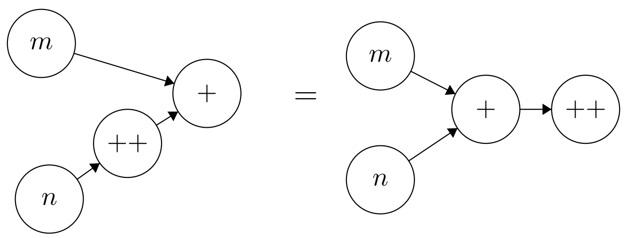
In other words,

$$m + 0 := m$$

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

with $m, n \in \mathbb{N}$

For example, $2 + 3$ is incrementing 2 three times. $2 + 3 = (((2 + +) + +) + +)$.



We will prove some algebraic properties of addition.

Property 1 (Additive group closure)**.** *Sum of two natural numbers are again a natural number.*

Proof. With $m, n \in \mathbb{N}$, we will prove that $m + n \in \mathbb{N}$ by induction axiom 5 on n. For the base case $n = 0$, we know by the definition of addition that $n + 0 = n \in \mathbb{N}$.

Now, assume that the statement is proven for n. We will demonstrate that it also holds for $n + +$. By recursive definition, $m + (n + +) := (m + n) + +$ Because by the inductive hypothesis $(m + n) = k$ is a natural number, $k + +$ is also a natural number by increment axiom 2. □

From definition 1.2.2, however, it is not clear that addition is commutative.

$$n + m = m + n$$

We will now prove a series of remarks by induction leading to the proof that addition is commutative.

Property 2 (Additive group identity)**.** $0 + m = m$

Proof. Note that definition 1.2.2 states that $m + 0 := m$ and we have not proven addition to be commutative. Thus, $0 + m$ are not necessarily the same as $m + 0$. Note that the very definition of the natural number m is defined recursively in axiom 2 that

$$m := \underbrace{(((0 + +) + +) ... + ++)}_{m \text{ times}} \tag{1.1}$$

So we can easily prove this remark by induction on m. For the base case $m = 0$, by definition 1.2.2, $n + 0 = n = 0 = m$. So the claim that $0 + m = m$ is true.

Now suppose that the claim is proven for m, we will prove it extends to $m + +$. By definition 1.2.2,

$$0 + (m + +) := (0 + m) + +$$

By the induction hypothesis,

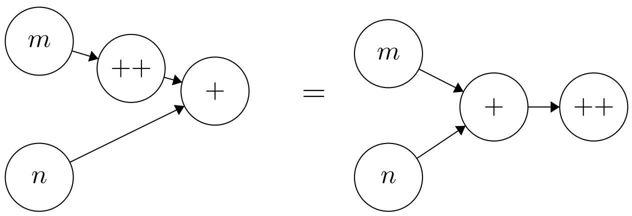
$$(0 + m) + + = (m) + +$$

Therefore,

$$0 + (m++) = (m)++$$

This proves the claim. $\square$

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



Proof. We will again use induction on n. For $n = 0$, by definition 1.2.2, $(m++) + 0 := m++$, and $(m+0)++ = (m)++$ as proven by the group identity property above.

Now, assume the claim is true for n, that is

$$(m++) + n = (m+n)++,$$

we will prove the claim for $n++$. That is we want to demonstrate that

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

By definition 1.2.2,

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

By the induction hypothesis,

$$= ((m+n)++)++.$$

Turning to the right hand side. Now using definition 1.2.2 again,

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

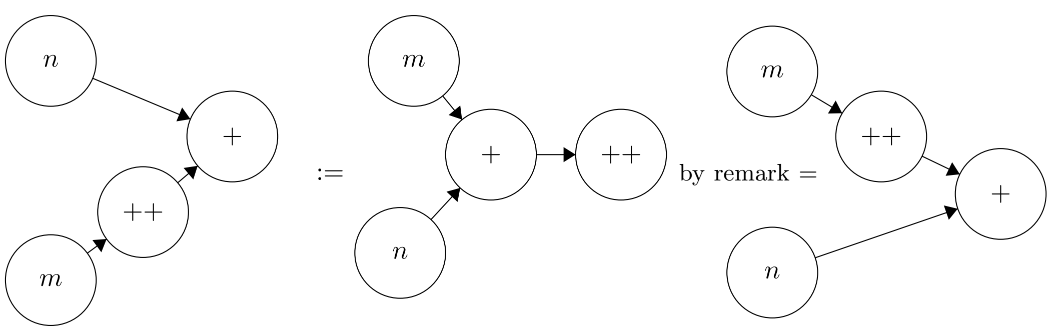
Obviously by the increment axiom 2, if $a = b$, then $a++ = b++$ (from now on, we will refrain from explicitly invoking axiom 2). This completes our claim. $\square$

Property 3 (Addition is commutative)**.** $n + m = m + n$

Proof. We will induct on m. For the base case $m = 0$, we already show in group identity property that $0 + n = n$. By definition 1.2.2, we also know that $n + 0 := n$ so the claim is proven for the base case.

Now, assume the claim is true for m, we extend it to $m++$. That is we want to demonstrate that

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



By definition 1.2.2,

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

By the proven remark, we also have

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

And by the induction hypothesis, we have

$$n + m = m + n$$

Thus, we conclude that

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

and addition is commutative. $\square$

Property 4 (Addition is associative)**.** $a + (b + c) = (a + b) + c$

Proof. We will prove the claim by inducting on c. For $c = 0$, We know that $b + 0 := b$, so it is certainly true that

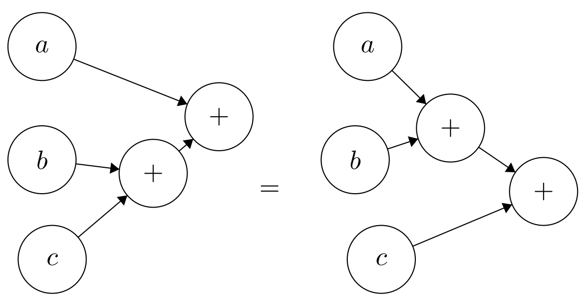
$$a + (b + 0) = a + b$$

By the group identity,

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

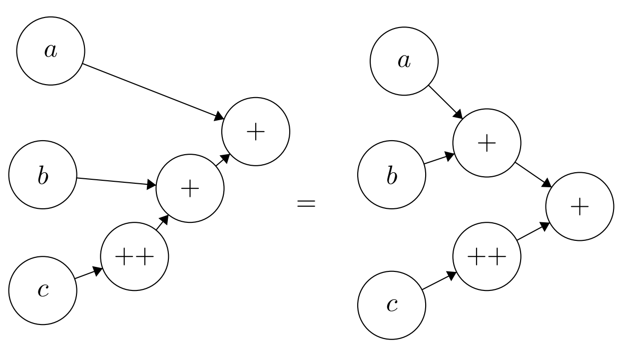
Thus, the claim is proven for base case. Now we assume that the claim is proven for a given c i.e.

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



Now, we need to show that

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



By the definition of addition, the equation is equivalent to

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

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

We know that by the induction hypothesis,

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

Thus, by the unique successor implication of axiom 4

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

This concludes our induction. $\square$

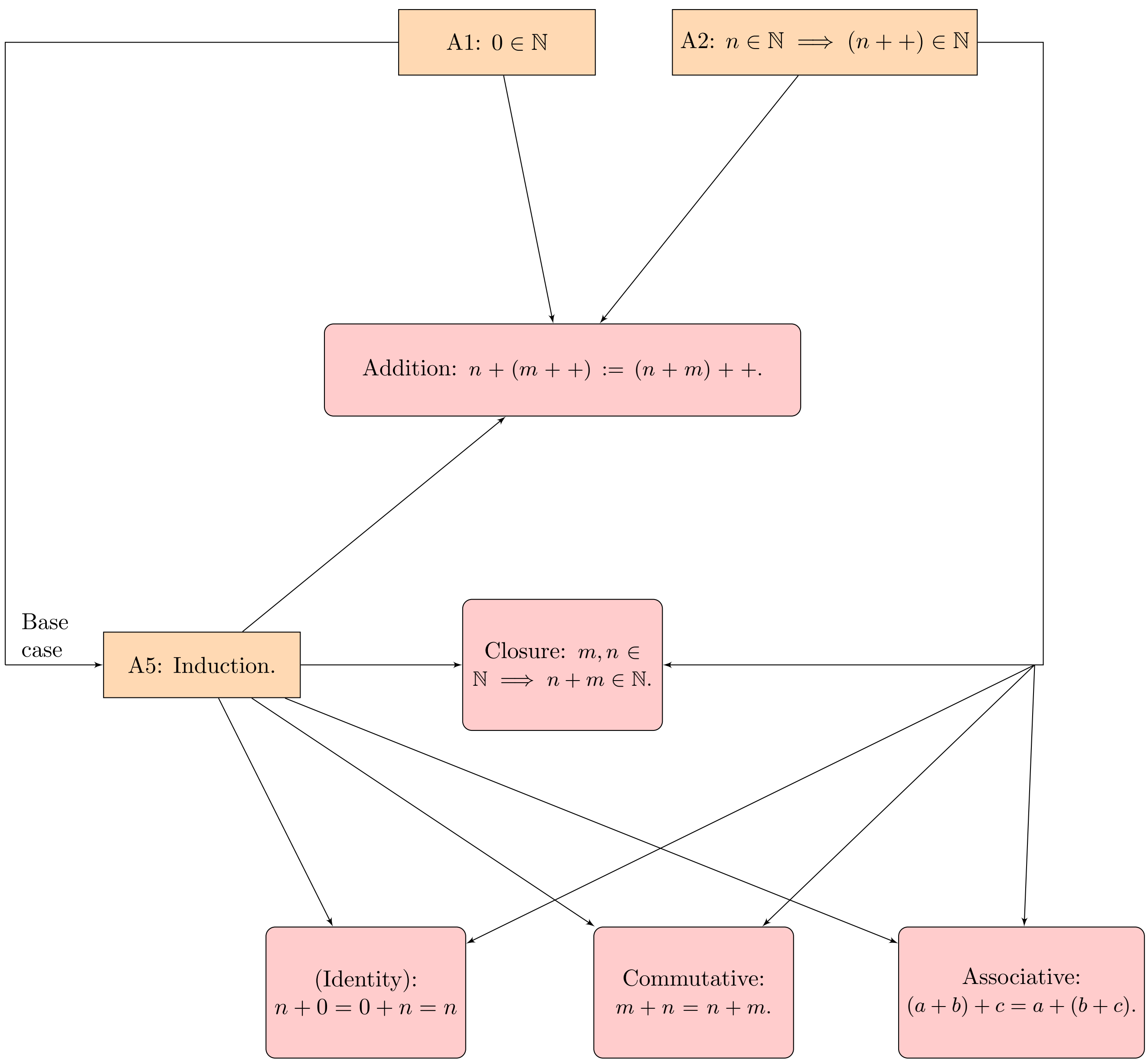


Figure 1.8: Algebraic properties of addition.

1.3 Multiplication

We now move on to multiplication. Similar to addition, multiplication can be defined recursively, but now in terms of addition instead of incrementation

Definition 1.3.1 (Multiplication). *Multiplication has a starting point*

$$n \times 0 := 0$$

and is defined recursively in terms of addition

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

with $m, n \in \mathbb{N}$.

Similar to addition, we will some algebraic properties of multiplication.

Property 5 (Multiplicative closure). *If* $m, n \in \mathbb{N}$*, then* $n \times m \in \mathbb{N}$.

Proof. We will induct on m. For the base case $m = 1$, we have

$$n \times 1 := n \in \mathbb{N}$$

Now, assume the claim is proven for m, we will show that it is also true for $m++$.

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

By induction hypothesis, $(n \times m)$ is a natural number, and we assume n to be natural. And by the additive closure, we know that sum of two natural numbers are a natural number. Thus, $(n \times m) + n$ is natural, and the claim is proven. □

Property 6 (Multiplicative identity). $n \times 1 = 1 \times n = n$

We can easily verify from the definition that $n \times 1 = n \times (0++) := n \times 0 + n = n$, so we only need to show $1 \times n = n$. First, we will show that

$$n + 1 = n++$$

This is because of the definition of addition,

$$n + (0++) := (n + 0)++ := n++$$

Now, induct on n. For the base case $n = 0$, we know from the definition of multiplication,

$$1 \times 0 := 0$$

Now, assume the claim is proven for n, we will show that it is also true for $n++$. That is

$$1 \times (n++) = n++$$

The left hand side is equal to

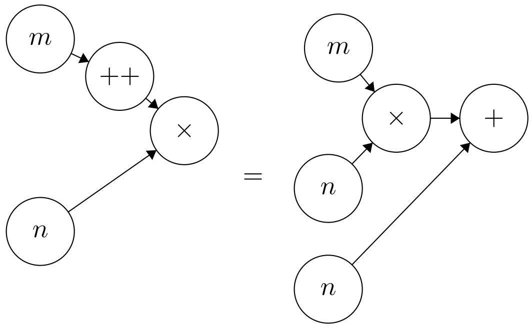
$$1 \times (n++) := (1 \times n) + 1$$

By the induction hypothesis, we know that $(1 \times n) = n$. And we also already show that $n + 1 = n++$. Therefore, the left hand side is equal to the right hand side. Induction is done.

Property 7 (Multiplication is commutative). $n \times m = m \times n$

Proof. The proof for multiplication is very similar to that of addition. First, we want to show that

$$(m++) \times n = m \times n + n$$



We will proceed by induction on n. For the base case $n = 0$, we need to show that

$$(m++) \times 1 = m \times 0 + 0$$

The left-hand side is equal to

$$(m++) \times 1 = (m++) \times (0++) := (m++) \times 0 + (m++)$$

By the additive identity,

$$(m++) \times 0 + (m++) = m++$$

By the multiplicative identity,

$$= m++ = m + 1 = m \times 1 + 1$$

So the base case is proven. Now, we assume the claim holds for n and extends it to $n++$. That is we want to show that

$$(m++) \times (n++) = m \times (n++) + (n++)$$

The left hand is equal to

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

By the induction hypothesis, this is equal to

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

The left-hand side is now equal to the right hand side, and our induction is complete.

We will now show that multiplication is commutative by induction on m. For the base case $m = 1$, we already have the multiplicative identity that $m \times 1 = 1 \times m = m$. We assume the claim is true for m and will extend it to $m++$. We need to show that

$$n \times (m++) = (m++) \times n$$

The left hand side is

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

By the induction hypothesis,

$$n \times m + n = m \times n + n$$

This is equal to the right-hand side as proven above. This concludes our induction. □

As we can see, the multiplicative commutativity is proven similarly to the way we prove additive multiplicative. We might expect to prove multiplicative associativity similarly to the way we prove additive associativity. As it turns out, we need distributivity before associativity.

Property 8 (Distributivity)**.** $(b + c) \times a = b \times a + c \times a$

Proof. We will induct on a. For the base case $a = 0$, we know that

$$(b+c)\times 0 := 0$$

and

$$b\times 0 + c\times 0 := 0+0 = 0$$

So the base case is proven. Now, assume the claim holds for a, we will prove that

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

The left-hand side is equal to

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

By the induction hypothesis, this is equal to

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

Addition is associative so we can move terms around

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

The left-hand side is now equal to the right-hand side, and our induction is complete. □

Property 9 (Multiplication is associative)**.** $a\times(b\times c) = (a\times b)\times c$

Proof. We will induct on c. For the base case $c = 1$, we know by the multiplicative identity that

$$a\times(b\times 1) = a\times b = (a\times b)\times 1$$

Now, we assume the claim holds for c and extends it to $c++$. That is we want to show that

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

The left hand side is equal to

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

Using the distributive law, we have

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

By the induction hypothesis, we have

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

The left-hand side is now equal to the right-hand side, and our induction is complete. □

One might wonder why we need distributivity to prove multiplicative associativity. After all, multiplication is just like addition. Both are defined recursively in terms of more primitive operations: multiplication in terms of addition, and addition in terms of incrementation. The proof for additive associativity did not invoke distributive law, so why would the above proof? It seems that this is because of the way we define these operations. Recall

$$a + (b++) := (a+b)++$$

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

Let the function $A_a(b) := a + b$ and $M_a(b) := a\times b$, $S(n) := n++$ respectively. The recursive definitions are equivalent to,

$$A_a(S(b)) := S(A_a(b))$$

$$M_a(S(b)) := A_a(M_a(b))$$

These look very familiar. As noted, addition is defined recursively in terms of incrementation ($S(A_a(b))$), and multiplication in terms of addition ($A_a(M_a(b))$). However, let us look closer at addition

$$A_a(S(b)) := S(A_a(b))$$

We can see the function composition commutes. $A_a \circ S(b) = S \circ A_a(b)$. This is not true in general: function composition is **not** necessarily commutative. We can see it is certainly not true for the multiplication definition. $M_a \circ S(b) \neq S \circ M_a(b)$ i.e. $a \times (b+1) \neq (a \times b) + +$. This makes addition somehow "special" than multiplication.

There's *exericse 3.5.12* in Tao's Analysis that frames the recursive definition in more rigorous language.

Theorem 1 (Recursion theorem)**.** *Let $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be a function, and c a natural number. Show that there exists a function $a : \mathbb{N} \to \mathbb{N}$ such that*

$$a(0) = c$$

$$a(n++) = f(n.a(n)) \textit{ for all } n \in \mathbb{N}$$

and furthermore this function is unique. Hint: first show inductively by modification of proof of Lemma 3.5.12 that for every natural number $N \in \mathbb{N}$ there exists a unique function $a_N : \{n \in \mathbb{N} : n \leq N\} \to \mathbb{N}$ such that $a_N(0) = c$ and $a_N(n++) = f(n, a(n))$ for all $n < N$. For additional challenge, prove this result using only Peano axioms (without using the ordering of the natural numbers and without appealing to Proposition 2.1.16 i.e. definition 1.2.1). Hint: first show inductively, using only Peano and basic set theory, that for every $N \in \mathbb{N}$, there exists a unique pair A_N, B_N of subsets of $\mathbb{N}$ which obeys the following properties: (a) $A_N \cap B_N = \emptyset$, (b) $A_N \cup B_N = \mathbb{N}$, (c) $0 \in A_N$, (d) $N++ \in B_N$, (e) whenever $n \in B_N$ we have $n++ \in B_N$ (f) whenever $n \in A_N$ and $n \neq N$, we have $n++ \in A_N$. Once one obtains these sets, use A_N as a substitute for $\{n \in \mathbb{N} : n \leq N\}$ in the previous statement.

Proof. Online proof from harmonicuser on math.stackexchange.

One has to be careful while forming infinite sets. Simply because $\alpha(n)$ is defined for each $n \in \mathbb{N}$ you cannot form the set $\{(n, \alpha(n)) : n \in \mathbb{N}\}$.

Existence: Let $\mathcal{C} = \{A \subseteq \mathbb{N} \times \mathbb{N} : (0, c) \in A, (n++, f(n, a(n))) \in A \text{ for all } n \in \mathbb{N}\}$ and order $\mathcal{C}$ by inclusion $\subseteq$. $\mathbb{N} \times \mathbb{N} \in \mathcal{C}$ and therefore the collection $\mathcal{C}$ is not empty. It is easy to see that $C = \bigcap_{A \in \mathcal{C}} A \in \mathcal{C}$ and is the smallest such element in $\mathcal{C}$ with this property.

We claim that C is a graph. ($(x, y) \in C$ and $(x, z) \in C \implies y = z$). If $(0, c), (0, d) \in C$, with $c \neq d$. Then $C - \{(0, d)\} \in \mathcal{C}$ - a contradiction. Similarly if $(n++, f(n, a(n)))$ and $(n++, d) \in C$ with $d \neq f(n, a(n))$ for some $n \in \mathbb{N}$, then $C - \{(n++, d)\} \in \mathcal{C}$ which contradicts the minimality of C. Hence C is a graph and therefore defines a function.

Uniqueness can be proved using induction. □

1.4 Exponent

We will now define exponent recursively in terms of multiplication

Definition 1.4.1 (Exponent)**.** *Exponentiation has a starting point*

$$a^0 := 1$$

and is defined recursively in terms of multiplication

$$a^{b++} := a^b \times a$$

with $a, b \in \mathbb{N}$

nlike two previous operations, exponentiation is not commutative nor associative. That is

$$a^b \neq b^a$$

$$a^{(b^c)} \neq (a^b)^c$$

Instead, we have different rules for exponent that we will prove

Property 10 (Exponent properties)**.** *We have the additive property of exponent*

$$a^{m+n} = a^m \times a^n$$

and the multiplicative property

$$(a^m)^n = a^{m \times n}$$

Proof. We will first prove the additive property by inducting on n. For the base case of $n = 0$,

$$a^{m+0} = a^m = a^m \times 1+ = a^m \times a^0$$

Now, we assume the claim holds for n and extends it to $n++$. That is we want to show that

$$a^{m+(n++)} = a^m \times a^{n++}$$

By the definition of addition, the left-hand side is equal to

$$a^{m+(n++)} = a^{(m+n)++} := a^{m+n} \times a$$

By the induction hypothesis,

$$a^{m+n} \times a = (a^m \times a^n) \times a$$

$$= a^m \times (a^n \times a) := a^m \times a^{n++}$$

Hence the additive property is proven. We will now prove the multiplicative property by induction on n. For the base case $n = 0$,

$$(a^m)^0 := 1 = a^{m \times 0}$$

Now, we assume the claim holds for n and extends it to $n++$. That is we want to show that

$$(a^m)^{n++} = a^{m \times (n++)}$$

The left-hand side is equal to

$$(a^m)^{n++} := (a^m)^n \times a^m$$

By the induction hypothesis, this is equal to

$$= a^{m+n} \times a^m$$

By the just proven additive property, we have

$$a^{m+n} \times a^m = a^{(m+n)+m}$$

By the definition of multiplication, this is equal to

$$= a^{m \times (n++)}$$

The left-hand side is equal to the right-hand side so the induction is complete. $\square$

1.5 Ordering

We now move on to the important question of ordering: what it means to say one number is greater than another and so on.

Definition 1.5.1 (Ordering of the naturals)**.** *For two natural number a and b, we define that a is greater or equal to b if and only if*

$$a \geq b \iff a = b + b', \text{ for some } b' \in \mathbb{N}$$

We also say that a is strictly greater than b if and only if

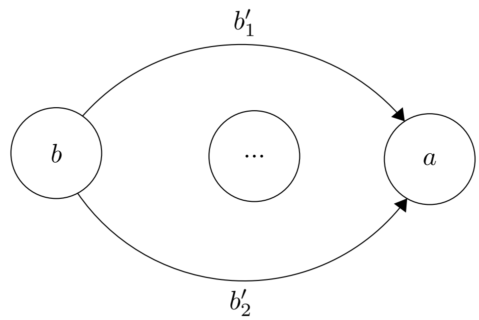
$$a > b \iff (a \geq b \wedge a \neq b)$$

ecause the definition of ordering involves the existence of a natural number b', we should prove that if this number exists, then it is unique.

Property 11 (Cancellation Law of naturals)**.** *We would prove that*

$$b + b_1' = b + b_2' \iff b_1' = b_2'$$

In other words, there could not be two separate paths starting from b and both path lead to the same number a.



Proof. We will first prove the "if" part. That is

$$b + b_1' = b + b_2' \implies b_1' = b_2'$$

As usual with properties concerning the natural number, we will prove this one using induction on b. For the base case $b = 0$,

$$0 + b_1' = 0 + b_2' \implies b_1' = b_2', \text{ by the additive identity}$$

Now, we assume the claim holds for b and extends it to $b{+}{+}$. That is we want to show that

$$((b++) + b_1' = (b++) + b_2') \implies b_1' = b_2'$$

By the definition of addition,

$$(b++) + b_1' = (b + b_1')++ = (b + b_2')++$$

By the unique predecessor axiom 4. we know that $m++ = n++$ implies that $m = n$. So we have

$$b + b_1' = b + b_2'$$

This, by our induction hypothesis, implies that $b_1' = b_2'$. Our induction is complete.

We now want to prove that if $b_1' = b_2'$, then $b_1' + b = b_2' + b$ for some $b \in \mathbb{N}$. This is an easy case of induction on b once again. For the base case $b = 0$, $b_1' := b_1' + 0 = b_2' + 0$. Now, we assume the claim holds for b and extends it to $b{+}{+}$. By the induction hypothesis, we know that if $b_1' = b_2'$ then

$$b + b_1' = b + b_2'$$

Thus, we have

$$(b++) + b_1' := (b + b_1')++ = (b + b_2')++$$

$$:= (b++) + b_2'$$

This implicitly uses the unique predecessor axiom 4, and our induction is complete. □

For convenience, let define the positive natural numbers as follows

Definition 1.5.2 (Positive natural number)**.** *A natural number a is called positive if $a \neq 0$*

We will now prove certain the closure-ness of positive numbers.

Property 12 (Positive closeness)**.** *If a be positive natural and b be natural (not necessarily positive), then*

$$a + b \text{ is positive natural}$$

Proof. Induct ob a. For the base case $a = 0$, $0 + b := b$ which is positive natural. Now, assume the claim for b, we will prove that $a + (b++)$ is positive natural. We have

$$a + (b++) := (a + b)++$$

Because of the induction hypothesis, $a+b$ is a natural number. Because of the axiom 3, 0 has no predecessor so $(a + b)++$ could not be 0. Hence, $(a + b)++$ is positive natural, and the claim is proven. $\square$

From property 12, we have this straightforward corollary

R For natural numbers a, b, we have

$$(a + b = 0) \iff a = b = 0$$

Proof. If $a = b = 0$, we have that $a + b = a + 0 := a = 0$. Now, turn to other direction. Now, because $a + b = 0$, by property 12 and proof by contradiction, we know that b must be 0. Similarly, because of the multiplicative commutativity, we can also show that $a = 0$. $\square$

Armed with the definition of positive natural numbers, we will now prove a property very similar to axiom 4.

Property 13 (Every positive natural has a predecessor)**.** *With S denoting the successor function, if a is positive natural, then there exists a natural number (not necessarily positive) b such that*

$$a = S(b)$$

In other words, the successor function is invertible for a

$$S^{-1}(a) = b$$

Proof. Notice that this is not quite axiom 4 which states that

$$S(m) = S(n) \iff m = n$$

Specifically, the part

$$S(m) = S(n) \implies m = n$$

$$\begin{tikzcd}
m \arrow[r] \arrow[d, bend right=60, "="'] & S(m) \arrow[d, bend left=60, "="] \\
n \arrow[r] & S(n)
\end{tikzcd}$$

assumes a priori that given a natural number k, k can be written in the form $k = S(m)$ i.e. $S^{-1}(k) = m$ exists. This certainly not true for 0 because $S^{-1}(0)$ does not exist by axiom 3.

We will prove this property by induction on a. For the base case $a = 0$, the statement is vacuously true because a is not a positive number.

$$(a \neq 0) \implies S^{-1}(a) \text{ exists}$$

evaluates to

$$\text{false} \implies \text{false}$$

which makes implication is true. If in doubt, look at the truth table for the implication $p \implies q$.

p	q	$p \implies q$
T	T	T
F	T	T
F	F	T
T	F	F

Now assume that the statement is true for a, we will prove it also true for $a{+}{+}$. Namely, there exists natural number b such that $a++ = S(b)$.This is trivial because all we need to do is to set $b = a$, then obviously $a++ = S(a)$. This proves the existence of the predecessor for positive natural. If we have existence, then axiom 4 provides that the predecessor is unique. □

Now, we are ready to verify the properties of *partial order* and *total order*.

Property 14 (Partial and Total order)**.** *We have these properties of ordering for the natural numbers with* $a, b, c \in \mathbb{N}$.

1. *Reflexive:* $a \geq a$
2. *Antisymmetry: If* $a \geq b$ *and* $b \geq a$*, then* $a = b$.
3. *Transitivity: If* $a \geq b$ *and* $b \geq c$*, then* $a \geq c$
4. *Connex property: Either* $a \geq b$ *or* $b \geq a$ *or both.*

Properties 1-3 define a partial order, while 2-4 define a total order.

Proof. Property 1 is easy. $a + 0 := a$ so $a \geq a$.

For property 2, because $a \geq b$, we know there exists a natural b' such that

$$a = b + b'$$

and because $b \geq a$, we also know that there exists a natural a' such that $b = a + b'$. Combined two equations, we have

$$a = a + a' + b' \implies a + 0 = a + a' + b'$$

Using the cancellation law, we know that

$$a' + b' = 0$$

As proven before, this only happens if $a' = b' = 0$. Therefore, $a = b + b' = b + 0 := b$. Property 2 is proven.

Property 3 is a simple substitution as well. Because $a \geq b$ and $b \geq c$, we know that

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

Due to additive closure, sum of two natural numbers $c' + b'$ is a natural number, so $a \geq c$ by definition of $\geq$ relationship.

Property 4 is the most challenging one. We will prove it by induction on b. For the base case $b = 0$, trivially $a \geq b$ because $a = b + b'$ with $b' = a$. Now assume that the statement is true for b, we will prove it also true for $b{+}{+}$. By the induction hypothesis, we know that at least one of $a \geq b$, $b \geq a$ must be true. Assume that $b \geq a$ is true. Then,

$$b = a + a', \text{ for some natural } a'$$

This means that

$$b++ = (a + a')++ := a + (a'++)$$

$(a'++)$ is a natural number due to axiom 2. Now, assume that $a \geq b$ is true. This means that either $a = b$ or $a > b$. If $a = b$, then

$$(b++) = a + 1$$

so $b \geq a$. Otherwise, assume that $a > b$. Then, it must be the case that

$$a = b + b' \text{ with } b' \neq 0$$

Otherwise, if $b' = 0$, then by the definition of addition, we deduce that $a = b$, which is a contradiction to $a > b$. Because b' is positive, by property 13, we know that there exists k such that $S(k) = b'$. Therefore, we have

$$a = b + (k + 1) = (b + 1) + k = (b++) + k \text{ with natural } k$$

Thus, we have $a \geq (b++)$. This proves that

$$((a \geq b) \vee (b \geq a)) \implies ((a \geq b++) \vee (b++ \geq a)), \text{ where } \vee \text{ is the logical "or" symbol}$$

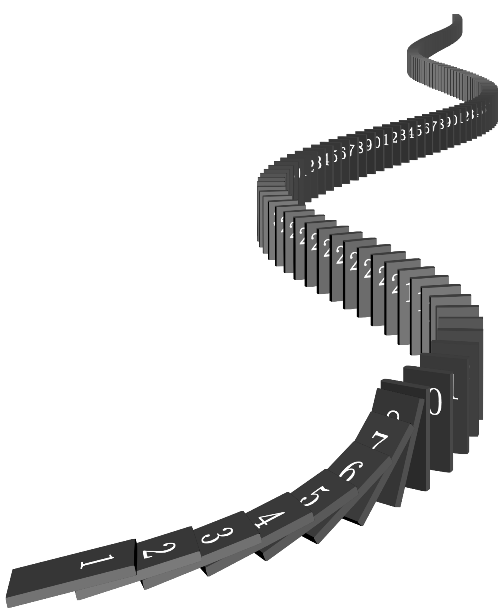
This concludes our induction and proves connex property. $\square$

Based on the proof for connex property, we can also prove the following

Property 15 (Trichotomy)**.** *For any natural numbers a, b, exactly one of the followings is true*

- $a = b$
- $a > b$
- $b > a$

1.6 Aside: Strong and backward induction



1.6.1 Strong induction

Axiom 5 establishes the principle of weak induction. For a property p, if $p(0) \wedge (p(n) \implies p(n++))$, then $p(n)$ is true $\forall n \in \mathbb{N}$ ($\wedge$ means "and" in logic).

Now we move to another kind of induction: strong induction. The claim is that if you assume $p(0), p(1), ..., p(n-1)$ are all true, and are able to deduce that $p(n)$ is also true, then $p(n)$ is true $\forall n \in \mathbb{N}$. Obviously, this induction is "stronger" than our usual induction as it allows us to assume that $p(0)$ up to $p(n-1)$ true. In weak induction, we merely assume $p(n-1)$ is true. Also, note that we do not need to prove the base case. This is because $p(0)$ is vacuously true given $(p(0) \wedge p(1)... \wedge p(n-1)) \implies p(n))$. Note that a vacuous truth is a statement that asserts that all members of the empty set have a certain property. This statement is automatically true. In other words, properties that can be said about members of an empty set (empty set has no member!) are assumed to be true for those non-existent members.

Theorem 2 (Strong induction)**.** *Let p be a property of the natural numbers. For an arbitrary $n \in \mathbb{N}$, assume $p(m)$ is true for all natural $m < n$. If we can deduce that $p(n)$ is also true, then we can conclude that $p(n)$ is true for all natural number n.*

Proof. Strong induction turns out to be just regular induction in disguise. Specifically, we can define a new property $Q(n)$ as the property that $p(m)$ is true for all natural $m < n$. In other words, $Q(n)$ is the intersection $p(0) \wedge p(1)... \wedge p(n-1)$. We need to prove $Q(n)$ is true for all $n \in \mathbb{N}$ and this will automatically imply that $p(n)$ is true for all $n \in \mathbb{N}$. We demonstrate this by induction.

First, we need to prove that $Q(0)$ is true. In other words, $p(m)$ is true for all natural numbers $m < 0$. This is vacuously true because the set of all natural numbers less than 0 is the empty set $\emptyset$. Every property that is attributed to $\emptyset$ is assumed to be true. Now for the inductive step, we prove that $Q(n) \implies Q(n+1)$. This is when our hypothesis for strong induction comes in. For strong induction, we assume that we are able to prove $(p(0) \wedge p(1)... \wedge p(n-1) \implies p(n))$. This also means that $Q(n) \implies p(n)$. And because $Q(n+1) := (p(0) \wedge p(1)... \wedge p(n-1)) \wedge p(n) := Q(n) \wedge p(n)$, this shows that $Q(n) \implies Q(n+1)$. We're done by weak induction, and we conclude that $Q(n)$ is true for all natural number n. This also means that $p(n)$ is true for all natural n. □

Of course, we do not need to start from 0. We can pick a new starting point $m_0 \in \mathbb{N}_{\neq 0}$. If we can prove that $p(m) \implies p(n)$ for $m_0 \leq m < n$, then we can conclude that $p(m)$ is true for all natural numbers starting from m_0 i.e. $m \geq m_0$.

1.6.2 Backward induction

Theorem 3 (Backward induction)**.** *If we can prove $p(m+1) \implies p(m)$ for all natural number m and that $p(m_0)$ is true for some starting point natural number m_0, then we can conclude that $p(m)$ is true for all natural $m \leq m_0$.*

Proof. Again, we can reduce this new type of induction to just regular weak induction. We will prove the validity of backward induction by using weak induction on m_0. For the base case $m_0 = 0$, If we know that $p(m_0) = p(0)$ is true, then it is certainly true that $p(m)$ is true for all $m \leq 0$.

Now, we assume that the principle of backward induction is valid up to m_0, we will prove that the principle is also valid for $m_0 + 1$. In the language of logic, we assume that

$$(p(m_0) \wedge (p(m+1) \implies p(m))) \implies (p(m), \forall m \leq m_0) \ (1)$$

And we want to prove that

$$(p(m_0+1) \wedge (p(m+1) \implies p(m))) \implies (p(m), \forall m \leq m_0 + 1) \ (2)$$

From $(p(m_0 + 1) \wedge (p(m + 1) \implies p(m)))$, we know that $p(m_0)$ is true. This coupled with the induction hypothesis (1) leads to $p(m)$ is true for all $m \geq m_0$. We also assume in (2) that $p(m_0+1)$ is true. This shows that $p(m)$ is true for all $m \leq m_0 + 1$. This concludes our induction. We conclude that backward induction is valid for all $m_0 \in \mathbb{N}$. $\square$

Chapter 2

Integers

We will now extend our number set from the natural to integers. Integers will extend the natural to include negative integers

$$..., -2, -1, 0, 1, 2, ...$$

This sequence extends in both left and right directions. To construct the integers, we will use equivalence classes and formal difference. Note that throughout many proofs in this chapter, we will rarely see the ubiquitous of induction as we have seen in natural numbers. The golden age of induction, with its clear advantage in proofs concerning the natural numbers, has indeed passed.

2.1 Equivalence Classes

Harvey Mudd lecture introduces equivalence classes.

Firstly, let us start with some preliminary definition

Definition 2.1.1 (Ordered tuple)**.** *An ordered of n-tuple is the sequence $(a_1, a_2, ..., a_n)$. For example, a 2-tuple is the ordered pair (a, b) of two objects a, b. Two n-tuple $(x_1, ..., x_n)$ and $(y_1, ..., y_n)$ are equal if and only if $x_i = y_i$ for all $i = 1, 2, ..., n$. Therefore, the order in which the objects appear is important.*

Using the definition of 2-tuple, we can define the Cartesian product as follows

Definition 2.1.2 (Cartesian product)**.** *The Cartesian product of a non-empty set A is the set of 2-tuple such that both the first and the second components of the tuple are elements of the set A.*

$$A \times A = \{(a_i, a_j) : a_i, a_j \in A\}$$

Example 3:
The Cartesian product of the set $A = \{1, 2, 3\}$ is $\{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)\}$. Visually,

A				
		1	2	3
A	1	(1,1)	(1,2)	(1,3)
	2	(2,1)	(2,2)	(2,3)
	3	(3,1)	(3,2)	(3,3)

Recall that a binary relation is a function that takes in two inputs and produce either true or false. That is, a relation R is a boolean-valued function that takes in 2-tuple $R : (x, y) \to \{\text{true, false}\}$. An equivalence relation defined on a set A is then a subset of the Cartesian product $A \times A$ where $p(a_i, a_j)$ is true. Recall from definition 1.1.1, that an equivalence relation $p(x, y)$ is a relation that is reflexive, transitive and symmetric.

We can now define an equivalence class

Definition 2.1.3 (Equivalence class)**.** *Given a set A and an element $a \in A$, an equivalence class containing a is*

$$\{x \in A : p(a, x) = true\}$$

With these definitions, we can now define the integers in the next section.

2.2 Formal difference

In hindsight knowledge, we know that the integer set is nothing but a set of natural numbers plus some "negative" integers, whatever "negative" might mean. We would inevitably have to introduce the idea of subtraction. But now, we will follow Tao's approach in his book *Analysis*. We would introduce a meaningless symbol $--$, and define the integers as all number of the form $a - -b$, where a, b are our known natural numbers.

Essentially, an integer is a pair of two natural numbers. This definition is rigorous because we know what natural numbers are. We postulate the existence of natural numbers, and we also define rigorously the idea of 2-tuple. So essentially, the integer set is the Cartesian product of the natural numbers $\mathbb{N} \times \mathbb{N}$. We also need a mechanism to tell when two integers are equal to each other as well.

Definition 2.2.1 (Integers)**.** *With the fictitious and meaningless symbol $--$ denoting the "formal difference" and $a, b \in \mathbb{N}$, we define an integer is one of the form*

$$a - -b$$

In other words, x is an integer if $x \in \mathbb{N} \times \mathbb{N}$.

Definition 2.2.2 (Equal integers)**.** *With $a, b, c, d \in \mathbb{N}$, we say that two integers are equal to each other*

$$a - -b = c - -d$$

if and only if

$$a + d = b + c.$$

We know $a + c \in \mathbb{N}$, and $b + d \in \mathbb{N}$. We also know how to tell when two natural numbers are equal to each other. Thus, we should also know when two integers are equal to each other.

We should check that this definition of equality satisfy all three properties of equivalence relation. Reflexive and symmetric are trivial. We will show that under 2.2.2 transitivity holds. That is

$$((a - -b) = (c - -d) \wedge (c - -d) = (e - -f)) \implies (a - -b) = (e - -f)$$

From $(a - -b) = (c - -d)$, we know that

$$a + d = b + c$$

And from $(c - -d) = (e - -f)$, we know that

$$c + f = d + e$$

From the Cancellation Law property 12, we know that $a = b \implies a + c = b + c$. Therefore, adding two equations above we have

$$a + d + c + f = b + c + d + e$$

We have not define subtraction. Thus, to eliminate common terms from both sides of the equation, we invoke the Cancellation law 12 again.

$$a + f + (c + d) = b + e + (c + d) \implies a + f = b + e$$

This by the equivalence relation if integers show that indeed $a - -b = e - -f$.

With the definition of equivalent integers, we can see the equivalence classes.

	0	1	2	3	...
0	0–0	0–1	0–2	0–3	...
1	1–0	1–1	1–2	1–3	...
2	2–0	2–1	2–2	2–3	...
3	3–0	3–1	3–2	3–3	...
...	...	...	...	...	...

Figure 2.1: Equivalence classes of $\mathbb{Z}$

All the unicolor diagonals of the table represent different equivalence classes. For example, all integers

$$0 - -0 = 1 - -1 = 2 - - - 2 = 3 - -3 = ...$$

forms an equivalence class. From the hindsight knowledge of subtraction, we know this makes sense as all our these subtraction results in 0. We will now proceed to define common operations on the integers.

2.3 Addition

Definition 2.3.1 (Integer addition)**.** *For two integers $a - -b$ and $c - -d$, we define*

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

ince we essentially reduce the problem of adding two integers to a problem of adding four natural numbers, we can easily the group axiom

Theorem 4 (Additive group of the integers)**.** *The integers under addition $(\mathbb{Z}, +)$ forms a group i.e.*

1. *(Identity) There exists a unique identity element $0 \in \mathbb{Z}$ such that $0 + x := x + 0 := x$ for all $x \in \mathbb{Z}$.*
2. *(Closure) For $x, y \in \mathbb{Z}$, $x + y \in \mathbb{Z}$.*
3. *(Associativity) For $x, y, z \in \mathbb{Z}$, we have*

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

4. *(Inverse) For all $x \in \mathbb{Z}$, there exists a unique inverse element $x^{-1} \in \mathbb{Z}$ such that*

$$x + x^{-1} = x^{-1} + x = 0$$

Proof. Let $x = a - -b$. First, we will define the identity as the equivalence class

$$0 = \{n - -n : n \in \mathbb{N}\}$$

We will verify that this definition indeed satisfies the identity property of group.

$$(a - -b) + (n - -n) := (a + n) - -(b + n)$$

By the definition of equivalent integers, $(a + n) - -(b + n) = a - -b$ when

$$(a + n) + b = a + (b + n)$$

This is true due to the algebraic properties of natural addition. Similarly, we can also easily verify that $(n - -n) + (a - -b) = a - -b$.

Closure property is also quite obvious. Let $y = c - -d$. By definition of integer addition, we have

$$x + y := (a + c) - -(b + d)$$

Let $a+c = n$ and $b+d = m$. Due to the closure of natural addition, n and m must also be natural numbers. Thus, $x + y = n - -m$ is an integer.

Let $z = e - -f$, we will verify the associative property.

$$x + (y + z) = (a - -b) + ((c - -d) + (e - -f))$$

$$:= (a - -b) + ((c + e) - -(d + f))$$

$$:= (a + (c + e)) - -(b + (d + f)))$$

$$= ((a + c) - -(b + d)) + (e - -f) = (x + y) + z$$

We can see that the associativity of integer addition comes straight from the associativity of natural addition.

Now, we will turn to the trickiest property: inverse. We want to find $x^{-1} = c - -d$ such that

$$(a - -b) + (c - -d) = 0 = n - -n$$

Follows the definition of addition, we demand

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

By the definition of equivalence relation, this is equivalent to

$$a + c + n = b + d + n$$

Using the principle of cancellation, we reduce it to

$$a + c = b + d \quad (1)$$

a, b are two variables that are dictated to us by x. c, d, on the other hand, are those that we can control, those that we seek. Thus, we should reduce two uncontrolled variables to only one for ease. To do this, we use the trichotomy of the naturals. If $a = b$, then

$$x^{-1} = 0 = \{c - -c : c \in \mathbb{N}\}$$

The equation (1) is obviously satisfied. If $a > b$, then $a = b + b'$. The equation (1) becomes

$$b + b' + c = b + d$$

Using cancellation law, we have

$$b' + c = d$$

Therefore, we have

$$x^{-1} = \{c - -(b' + c) : c \in \mathbb{N}\}$$

Similarly, if $b > a$, then

$$x^{-1} = \{(a' + d) - -d : d \in \mathbb{N}\}$$

We can later verify that addition in integer is commutative so $x + x' = x' + x = 0$. Therefore, x^{-1} is an appropriate inverse. Actually, the process of finding the inverse by looking at different ordering relation between a and b here is not totally necessary. In fact, we can just flip the pair. We can easily verify that $b - -a$ is an inverse of $a - -b$.

In our definition. we require that $x + 0 = 0 + x$ and that $x + x^{-1} = x^{-1} + x$. That is binary operation between x and the identity or between x and its inverse is commutative. We know that binary operation doesn't have to be commutative in general. In fact, the set of all invertible matrices under matrix multiplication forms a group. But we know that matrix multiplication is not commutative i.e. $A \times B \neq B \times A$ for matrices A, B. But when applying to the identity and inverses, we know that $A \times I = I \times A = A$, and $A \times A^{-1} = A^{-1} \times A = I$ with the identity matrix I. Question like does $A \times I = A \iff I \times A = A$ and $A \times B = I \iff B = \times A = I$ are of pure group theory interest, and will not be pursued here.

In general, if we assume that $x + 0 = x \iff 0 + x = x$ and $x + x^{-1} = 0 \iff x^{-1} + x = x$, then identity and inverse elements are unique for group. For identity elements, suppose we have two different identities 0 and $0'$ such that

$$x + 0 = x + 0'$$

because each group element has an inverse, this is equivalent to

$$(x^{-1} + x) + 0 = (x^{-1} + x) + 0' \implies 0 = 0'$$

a contradiction. So, the identity element must be unique. Similarly, suppose we have two different inverses x_1' and x_2' of x, then

$$x + x_1^{-1} = x + x_2^{-1} = x$$

$$\implies (x_1^{-1} + x) + x_1^{-1} = (x_1^{-1} + x) + x_2^{-1}$$

$$\implies x_1^{-1} = x_2^{-1}$$

In our additive group of the integers, identity and inverses are "unique" up to equivalence classes, because we have define the notion of equality as equivalence relation. □

Not only $(\mathbb{Z}, +)$ is a group but it is also a special kind of group – Abelian group.

Definition 2.3.2 (Abelian group)**.** *$G = (S, \circ)$ is a group. If G is commutative under $\circ$ then we call G an Abelian group. That is G is Abelian when*

$$x \circ y = y \circ x, \text{ for } x, y \in S$$

Theorem 5. *Integers under addition form Abelian group*

Proof. We already show $(\mathbb{Z}, +)$ is a group so we only need to show that $x + y = y + x$ for $x, y \in \mathbb{Z}$. This is quite trivial. Let $x = a - -b$ and $y = c - -d$, we have

$$x + y := (a + c) - -(b + d) = (c + a) - -(d + b) = y + x$$

□

We will now move on to multiplication.

2.4 Multiplication

Definition 2.4.1 (Integer multiplication)**.** $(a--b)\times(c--d) := (ac+bd)--(ad+bc)$

Like natural multiplication, integer multiplication has many properties that we would characterize them as a monoid.

Definition 2.4.2 (Monoid)**.** $(S,\circ)$ *is a monoid if it*

1. *has an identity element.*
2. *is associative*
3. *is closed under* $\circ$

That is a monoid is a group without the inverse property.

Theorem 6. *Integers under multiplication is a monoid.*

Proof. First, we will identify the identity element as

$$\{(a++)--a : a\in\mathbb{N}\}$$

Think of it as the equivalence class for the number 1. We will check this is indeed the identity element. Let $x = b--c$.

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

We have

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

and

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

To prove that $x\times((a++)--a) = x$, we need to show that

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

This is true.

Next, we will prove that multiplication is associative. Let $x = a--b, y = c--d, z = e--f$, we aim to prove that

$$x(yz) = (xy)z$$

$$(a--b)((ce+df)--(cf+de)) = ((ac+bd)--(ad+bc))(e--f) \quad (1)$$

The left-hand side is equal to

$$(a(ce+df)+b(cf+de))--(a(cf+de)+b(ce+df))$$

$$= (ace+adf+bcf+bde)--(acf+ade+bce+bdf)$$

(Note that we have $8 = 2^3$ terms in total. This makes sense when we consider the product $xyz = (a-b)(c-d)(e-f)$, at the first bracket we have two choices: a or b, at the second: c or d, and at the third: e or f. So we have $2\times2\times2 = 8$ terms in total.)

We can expand the right-hand side of (1) and see that it matches with the left-side expansion. This is tedious and is left as an exercise for the reader.

Finally, we have closure because

$$x \times y = (a - -b) \times (c - -d) = (ac + bd) - -(ad + bc)$$

We know that both $ac+bd$ and $ad+bc$ are natural numbers due to the closure of addition and multiplication. Thus, $x \times y$ is also an integer because it is the formal difference of a pair of naturals. $\square$

2.5 Commutative Ring

We introduce the concept of a ring and show that the integers form a commutative ring.

Definition 2.5.1 (Ring). *A set R with two binary operations $+$ and $\times$ is a ring if satisfying*

1. *R is an Abelian group under $+$.*
2. *R is a monoid under $\times$.*
3. *$\times$ is distributive over $+$. i.e.*

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

A ring is commutative if $\times$ is commutative.

e will need to show distributivity and multiplicative commutativity of the integers.

Property 16 (Distributivity). *For $x, y, z \in \mathbb{Z}$, we have*

$$x \times (y + z) = x \times y + x \times z$$

Proof. Let $x = a - -b, y = c - -d, z = e - -f$, we have

$$x \times (y + z) := (a - -b) \times ((c + e) - -(d + f))$$

$$:= (a(c + e) + b(d + f)) - -(a(d + f) + b(c + e)$$

Using the algebraic properties of the naturals including the distributive law to rearrange terms, we have

$$= ((ac + bd) + (ae + bf)) - -((ad + bc) + (af + be))$$

$$= x \times y + x \times z$$

$\square$

Property 17 (Commutativity). *For $x, y \in \mathbb{Z}$, we have*

$$x \times y = y \times x$$

Proof. Proof by intimidation invented by Croatian-American mathematician William Feller in his lectures. The proof goes as following

Any moron will know that ...

The rest of the proof is left as an exercise for the reader. $\square$

2.6 Natural embedding in the integers

We will now show that the natural numbers are, in some sense, a subset of the integers. We define a concept in group theory: homomorphism

Definition 2.6.1 (Group homomorphism). *Let $(G, \circ)$ and $(H, \diamond)$ be two groups. If there exists a function $f : G \to H$ such that*

$$f(a \circ b) = f(a) \diamond f(b), \text{ for all } a, b \in G$$

Then, f is called a homomorphism between G and H.

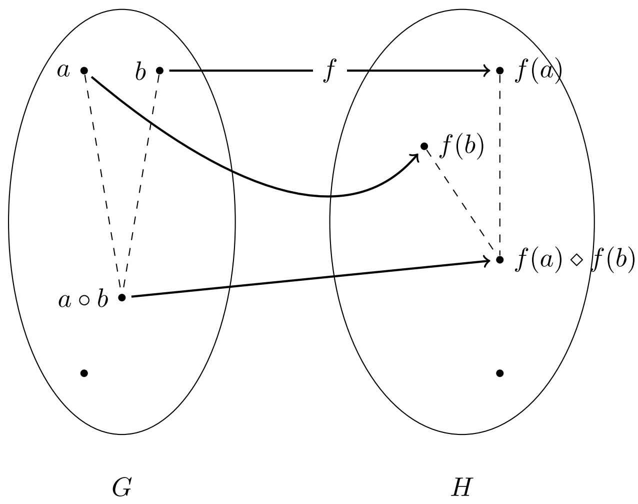


Figure 2.2: Group homomorphism

Because the natural number is not technically a group, we will use the word pseudo-homomorphism to refer to a function $f : \mathbb{N} \to \mathbb{Z}$ such that

$$f(m+n) = f(m) + f(n), \text{ for all } m, n \in \mathbb{N}$$

The mapping will be defined as follows

Definition 2.6.2 (Natural embedding in the integers)**.** *For each $n \in \mathbb{N}$, the corresponding element in $\mathbb{Z}$ is the equivalence class of $n - -0$. In other words, the pseudo-homomorphism $f : \mathbb{N} \to \mathbb{Z}$ is*

$$f(n) = n - -0, \text{ for all } n\mathbb{N}$$

This pseudo-homomorphism is valid for both addition and multiplication. That is

$$m + n = (m - -0) + (n - -0)$$

and

$$m \times n = (m - -0) \times (n - -0)$$

We cam verify these two equations very quickly using the operations' definitions. Due to this homomorphism, we can think of any given natural number as strictly a natural number or as in terms of its corresponding integer. For example, we can think of the number $2 \in \mathbb{N}$ as 2 or as $2 - -0, 3 - -1, 4 - -2, ... \in \mathbb{Z}$. It does not matter much since all the usual operations are the same. To add two natural numbers 2 and 3 for example, we can think of it as adding 2 and 3 in the natural which yields 5. Equivalently, we can think of the addition as adding $(2, 0) = 2 - -0$ and $(3, 0) = 3 - -0$, which yields $(5, 0) = 5 - -0$ the integer counterpart of the natural number 5.

Strictly speaking, $\mathbb{N}$ is **not** really a subset of $\mathbb{Z}$. This is because $\mathbb{Z}$ is essentially the Cartesian product $\mathbb{N} \times \mathbb{N}$. That is $\mathbb{Z}$ and $\mathbb{N}$ are two totally different types of objects: each element in $\mathbb{N}$ consists of a single number while each element of $\mathbb{Z}$ is a 2-tuple of some elements in $\mathbb{N}$. But through this pseudo-homomorphism, $\mathbb{N}$ is just as good as a subset of $\mathbb{N}$.

2.7 Subtraction

Now because of the additive integer inverse, we have a new notion of subtraction, which is not possible under the natural number set.

Definition 2.7.1 (Integer subtraction). *For $x, y \in \mathbb{Z}$, we define subtraction as follows*

$$x - y = x + y^{-1}$$

R [Closure of subtraction] For any $x, y \in \mathbb{Z}$, $x - y \in \mathbb{Z}$.

Proof. This is because $x - y := x + y^{-1}$. By the group inverse property of integers, we know that $y^{-1} \in \mathbb{Z}$. Additive closure leads us to conclude that $x + y^{-1} \in \mathbb{Z}$. □

Now, we can verify that our "formal difference" is just subtraction. In other words, if $a, b \in \mathbb{N}$ and f is the pseudo-homomorphism between $\mathbb{N}$ and $\mathbb{Z}$ and let $x = a - -b \in \mathbb{Z}$, then we have

$$x = f(a) - f(b)$$

Just to quickly check,

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

Thus, from now on, we can replace the formal difference by just our regular subtraction operation.

2.8 Ordering

Similar to the ordering of the naturals, we now define some ordering relation on the integers.

Definition 2.8.1 (Positive/Negative). *An integer $x = a - -b$ is called positive if $a > b$. If $a = b$, then as we defined before, x is 0. Otherwise, if $a < b$, then x is negative.*

Theorem 7 (Sign of integers). *For any given integer x, we can expect exactly one of the following case to be true*

- *x is positive*
- *x is negative*
- *x is 0*

As we know from the trichotomy, given two natural numbers a and b, exactly one of the following must be true

- $a > b$
- $b > a$
- $a = b$

Thus, the trichotomy of the integers follow directly from that of the naturals.

Definition 2.8.2 (Trichotomy of integers). *For $x, y \in \mathbb{Z}$, we define*

- *$x > y$ if and only if $x - y$ is positive*
- *$x = y$ if and only if $x - y = 0$*

We know that exactly one of the following must be true

1. *$x > y$*

2. $x = y$

3. $y > x$

Proof. This is quite straightforward. We define the ordering between x and y in terms of the sign of $x - y$. Because of the closure of subtraction, $x - y$ is an integer. Any given integer must either be 0, positive, or negative, leading uniquely to one ordering between x and y. □

2.9 Substitution axiom

Tao in page 329 A.7: Equality of *Analysis* defines a substitution axiom.

Definition 2.9.1 (Substitution axiom)**.** *If x, y are objects of the same type, and $x = y$, then $f(x) = f(y)$ for all functions or operations f.*

This coupled with the three properties of equivalence relation define an equality relation. This substitution axiom should be very straightforward to verify for the natural numbers based on axiom 4 and induction. We will now quickly verify this axiom for integers in the case of addition, subtraction and multiplication.

For addition, assume that $a - -b = a' - -b'$, we want to prove that

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

Theis equivalent to

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

By the definition of equivalence in integers, this is only true when

$$a + c + b' + d = a' + c + b + d$$

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

This is true because $a - -b = a' - -b'$.

For subtraction, we want to prove that

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

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

This is reduced to the case of addition above. The other case of $(c - -d) - (a - -b) = (c - -d) - (a' - -b')$ can be proved similarly.

For multiplication, we want

$$(a - -b) \times (c - -d) = (a' - -b') \times (c - -d)$$

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

The left-hand side is equal to the right-hand side if

$$ac + bd + a'd + b'c = a'c + b'd + ad + bc$$

Using the distributive law of the natural numbers, this is equal to

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

Because $a - -b = a' - -b'$, we know that $b' + a = a' + b$ and the claim is proven.

Chapter 3

Rational Numbers

Note that $\mathbb{Z}$ is a group under addition. However, it is not a group under multiplication. The only thing that prevents it from becoming so is the lack of multiplicative inverse. For inverse of multiplication to make sense, we need to introduce the rational numbers and define division.

3.1 Formal division

Similar to how we build the integers from the naturals, we will build the one rational from a pair of two naturals.

Definition 3.1.1 (Rational Numbers)**.** *Let the meaningless // symbol denote "formal division". For any $x, y \in \mathbb{Z}$ and $y \neq 0$, object of the form $x//y$ is defined to be a rational number.*

In other words, the rational number set $\mathbb{Q}$ is the Cartesian product $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$.

We will now define the equivalence classes from our hindsight knowledge of how rational numbers work.

Definition 3.1.2 (Equivalent rationals)**.** *Two rationals number $a//b$ and $c//d$ are equal if and only if*

$$ad - bc = 0$$

This is a rigorous definition since we know how to multiply, subtract two integers and tell when an integer is equal to 0.

	1	2	3	4	...
0	0 //1	0 //2	0 //3	0 //4	...
1	1//1	1//2	1//3	1//4	...
2	2//1	2//2	2//3	2//4	...
3	3//1	3//2	3//3	3//4	...
...	...	...	...	...	...

3.2 Addition

3.3 Multiplication

3.4 Ordering

3.5 Integer embedding in the rationals

3.6 Ordered field

3.7 Reciprocal and Field

$\mathbb{Q}$ is an ordered field.

Chapter 4

Real Numbers

4.1 Cauchy Sequence

4.1.1 Absolute value

4.1.2 ϵ-closeness

4.1.3 Definition and Properties

4.2 Real Numbers

4.2.1 Definition

4.2.2 Addition

4.2.3 Multiplication

To prove $xy \in \mathbb{R}$ given that $x \in \mathbb{R}$ and $y \in \mathbb{R}$. Note that the properties that the Cauchy sequence is always bounded is very important. So we can pick an arbitrary ϵ and find one of the many possible bounds for the sequence. We then can replace any $|a_n|$ and $|b_n|$ by the bound so that we no longer need to use $a_n = f(n) = f(n(\epsilon))$ (A sequence by definition is $f : \mathbb{N} \to \mathbb{Q}$), which is dependent on ϵ.

exercise 5

NOTE: definition of Cauchy sequence VS of two equivalent Cauchy sequences are VERY different.

4.2.4 Group Homomorphism: Embedding Rationals into the Reals

There's no exception. Such time is not finite

Chapter 5

Series

First, just to remind the readers of convergence and absolute convergence. Sequence (a_n) is conditionally convergent to L if $(a_n) \to L$. We say (a_n) is conditionally convergent to L if $|a_n| \to L$. Absolute convergence is a stronger condition than conditional convergence. If (a_n) is absolute convergent then it must conditionally convergent as well. The converse is not true however. Think of absolute convergence to conditional convergence as uniform continuity to regular point-wise continuity.

We will now define the notion of finite series, which is the sum of a finite sequence. Given a finite sequence $(a_n)_{n=1}^N$, the series

$$\sum_{i=1}^{N} a_n$$

is defined recursively as follow

$$\sum_{i=1}^{N} a_n = 0, \text{ whenever } m > N.$$

Otherwise,

$$\sum_{i=m}^{N+1} a_n = \sum_{i=m}^{N} a_n + a_{N+1}$$

We also have the definition for infinite series for the sequence $(a_n)_{n=1}^{\infty}$. Define the partial sum

$$S_N = \sum_{i=1}^{N} a_n$$

We say that the infinite sum

$$\sum_{i=1}^{\infty} a_n$$

is finite and is equal to a real number L when the sequence of partial sums $(S_n)_{n=1}^{\infty}$ is convergent to L. Otherwise, the infinite sum is divergent.

We have a couple of important theorems for mostly re-arranging terms. Firstly, for finite series. re-arranging terms do not change the values of the series. Concretely for the sequence $(a_n)_{n=1}^N$ if $f(n)$ is the bijection from $\{1, ..., N\}$ to $\{1, ..., N\}$. then we claim

$$\sum_{i=1}^{N} a_n = \sum_{i=1}^{N} a_{f(n)}$$

For example takes N and $f(1) = 2, f(2) = 3, f(3) = 1$, then the claim is

$$a_1 + a_2 + a_3 = a_2 + a_3 + a_1$$

Rigorous proof is required here.

Unfortunately, this is not the case with infinite series. For infinite series, you can get different results from re-arranging terms. For example, consider the series

$$S = 1 - 1 + 1 - 1 + 1 - 1 + ...$$

If we carry out naively, then

$$S = (1 - 1) + (1 - 1) + (1 - 1) + ... = 0$$

However, if we regroup them, we have

$$S = 1 - (1 - 1 + 1 - 1 + ...) = 1 - S$$

$$\implies S = \frac{1}{2}$$

This is because we can quickly notice that S is not convergent in the first place. The sequence of partial sums look like $1, 0, 1, 0, 1, ...$ which is not convergent.

Theorem 8 (Re-arrange infinite series)**.** *Let $f : \mathbb{N} \to \mathbb{N}$ be a bijection. Consider the sequence $A = (a_n)_{n=0}^{\infty}$. If A is* ***absolutely convergent*** *then*

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

It is not surprising that we can take the sum over a countable set as well. Suppose X is infinitely countable, then X can be put into an one-to-one correspondence with the natural numbers $\mathbb{N}$ by a bijection a. That is for all $x \in X$, we have $a(x) = n$ for $n \in \mathbb{N}$. Let define $(a_n)_{n=1}^{\infty}$ such that $a_n = x$ for $a(x) = n$. Therefore,

$$\sum_{x \in X} x = \sum_{i=1}^{\infty} a_i$$

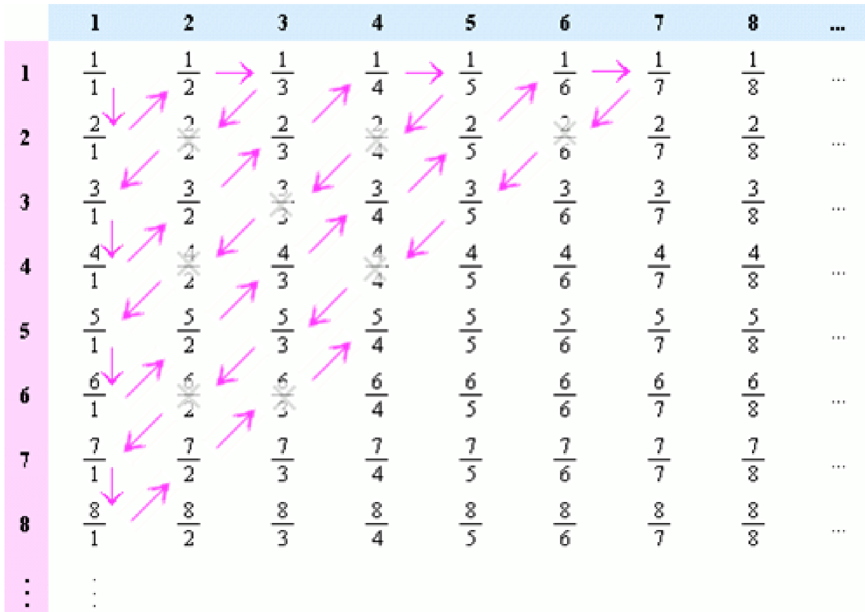
This just shows that we can in fact sum over a countable set. Now, we move to an important theorem about double sums

Theorem 9 (Fubini's theorem of double sums)**.** *Given two infinitely countable sets X and Y, and let $f : X \times Y \to \mathbb{R}$. If $\displaystyle\sum_{(x,y) \in X \times Y} f(x, y)$ is absolute convergent then*

$$\sum_{(x,y) \in X \times Y} f(x, y) = \sum_{(y,x) \in Y \times X} f(x, y)$$

$$= \textstyle\sum_{x \in X} \sum_{y \in Y} f(x, y) = \sum_{y \in Y} \sum_{x \in X} f(x, y)$$

First, note that $\sum_{(x,y) \in X \times Y} f(x, y) = \sum_{(y,x) \in Y \times X} f(x, y)$ are not at all trivial. First, recall that if X and Y are countable then $X \times Y$ is countable. This is a pictorial proof that the rational number, which is essentially the Cartesian product of $\mathbb{N}$ is countable.



We prove the countability by traversing the table in a zig-zag formation, hitting every diagonals. It establishes a bijection between $\mathbb{N}\times\mathbb{N}$ and $\mathbb{N}$. Now, consider a finite Cartesian product $X\times Y$ and the zig-zag formation.

$$\begin{tikzcd}
X/Y & 4 & 5 & 6 \\
1 & (1,4) \arrow[r] & (1,5) \arrow[dl] & (1,6) \arrow[d] \\
2 & (2,4) \arrow[d] & (2,5) \arrow[ur] & (2,6) \arrow[dl] \\
3 & (3,4) \arrow[ur] & (3,5) \arrow[r] & (3,6)
\end{tikzcd}$$

Instead of $X\times Y$, we consider $Y\times X$ and the same zig-zag formation.

$$\begin{tikzcd}
Y/X & 1 & 2 & 3 \\
4 & (1,4) \arrow[r] & (2,4) \arrow[dl] & (3,4) \arrow[d] \\
5 & (1,5) \arrow[d] & (2,5) \arrow[ur] & (3,5) \arrow[dl] \\
6 & 1,6) \arrow[ur] & (2,6) \arrow[r] & (3,6)
\end{tikzcd}$$

Note that the table for $Y\times X$ if thought as a matrix is the transpose of $X\times Y$. Also note that the order in which each pair (x,y) is accessed has changed. For example, in $X\times Y$ we sum will sum $f(1,4)+f(1,5)+...$ but in $Y\times X$ we sum $f(1,4)+f(2,4)+....$ Re-arranging terms like this is fine for finite series but we might run into problem for infinite series.

Chapter 6

Compactness

Theorem 10 (Nested interval theorem)**.** *In a bounded interval $I_0 = [a, b]$ of $\mathbb{R}$, we build nested intervals $I_n := [a_n, b_n]$ such that $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$. Then*

$$\bigcap_{i=0}^{\infty} I_n \neq \emptyset$$

Proof. We will show that $L = sup(a_n)$ is in $\bigcap_{i=0}^{\infty} I_n$. Note that by definition of being upper bound $L \geq a_n$ for all n. Also, $L \leq b_n$ for all n. Otherwise, if $L > b_n$ for some n, then note that by definition of nested interval $b_n \geq a_n \geq a_{n-1} ... \geq a_0$, and $b_n \geq b_{n+1} \geq a_{n+1}$, $b_n \geq b_{n+1} \geq b_{n+2} \geq a_{n+2}$ and so on. In other words, b_n is an upper-bound, which is less than L a contradiction of least upper bound. Then, $L \leq b_n$ for all n as claimed. We now can see that $a_n \leq L \leq b_n$ for all n, so L is in every I_n. Thus, L is in the intersection. □

Theorem 11 (Cantor intersection theorem)**.** *Suppose (C_k) is a sequence of nested non-empty compact subsets of a metric space X*

$$C_1 \supset C_2 \supset C_3 ...$$

it follows that

$$\bigcap_{i=1}^{\infty} C_n \neq \emptyset$$

Proof. Let the complement $K_n = C_n^c := X \setminus C_k$. We see that K_n is open because compact sets are closed. For the sake of contradiction, assume that $\bigcap_{i=1}^{\infty} C_n = \emptyset$. Then, $\bigcap_{i=1}^{\infty} U_n = X$ covers the entire metric space. Thus, (U_n) is an open cover of C_1. Since C_1 is compact, there exists a finite subcover $\{U_{n_1}, U_{n_2}, ..., U_{n_M}\}$. Let U_{n_K} be the largest one. The ordering makes sense because $U_1 \subset U_2 \subset U_3 \subset ....$ Then, U_{n_K} covers C_1. However, $U_{n_K} := X \setminus C_{n_k}$, and $C_{n_k} \subset C_1$. Because C_{n_k} and C_1 are non-empty, there exists a point p in both C_{n_k} and C_1, which is not in U_{n_K}. This contradicts the statement that U_{n_K} covers C_1. Therefore, $\bigcap_{i=1}^{\infty} C_n \neq \emptyset$ □

This proof decisively relies on the fact that if the intersection is empty then we can construct some complement collection that covers the whole space (clever set theory manipulation) and the fact compact set induces finite subcover (this is sort of like black-magic. Very perplexing indeed).

A compact set is defined to be a set S such that *any* open cover $\{G_\alpha\}$ has a finite sub-cover $G_{\alpha_1}, ..., G_{\alpha_n}$.

This concept is very hard to wrap my head around. I have no good intuition or way to visualize this.

Trouble comes from a proof from Lecture 13: Compactness and the Heine-Borel of Harvey Mudd Lectures on Real Analysis.

https://www.youtube.com/watch?v=AQHVdiXRXQA

Prof. Francis Su proved that a bounded interval $[a, b] := \{x \in \mathbb{R} : a \leq x \leq\}$ in $\mathbb{R}$ is compact. The proof uses **nested interval theorem** (generalization is the **Cantor intersection theorem**) to eventually invoking the fact that if a point p is contain in an open set U then a very small neighborhood (bounded interval of radius ϵ) around p must be contained in U as well by the property of *open* set.

But instead of starting with interval $[a, b]$ and start shrinking down to smaller and smaller nested intervals $I_1 \supset I_2 \supset I_3 ...$ to a point p, what if we start with p and a small neighborhood around p, building up to the interval $[a, b]$.

We see that $\mathbb{N}$ is obviously not bounded and hence not compact. But if we can still take a small interval around each integer $(n - \epsilon, n + \epsilon)$. There're just infinitely many such neighborhoods and the cover cannot have finite subcover. Why couldn't we do the same thing with interval $[0, 1]$ for example. If we zoom in enough to $[0, 1]$, this interval looks exactly like the real line $\mathbb{R}$ because the real number line is very dense. i.e. The cardinality of open interval $(0, 1)$ and the real line is the same $\mathbb{R}$.

https://math.stackexchange.com/questions/200180/is-there-a-bijective-map-from-0-1-to-mathbbr

Couldn't we wrap around each point in $[0, 1]$ a small neighborhood like we wrap the natural numbers and conclude that a subset of $(0, 1)$ i.e. the image of the natural numbers on $(0, 1)$ is not compact so the superset cannot possibly be compact !

We interchangeably use $(0, 1)$ and $[0, 1]$ because the difference between closed and open set is very subtle and yet somehow dramatic.

See my GoodNotes on my iPad's analysis note.

This proof is kinda saying that in order for a big interval to have no finite cover, it must be the case that a very small closed interval around a point have no finite cover.

Trouble: Can it be the case that a small interval wrapped around a point is covered by a finite open set but a bigger interval is not? (Yes with the natural numbers?)

It kinda reminds me of infinite series. In order for an infinite series $\sum_{n=1}^{\infty} a_n$ to converge , the sequence $(a_n)_{n=1}^{\infty}$ converges to 0. This is a necessary condition but not sufficient. An example is the divergent harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$. The key is the difference in partial sum $\sum_{n=N}^{N+M} a_n$ must converge to zero . This is no guarantee that $\sum_{n=N}^{N+M} a_n$ converge to zero even if $\sum_{n=N}^{N+M} a_n \approx \sum_{n=N}^{N+M} \epsilon$. We know by the Archimedean property that no matter how small ϵ is, if we have enough of them, say $N + M$ of them, then the sum can be big enough and not converging to 0. This is related to a finite open set that covers a neighborhood around a point. Although one open set is finite, if we have enough point then $1 + 1 + 1 + 1 + ... = \infty$.

https://www.math.upenn.edu/ kazdan/508F14/Notes/compactness-L.pdf

Intuitive remark: a set is compact if it can be guarded by a finite number of arbitrarily nearsighted policemen.

https://math.stackexchange.com/questions/371928/what-should-be-the-intuition-when-working-with-compactness

Pedro Tamaroff's comment Look at $(0, 1)$. You can "stretch it out infinitely to the side" and it will look just look $\mathbb{R}$. In fact, you can "see" it is similar to $\mathbb{R}$ since it has no "visible end", i.e., you can zoom near 1 or 0 and you would keep on going indefinitely, never getting to 0 or 1. However, when you look at $[0, 1]$, you will "bump" into 0 or 1, and you'll see that this, in some sense, "ends". From the MO thread, "whenever one takes an infinite number of "steps" in the space, eventually one must get arbitrarily close to some other point of the space."

Christian Blatter's answer.

You may read various descriptions and consequences of compactness here. But be aware that compactness is a very subtle finiteness concept. The definitive codification of this concept is a fundamental achievement of 20th century mathematics.

On the intuitive level, a space is a large set X where some notion of nearness or neighborhood is established. A space X is compact, if you cannot slip away within X without being caught. To be a little more precise: Assume that for each point $x \in X$ a guard placed at x could survey a certain, maybe small, neighborhood of x. If X is compact then you can do with finitely many (suitably chosen) guards.

Complicated joke about compactness!

$$http : //spikedmath.com/505.html$$

A website of math jokes

$$https : //jcdverha.home.xs4all.nl/scijokes/1_6.html$$

Mathoverflow's question on compactness

https://mathoverflow.net/questions/25977/how-to-understand-the-concept-of-compact-space

Charles Matthews' answer
Some heuristic remarks are helpful only to a subset of readers. (Maybe that's true of all heuristics, as a meta-heuristic - if everyone accepts a rough explanation, it's something rather more than that.)

Non-compactness is about being able to "move off to infinity" in some way in a space. On the real line you can do that to the left, or right: but bend the line round to fill all but one point on a circle (which is compact) and you see the difference having the "other point" near which you end up. This example of real line versus circle is too simple, really. Another way you can "go off to infinity" in a space is by having paths branching out infinitely (as in König's lemma, which supplies another kind of intuition).

Compactness is a major topological concept because the various ways you might try to "trap" movement within a space to prevent "escape" to infinity can be summed up in a single idea (for metric spaces, let's say). The definition by open sets is cleaner, but the definition by sequences having to accumulate on themselves (not necessarily to converge, but to have at least one convergent subsequence) is somewhat quicker to say. If you restrict attention to spaces that are manifolds, you can think of continuous paths and whether they have to wind back close to themselves or not.

Terry Tao has a nice explanation in the Princeton companion to maths. The article's also on his website:

http://www.math.ucla.edu/ tao/preprints/compactness.pdf

In some metric spaces like the Euclidean $\mathbb{R}^n$, compact sets are just closed and bounded. Bounded sets are very easy to visualize but closed sets are much more subtle. The difference between closed set like $[0, 1]$ and open set like $(0, 1)$ seem to be dramatic. Refer to Pedro Tamaroff's comment.

Important stuff Heine-Borel theorem, Cantor intersection theorem, Bolzano–Weierstrass theorem, Finite intersection property.

Definition 6.0.1 (Finite intersection property)**.** *Let X be a set (think metric space) with a collection of subset $S = \{S_\alpha\}$ with S_α non-empty. Then S is said to have finite intersection property if any finite intersection of A is non-empty. That is $\cap_{i=1}^{N} S_{\alpha_i} \neq \emptyset$.*

Theorem 12. *If a collection of compact sets $S = \{S_\alpha\}$ in a metric space X has the finite intersection property, then the (possibly infinite) intersection over the entire collection has non-empty intersection i.e. $\cap_\alpha S_\alpha \neq \emptyset$.*

Proof. We proceed by a proof by contradiction. Let define the open complement of each compact set $U_\alpha := S_\alpha^c$. Assume that $\cap_\alpha S_\alpha = \emptyset$, then $\cup_\alpha U_\alpha = (\cap_\alpha S_\alpha)^c = (\emptyset)^c = X$ (DeMorgan's law) [1].

Pick an arbitrary set S' from S. $\cup_\alpha U_\alpha$ covers the whole metric space so it is an open cover of S'. Because S' is compact, there exists a finite sub-cover (U_{α_i}) of S'. We claim that $S' \cap (\cap_{i=1}^{N} S_{\alpha_i}) = \emptyset$, which is a contradiction of the hypothesis. To see why it is true, note that because (U_{α_i}) covers S' (i.e S' is a subset of $\cup_i U_{\alpha_i}$), $S' \cap (U_{\alpha_i})^c = S' \cap S_{\alpha_i} = \emptyset$. □

6.1 Cantor Set

Harvey Mudd

PBS Infinite Series

Define $K_0 = [0, 1]$. Then we take the middle piece out. Precisely, we define $K_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Continue recursively to define $K_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$. The recursive definition is

$$K_n = \frac{K_{n-1}}{3} \cup (\frac{K_{n-1}}{3} + \frac{2}{3})$$

. The cantor set is defined to be the points that are left after we have done this process infinitely many times.

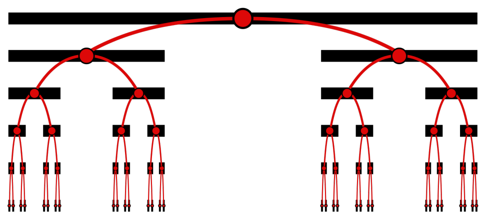
$$C := \cap_{i=1}^{\infty} K_i$$

Note that C is not empty. For example, 0 is always in C. More strongly, if a point p is an endpoint of a bounded interval in any of K_n's, then p is in C. For example, $\frac{1}{3}$ first becomes an endpoint of $[0, \frac{1}{3}]$ in K_1, and $\frac{1}{3}$ remains in $K_2, K_3, \ldots$.

One way to look at the Cantor set is through the lens of ternary base (base 3). In base 3, the number $\frac{1}{3} = 0.1$ because 0.1 in base three is $3^0 \times 0 + 3^{-1} \times 1$. Alternatively, $\frac{1}{3} = 0 \times \frac{1}{3} + 2 \times \frac{1}{9} + 2 \times \frac{1}{27} + \ldots = 0.0222\ldots = 0.0\overline{2}$. We can label each segment as 0 or 2. If the segment in the set K_n is label 0 that means that all numbers in that segment can have 0 in their n^{th} digit. For example, K_0 the segment $[0, 1]$ is labeled 0 in this picture because every point in this has the form 0.*something*, although 1 can be written as 1.00000... or $0.2222\ldots = 0.\overline{2}$.

[1] One way to think about this is to imagine each A_α as a house and U_α as a wagon directly outside the house. Every item in the world (metric space X) is either in the house or the wagon outside of it. If we claim that $\cup_\alpha U_\alpha$ does NOT contain everything in the world, then there exists at least an item p that is not in any of the wagon. But it means that item must be in EVERY house otherwise if it is not in house i^{th} then it is in the wagon i^{th}. So the intersection of the houses are not empty i.e. it contains p.

To reach $\frac{1}{3}$ in the picture, starting from K_1 we will continuously descend down to the right (going right corresponding to the digit 2).



We see that C contains all numbers that have ternary point (as opposed to decimal point) being 0 or 2. No number has ternary point of 1 because we always take the middle piece out at every step of constructing the next K_n.

Using Cantor diagonal argument, we can show that C has uncountably many elements. Suppose C has countably many elements, we will construct a counter-example p. Because C is assumed to have countably many elements, we can list them in a sequence like $x_1, x_2, x_3, \ldots$. Now, construct p by letting the n^{th} digit of p be 0 if n^{th} digit of x_n be 2, and the n^{th} digit of p be 2 otherwise. This way, we see that p differs from every single element x_n by at least the n^{th} digit so p is not included in the list. But p is included in C since it has only 0 or 2 in its digits and we arrive at a contradiction.

Chapter 7

Continuity

There're a lot of equivalent definitions of continuity. There are three main ones: epsilon-delta, convergent series, and open-set topology.

Definition 7.0.1 (Continuity: Epsilon-delta)**.** *$f : X \to Y$ is continuous at x_0 if for all $\epsilon > 0$, there exists $\delta(\epsilon)$ such that*

$$|f(x_0) - f(x)| < \epsilon$$

for all $x \in X$ that

$$|x - x_0| < \delta$$

The definition in terms of limit is just a more compact way to rephrase epsilon-delta definition

Definition 7.0.2 (Continuity: Limit definition)**.** *$f : X \to Y$ is continuous at x_0 if the left limit is equal to the right limit and the value of the function at x_0*

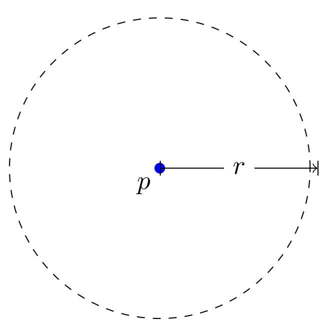
$$lim_{x \to x_0, x \in X} f(x) = lim_{x \to x_0^+, x \in X} f(x) = lim_{x \to x_0^-, x \in X} f(x) = f(x_0)$$

Definition 7.0.3 (Continuity: Convergent series)**.** *$f : X \to Y$ is continuous at x_0 if for all convergent series $(x_n)_{n=1}^{\infty} \to x_0$, where each $x_n \in X$. We have $(f(x_n))_{n=1}^{\infty} \to f(x_0)$.*

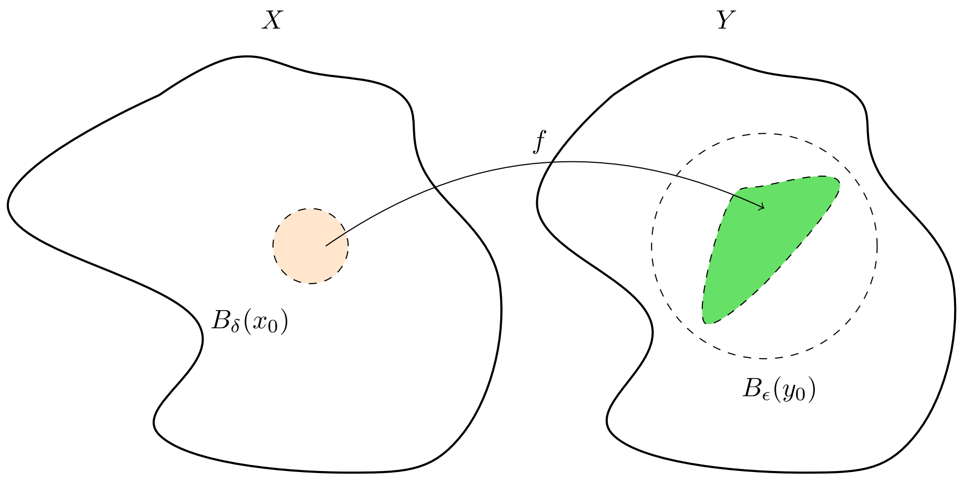
First, a point of notation. A open-ball of radius r around a point p in a metric space X is

$$B_r(p) = \{x \in X : d(x, p) < r\}$$

, where $d(x, p)$ is the distance between two points x and p defined by the metric of X.



Definition 7.0.4 (Continuity: open-ball)**.** *$f : X \to Y$ is continuous at x_0 if for all open-ball around y_0 in Y i.e. $B_\epsilon(y_0)$ for all $\epsilon > 0$, there exists an open-ball $B_\delta(x_0)$ in x_0 such that the image of $B_\delta(x_0)$ under function f is a subset of $B_\epsilon(y_0)$.*



We can extend the topological definition of continuity around one point to define continuity on the whole domain.

Definition 7.0.5 (Continuity: Open-set topology)**.** *$f : X \to Y$ is continuous on X if and only if for every open set U in Y, the inverse image (pre-image) in X*

$$f^{-1}[U] = \{x \in X | f(x) \in U\},$$

(note that f^{-1} is just notation. We do not require f to be injective) is also open.

An important property of continuous function. Continuous function $f : X \to Y$ maps a compact set $K \in X$ to a compact set $f[K] \in Y$. Note that compactness usually means "small" (next best thing to being finite). If K is small, then its image cannot be too big. This fits our idea of continuity at x_0. Take an "island of stability" around x_0, if this island is small enough then $f(x)$ cannot change too dramatically around $f(x_0)$ – that is why it is called island of **stability** anyway.

Theorem 13. *$f : X \to Y$ is continuous on X. If $K \subset X$ is a compact set, then $f[K] \subset Y$ is a compact set.*

Proof. Let $U = f[K]$, and $\{U_\alpha\}$ be an open cover of U. Then let $K_\alpha = f^{-1}[U_\alpha]$ which is open because f is continuous. Note that $\{K_\alpha\}$ will be an open cover of K. Because K is compact, there exists a finite subcover $\{K_{\alpha_i}\}_{i=1}^N$. Then we can note that $\{U_{\alpha_i}\}_{i=1}^N$ will cover U so U is compact. □

R As a corollary, we see that it is possible to map a small closed set to the whole real line. For example, $f(x) = \dfrac{1}{x}$ for $x \neq 0$ and $f(0) = 20$ can maps $X = [0, 1]$ to the upper half of the real line $Y = (0, \infty)$. But $f(x)$ is of course **not** continuous on this domain X. This is because $(0, \infty)$ is not compact. Being continuous somehow demands a fair scaling: if you want your codomain to be very big, you have to give me a big enough domain because continuous will prohibit magically approach an asymptote or jump.

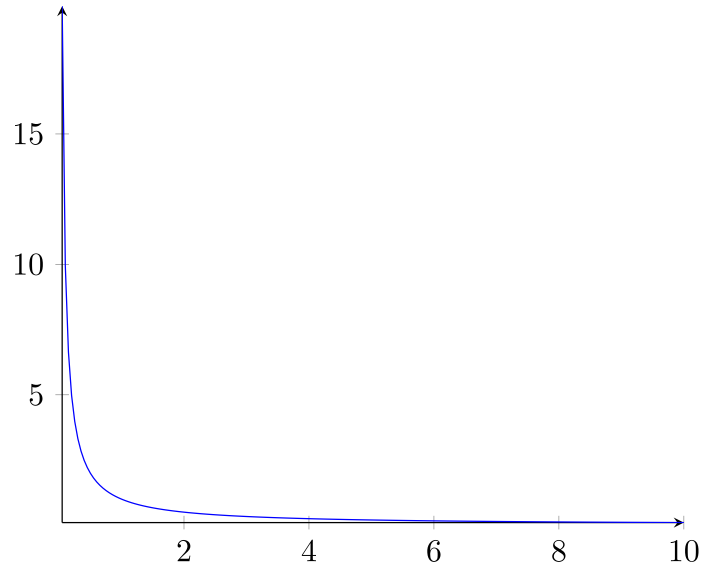
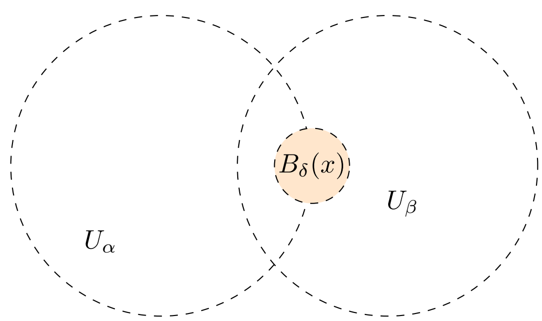


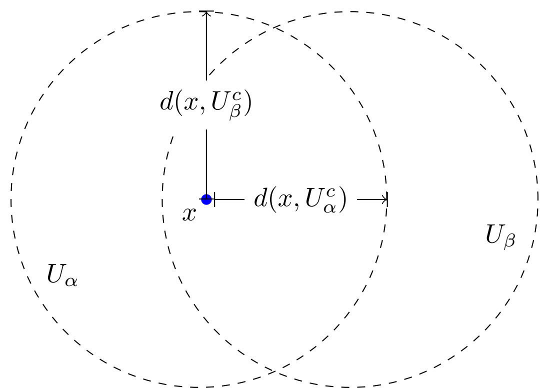
Figure 7.1: Function $\dfrac{1}{x}$

Theorem 14 (Lebesgue covering theorem)**.** *On a compact set K with an open cover $\{U_\alpha\}$, there exists a Lebesgue number $\delta > 0$ such that $B_\delta(x) \subset U_\alpha$ for all $x \in K$.*

Proof. The covering theorem basically says that it is possible to choose a small enough ball such that at every point $x \in K$, that ball lies *entirely* in one of the covering open set. In other words, it is impossible to have situation as follows.



This ball $B_\delta(x)$ half lies in U_α and half in U_β. This is bad. We want the ball to lie entirely inside of at least one open set. Now, focus on any particular point x. When we say $B_\delta(x)$ lies entirely in, say, U_α, it means that there is no point $p \in B_\delta(x)$ such that p is in the complement U_α^c. This can be achieved by setting δ to be smaller than the smallest distance between the center of the ball x and U_α^c. That way every point in the ball will be further away from U_α^c. The smallest distance between x and U_α^c, $d(x, U_\alpha^c)$, is defined to be $inf\{d(x,y)\}_{y \in U_\alpha^c}$. This infimum makes sense because metric distance is bounded above 0. Moreover, $d(x, U_\alpha^c) > 0$ whenever $x \in U_\alpha$ because U_α is open.



We can easily show by the triangle inequality that $d(x, U_\alpha^c)$ as a function of x (fixed U_α) is continuous. Now, because K is compact, there exists a finite subcover $(U_i)_{i=1}^N$. Define a new function

$$f(x) = max\{d(x, U_i^c)\}_{i=1}^N$$

$f(x)$ is a maximum function of finite continuous functions so $f(x)$ is also continuous. On a compact set K, continuous $f(x)$ must attained a minimum value δ. We see that this is our desired Lebesgue number δ. That is for every x, there exists U_i such that $d(x, U_i^c) \geq \delta$. This is straightforward because $\delta \leq f(x) = max\{d(x, U_i^c)\}_{i=1}^N$. □

Definition 7.0.6 (Tao's equivalent-sequences uniform continuity)**.** *For a function $f : X \to Y$, these statements are equivalent.*

1. *f is uniformly continuous.*
2. *If (a_n) and (b_n) (necessarily bounded or convergent) are equivalent with each $a_n, b_n \in X$, then $(f(a_n))$ and $(f(b_n))$ are equivalent.*

Proof. We re-iterate again that (a_n) and (b_n) are not necessarily convergent or even converges to an element in X. The two sequences are equivalent when they get arbitrary close to each other. In other words, for every $\epsilon > 0$, there exists N such that

$$|a_{N+k} - b_{N+k}| < \epsilon, \text{ for all } k \geq 0.$$

We will now prove that (1) $\implies$ (2). Given two equivalent (a_n) and (b_n), we will show that $(f(a_n))$ and $(f(b_n))$ are also equivalent. That is for every ϵ, there exists N such that

$$|f(a_{N+k}) - f(b_{N+k})| < \epsilon \text{ (a) , for all } k \geq 0.$$

We will now rely on uniform continuity to pick N indirectly through $\delta(N)$. Now because (a_n) and (b_n) are equivalent, we can find N such that

$$|a_{N+k} - b_{N+k}| < \delta$$

This N also guarantees the inequality (a) by uniform continuity.

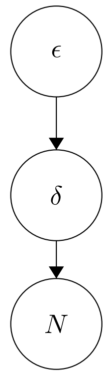


Figure 7.2: The choice of ϵ relies on δ. And δ can be used to pick N.

We now prove a somewhat more difficult implication (2) $\implies$ (1). $\square$

Theorem 15 (Uniform continuity on compact set)**.** *$f : X \to Y$ is continuous on the compact set $K \subset X$, then f is also uniformly continuous on K.*

Proof. We have 3 different but very similar proofs. One of them by Harvey Mudd uses Lebesgue covering lemma, one by Tao uses two equivalent sequences, and the last one by me just using raw definitions of uniform convergence.

Proof by me. For the sake of contradiction, suppose that f is not uniformly continuous. That means there exists an ϵ for which none of positive δ works for every x of K i.e. $inf\{\delta_x(\epsilon)\}_{x\in K} = 0$. Fix that ϵ. More precisely, for every $\delta > 0$, there exists $x \in K$ that for some $y \in B_\delta(x)$, we have $|f(x) - f(y)| \geq \epsilon$. Now, by axiom of choice we form the sequence $(x_n)_{n=1}^\infty$, where $x_n \in K$ such that for some $y \in B_{\frac{1}{n}}(x)$ we have $|f(x) - f(y)| \geq \epsilon$. In other words, our sequence has roughly $(\delta_{x_n})_{n=1}^\infty \approx \left(\frac{1}{n}\right)_{n=1}^\infty \to 0$.

Now, because $(x_n) \in K$, which is bounded, by the Bolzano–Weierstrass theorem we have a subsequence $(x_{n_i}) \to L$. This number L is in K because K is bounded. We will try to arrive at some contradiction. Because f is continuous on K, namely at L, there exists δ_L such that for every $x \in B_{\delta_L}(L)$ we have $|f(L) - f(x)| < \frac{\epsilon}{2}$. Because $(x_{n_i}) \to L$, there exists $N(\delta_L)$ such that $|x_{n_{N+k}} - L| < \frac{\delta_L}{2}$ for all $k \geq 0$. By the triangle inequality, we have

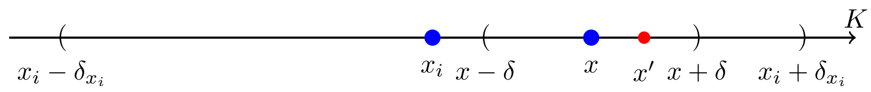
$$|f(x_{n_{N+k}}) - f(x)| \leq |f(x_{n_{N+k}}) - f(x_0)| + |f(x_0) - f(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

for all $x \in B_{\delta_L}(L)$. Know clearly, $\frac{\delta_L}{2}$ works for all $x_{n_{N+k}}$. But this contradicts the fact that we construct (x_n) so that $\delta_{x_{n_{N+k}}} < \dfrac{1}{N+k}$. Therefore, f must be uniformly continuous.

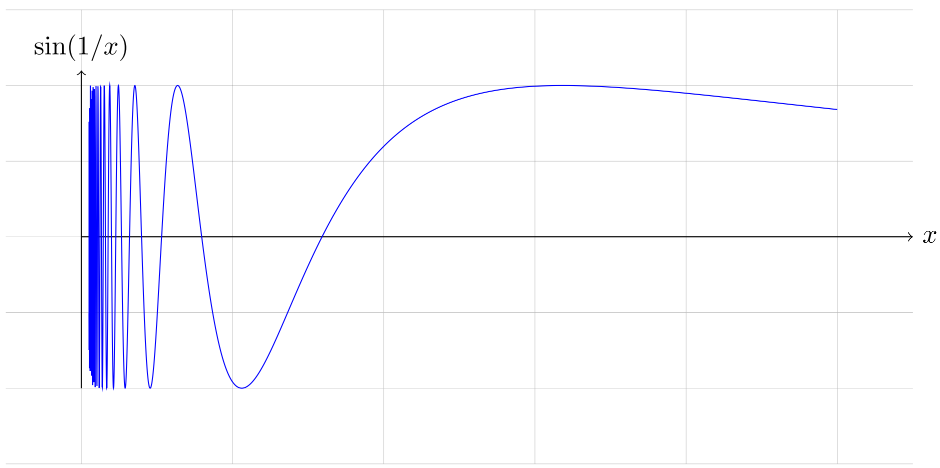
Now, the proof by Lebesgue covering lemma is very similar. First, because f is continuous on K, for every $\frac{\epsilon}{2}$, we can find $\delta_x(\epsilon)$ for every x in K. Fix $\frac{\epsilon}{2}$, and form the ball $B_{\delta_x}(x)$ for each x. We easily see that this is an open cover of K. By Lebesgue covering lemma, there exists Lebesgue number δ. We will show that this δ is our uniform island of stability. For every x, the ball $B_\delta(x)$ with Lesbegue number radius falls entirely in a ball $B_{\delta_{x_i}}(x_i)$. By the triangle inequality,

$$|f(x_i) - f(x')| \leq |f(x_i) - f(x)| + |f(x) - f(x')| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

for all $x' \in B_\delta(x)$.



□



Discontinuous function. Dirichlet's function is

$$f(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \\ 0, & \text{if } x \in \mathbb{I} \end{cases}$$

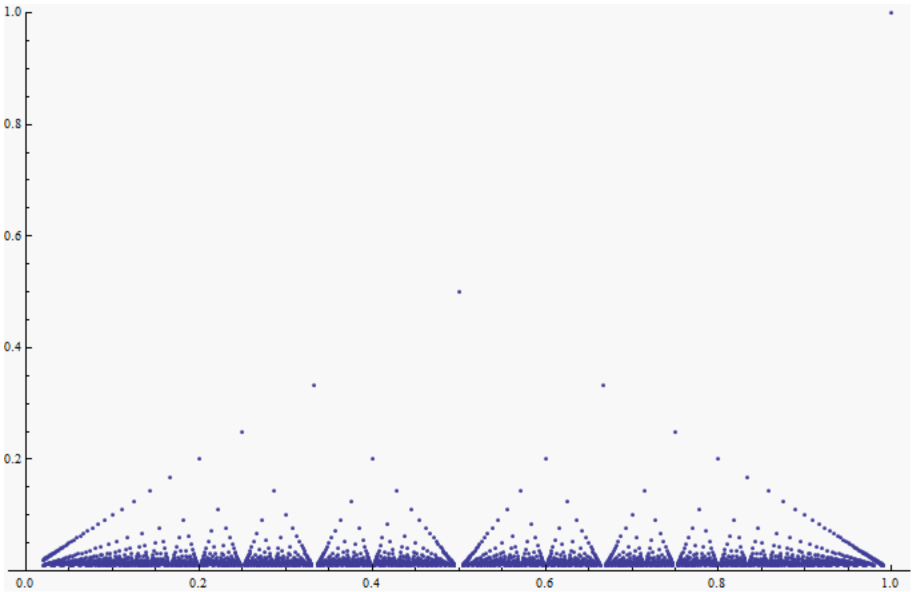
This function is discontinuous everywhere due to the denseness of rational $\mathbb{Q}$ and irrational $\mathbb{I}$ in $\mathbb{R}$ [1]. Whenever we take a small ball of radius ϵ around a point x, the ball $B_\epsilon(x)$ will contains both rational and irrational numbers.

Small-Remann is a modification to Dirichlet's function.

$$f(x) = \begin{cases} \frac{1}{q}, & \text{if } x \in \mathbb{Q} \text{ and } x \text{ is of the reduced form } \frac{p}{q}. \\ 0, & \text{if } x \in \mathbb{I} \end{cases}$$

This small-Riemann function is continuous at irrational point, where we can find $\frac{p}{q}$ such that $q \to \infty$ and $\frac{1}{q} \to 0$. However, it is discontinuous at every rational point due to the denseness of rational $\mathbb{Q}$ and irrational $\mathbb{I}$ in $\mathbb{R}$.

[1] between any two real numbers exist a rational and an irrational number



Note that if we set $y = q$ for some $q \in \mathbb{N}$ then the function goes above q finitely many times. For example, take $q = 5$ then the fun

7.1 Limit at infinity

We have a good idea of when some sequence is converging to a number $L \in \mathbb{R}$ and $L \notin \{\infty, -\infty\}$, however what it means for a sequence to converge to ∞?

Simply enough, it means that this sequence (a_n) grows without bound i.e. there does not exists $M \in \mathbb{R}$ such that $M > |a_n|$ for all n. This motivates our definition of an adherent point being infinite.

Definition 7.1.1 (Infinity as adherent point)**.** *Recall that a point p is an adherent point to the set K if in a sense the set K can get arbitrarily close to p. We use the same idea here. When the set has infinity as its adherent point, K is unbounded. Precisely, if for every $M \in \mathbb{R}$, there exists $x \in K$ such that $x > M$ then we have that $+\infty$ is adherent to K. Similarly, if for every $M \in \mathbb{R}$, there exists $x \in K$ such that $x < M$ then we have that $-\infty$ is adherent to K*

Note that topological structure can unify this definition with the usual ϵ-ball definition of finite adherent point.

Definition 7.1.2 (Limit at infinity)**.** *$f : X \to \mathbb{R}$ is a function. We say that $f(x)$ converges to a finite number L at infinity and write $lim_{x\to+\infty,x\in X} f(x) = L$ iff for every $\epsilon > 0$ there exists M such that f is ϵ-close to L for every $x > M$.*

ote that this definition is very similar to the definition of limit as x approaches a finite number x_0. For finite x_0, we take an ϵ-ball around $f(x_0)$ and consider island of stability around x_0 or how far we need to get close to x_0 to guarantee the stability of f around $f(x_0)$. The above definition is also very similar in this regards. It cares about how close to x_0 we need to get in order to guarantee ϵ-closeness to f. Because $x_0 = +\infty$, we can get close to $+\infty$ not by a both-sided ball but by approach from the left-side, increasing unboundedly.

Example 4:
Let $f : (0, +\infty) \to \mathbb{R}$ be $f(x) = 1/x$. Then we have $lim_{x\to\infty,x\in(0,+\infty)} 1/x = 0$

We need to show that for every ϵ there exists $M(\epsilon)$ such that

$$|f(x) - 0| = \frac{1}{x} < \epsilon$$

whenever

$$x > M$$

. Take M to be

$$M > \frac{1}{\epsilon}$$

will satisfy this.

Chapter 8

Differentiation

NOTE that differentiability and continuity are local properties i.e. it only depends on what happens to the function around an arbitrarily small neighborhood around a point in the domain. Ironically enough, these properties depend on the domain. For example, $1/x$ is continuous on $(0, \infty)$ but not $[0, \infty)$.

Definition 8.0.1 (Continuity)**.** *$f : X \to \mathbb{R}$ is continuous at $x_0 \in X$ if*

$$lim_{x \to x_0, x \in X} f(x) = f(x_0)$$

Definition 8.0.2 (Differentiability)**.** *$f : X \to \mathbb{R}$ is differentiable at $x_0 \in X$ and x_0 is a* ***limit point*** *of X if*

$$lim_{x \to x_0, x \in X - \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0}$$

If we let $f(x_0) = c$ be a constant and $g(x) = \dfrac{f(x) - f(x_0)}{x - x_0}$ is a function of $g(x)$. So we have $f'(x_0) = lim_{x \to x_0, x \in X - \{x_0\}} g(x)$. We see that differentiability is pretty much the continuity property of $g(x)$ at an outside point x_0 on a slightly different domain $X - \{x_0\}$ except the limit does not have to equal to $g(x_0)$ which is undefined $\left(\dfrac{0}{0}\right)$.

C^k function. $x^{5/4} sin(\frac{1}{x})$ is a C^0 function but not C^1 although it derivative exists. ($sin(1/x)$ is the topologist's sine curve).

Weierstrass function: continuous everywhere but differentiable nowhere Weierstrass function:.

Theorem 16 (Mean value theorem)**.** *If $f : \mathbb{R} \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that*

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

Think of generalized mean value theorem as distance and velocity. $(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c)$

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}$$

Imagine we let $f(t)$ be the distance of the rabbit from a starting point p at time t and $g(t)$ be the distance of a turtle. f and g defined on the same time interval $[a, b]$ so they have the same time to move. For simplicity, assume they started a same point. Positive $f(t)$ means that the rabbit is at the right-side of p and negative means in the right. $f'(t)$ is the speed. Think of $f'(c)$ is the average speed.

Cauchy's mean value theorem (generalized) Rolle's theorem Taylor approximation Derivative of monotone functions. L' Hopital rule.

Chapter 9

Riemann integration

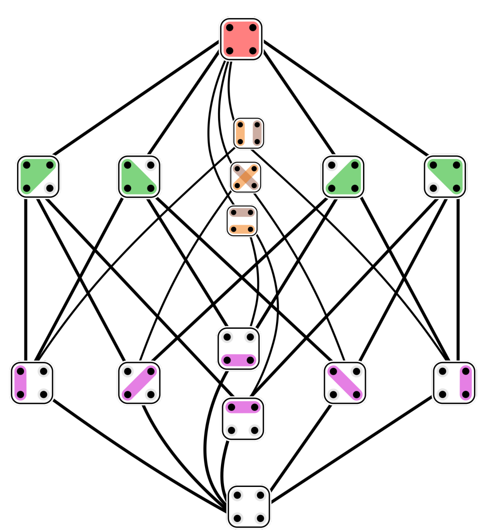
IDEA: use piecewise constant functions to approximate arbitrary functions.

9.1 Partition

A partition $S = \{S_1, ..., S_n\}$ of a set X is a finite set of S_i such that every $x \in X$ belongs to exactly one S_i. In other words, S_i's are all disjoint in the sense that $S_i \cap S_j = \emptyset$ for $i \neq j$, and their union forms the whole X i.e. $X = \bigcup_{i=1}^{N} S_i$.

Refinement of partition: Partition S is a refinement of partition S' if every $S_i \in S'$ is a subset of some $S'_j \in S'$. In that case, we write $S \leq S'$, because we can verify that this definition satisfies properties of partial order.

This picture is the partial order of partition refinement on a 4-element set $\{1, 2, 3, 4\}$. A color region indicates a subset of X while uncolored dots indicate single-element subsets (singletons). The red partition is simple $\{1, 2, 3, 4\}$. The first green partition (from the left) is $\{\{1, 2, 3\}, \{4\}\}$. The first purple partition is $\{\{1, 3\}, \{2\}, \{4\}\}$. We can verify that indeed $\{\{1, 3\}, \{2\}, \{4\}\} \leq \{\{1, 2, 3\}, \{4\}\}$. Specifically, $\{1, 3\} \subset \{1, 2, 3\}$, $\{2\} \subset \{1, 2, 3\}$,$\{4\} \subset \{4\}$.



One way to think about this lattice is to start from the top (red partition). Then, we go into a color region and start throw element out to form singleton. That way, we start fragmenting or mashing "coarser" partition into finer pieces. For example, the first purple is obtained by mashing the green group of the first green partition by throwing 2 out. $\{\{1,2,3\},\{4\}\} \to \{\{1,3\},\{2\},\{4\}\}$. That first purple can also obtained by mashing the brown-group of the first (from top to bottom) orange-brown partition.

Definition 9.1.1 (Common refinement)**.** *The common refinement of partitions P and P' is defined as*

$$P\#P' = \{M \cap N : (M,N) \in P \times P'\}$$

e can prove that $P\#P'$ be the partition of X. $P\#P'$ is finer than both P and P'. $P\#P'$ is also in some sense the closest child to P and P' in the lattice. If we want to produce a partition P'' that is finer than both P and P' than P'' is the finer than $P\#P'$. That is, we can fragment $P\#P'$ to get P''.

9.2 Piecewise constant functions

Definition 9.2.1 (Constant function)**.** *Let $f : X \to \mathbb{R}$. We say that f is constant on X iff there exists $c \in \mathbb{R}$ such that $f(x) = c$.*

Definition 9.2.2 (Piecewise constant function)**.** *Let I be a bounded interval $[a,b]$ let $f : I \to \mathbb{R}$ and P be a partition of I. We say that f is piecewise constant with respect to P if for every $J \in P$, f is constant on J.*

We say that f is piecewise constant if there exists a partition P of I such that f is piecewise constant with respect to P.

Lemma 17. *If $f : I \to \mathbb{R}$ is piecewise constant with respect to P, then f is also piecewise constant with respect to P', which P' as a partition of I finer than P.*

The space of piecewise constant functions is closed under algebraic operations

Lemma 18. *Let I be a bounded interval and $f : I \to \mathbb{R}$, $g : I \to \mathbb{R}$. Then these functions are all piecewise constant*

1. *$f + g$.*
2. *$f - g$.*
3. *$max(f,g)(x) := max(f(x), g(x))$for each x.*
4. *$min(f,g)$.*
5. *f/g if $g(x) \neq 0$ for every x.*

As an illustration, let consider the function $f + g$. Let P and P' be the partitions of I such that f is piecewise constant with respect to P and g with respect to P'. We claim that $f+g$ is piecewise constant with respect to the common refinement $P\#P'$. We already know that $P\#P'$ is a partition of I so it is sufficient to show that f is constant for every $J \in P\#P'$. Indeed, every $J = M \cap N$ for some M in P and N in P'. f is constant in M with the value, say, m and constant in N with the value n. Therefore, f is constant in $M \cap N$ with the value $m + n$.

Definition 9.2.3 (Piecewise constant integral)**.** *$f : I \to \mathbb{R}$ is piecewise constant with respect to the partition P. We define the piece constant integral over P as follows*

$$\int_P f := \sum_{J \in P} f(x \in J)|J|$$

where $|J|$ is the measure (length) of interval J. The measure of $[a,b], (a,b], [a,b), (a,b)$ by definition is $b - a$.

ote that piecewise constant integral is a well-defined sum. We require the partition P to be finite, so the sum will be finite.

We can also show that piecewise constant integral is independent of partitions.

Lemma 19. *$f : I \to \mathbb{R}$ is piecewise constant with partitions P and P' of I, we have*

$$\int_P f = \int_{P'} f$$

We can show they are both equal to $\displaystyle\int_{P\#P'} f$. An every important trick about partition that we can recall from probability is that

$$p(y) = \textstyle\sum_{x\in X} p(x, y)$$

In our set language, let P be a partition of P' and M be a subset of I. Then we have this disjoint additive property

$$M = \textstyle\bigcup_{J\in P}(M \cap J)$$

The proof is straightforward, $M \subset I$ so every point $x \in M$ must be in the partition P, which is the cover of I. Moreover, x must be in exactly one of the J because P being a partition is comprised of disjoint intervals. Now, let consider

$$\int_{P\#P'} f = \sum_{(M,N)\in P\times P'} |M \cap N| f(x \in M \cap N).$$

Let $c_{(m,n)}$ be $f(x \in M \cap N)$ Because this is a finite sum, we can re-arrange terms and make this into a double sum as we wish

$$\int_{P\#P'} f = \sum_{M\in P}\sum_{N\in P'} |M \cap N| c_{(m,n)} = \sum_{M\in P} (c_{(m,n)} \sum_{N\in P'} |M \cap N|)$$

By the disjoint additive property,

$$\int_{P\#P'} f = \sum_{M\in P} c_{(m,n)}|M| = \int_P f$$

A different grouping to create a double sum $\sum_{N\in P'}\sum_{M\in P} |M \cap N| c_{(m,n)}$ also shows that $\int_{P\#P'} f = \int_{P'} f$

Theorem 20 (Law of integrations)**.** *Let I be a bounded interval and let $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ be piecewise constant on I then*

1. $\displaystyle\int_I (f + g) = \int_I f + \int_I g.$
2. $\displaystyle\int_I cf = c\int_I f$ *for some constant $c \in \mathbb{R}$.*
3. $\displaystyle\int_I (f - g) = \int_I f - \int_I g.$
4. *If $f(x) \geq 0$ for all $x \in I$, then* $\displaystyle\int_I f \geq 0.$
5. *If $f(x) \geq g(x)$ for all $x \in I$ then* $\displaystyle\int_I f \geq \int_I g.$
6. *If $f(x) = c$ for all $x \in I$ then* $\displaystyle\int_I f = c|I|.$

7. *Let J be a bounded interval containing I i.e. $I \subset J$, and define a new function $F : J \to \mathbb{R}$ as*

$$F(x) = \begin{cases} f(x), \text{ if } x \in I. \\ 0, \text{ if } x \notin I \end{cases}$$

8. *Suppose that $\{J, K\}$ partition I into two intervals J and K. Then f is piecewise constant on J and K. Furthermore, we have*

$$\int_I f = \int_J f + \int_K f$$

9.3 Upper and lower Riemann integrals

This is similar to how we define Jordan measure. We "squeeze" the Jordan-measurable region by a outer elementary and inner elementary regions.

This time, we will define Riemann integrability by defining the upper and lower Riemann integrals. But first, let just quickly define some simple terms about one function "majorizing" or "minorizing" another function

Definition 9.3.1 (Majorization of function)**.** *Let $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$. We say f majorizes g on I if $f(x) \geq g(x)$ for all $x \in I$. Similarly, if $f(x) \leq g(x)$ for all $x \in I$ then we say f minorizes g on I.*

Definition 9.3.2 (Upper and lower Riemann integrals)**.** *Let $f : I \to \mathbb{R}$, f is not necessarily piecewise constant. We define the upper Riemann integral of f as follows*

$$\overline{\int_I} f := inf\{\int_I g \text{ for piecewise constant function } g \text{ on } I \text{ that majorizes } f \text{ }\}.$$

Similarly, we define the lower Riemann integral of f as follows

$$\underline{\int_I} f := sup\{\int_I g \text{ for piecewise constant function } g \text{ on } I \text{ that minorizes } f \text{ }\}$$

We also show that the upper and lower Riemann integrals are finite (not $+\infty$ or $-\infty$) if f is bounded.

Lemma 21. *$f : I \to \mathbb{R}$ is bounded by $M \geq 0$ on I i.e. $-M \leq f(x) \leq M$ for all $x \in I$. Then*

$$-M|I| \leq \underline{\int_I} f \leq \overline{\int_I} f \leq M|I|.$$

Proof. We will first show that $-M|I| \leq \underline{\int_I} f$. Let define $g(x) = -M$ for all $x \in I$. By boundedness of f, we know that $f(x) \geq g(x)$ or the piecewise constant g minorizes f. Therefore, $-M|I| = \int_I g \leq \underline{\int_I} f$. Otherwise, $\underline{\int_I} f$ would not be the infimum, the **greatest** lower bound. Similarly, consider a piecewise function h that majorizes f and k that minorizes f. Easily, h also majorizes k. Therefore, by law of integrations 5, we have that $\int_I h \geq \int_I k$. Therefore, $\int_I h \geq \underline{\int_I} f$ or otherwise h would be the least upper bound of the infinitely uncountable set of real numbers $K = \{\int_I g$ for piecewise constant function g on I that minorizes f $\}$ [1]. Similarly, we can take the infimum on h to arrive at the conclusion that $\underline{\int_I} f \leq \overline{\int_I} f$. □

[1] note that h does not have to belong K i.e. h does not have to minorizes f in order to be the least upper-bound of K. It all boils down to real numbers once Riemann integrals of functions are evaluated.

Definition 9.3.3 (Riemann integrability)**.** *Let* $f : I \to \mathbb{R}$ *be bounded function on bounded interval* I*. If* $\underline{\int_I} f = \overline{\int_I} f$*, then we say that* f *is Riemann integrable and write*

$$\int_I f := \underline{\int_I} f = \overline{\int_I} f$$