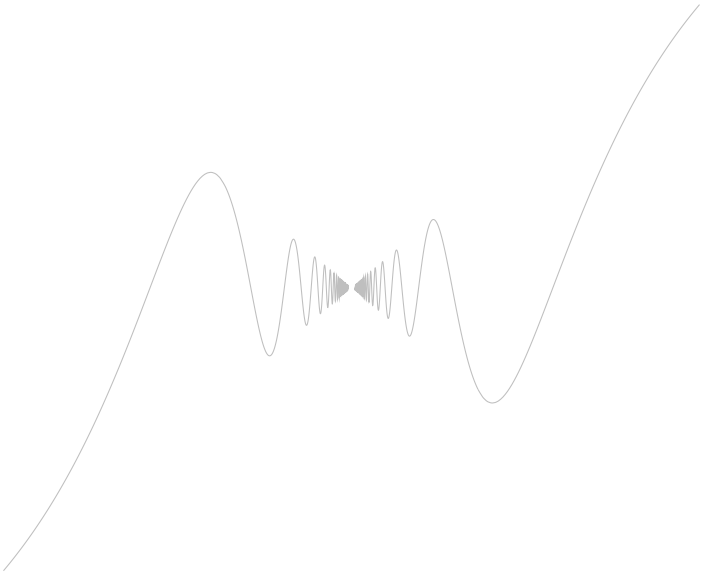
# ARITHMETIC

*of the*

# CONTINUUM



# Contents

# PREFACE

These notes form an introduction to doing mathematics.

Einstein once said to

> "make things as easy as possible, but no easier."

I have done my best to follow this advice.

Justin T. Chun
January 2025

# Chapter 0

# ZEROTH ORDER LOGIC

Logical reasoning forms the cornerstone of mathematical thought.

We thus start by summarizing the basic notions.

## FUNCTIONS AND TUPLES

**Definition 0.1.** A **function** is a way to map elements from one set, the *domain*, to elements of another set, the *codomain*, such that every element of the domain is paired with a unique element of the codomain.

We codify logical connectives (e.g. *and*, *or*, *not*) as functions taking values in the set

$$\mathcal{V} = \{\bot, \top\}$$

where $\bot$ denotes *false* and $\top$ denotes *true*.

We collect together finitely many objects using $n$-tuples,[1] where for example

$$(a, b, c, d, a)$$

denotes a 5-tuple containing the entries $a$, $b$, $c$, $d$, and $a$.

We denote by $\mathcal{V}^n$ the set of all $n$-tuples with entries in $\mathcal{V}$, and by

$$f : \mathcal{V}^n \to \mathcal{V}$$

an $n$-ary truth function.

[1] Note that order matters for $n$-tuples but does not matter for sets. Also, $n$-tuples may have repeats whereas sets may not.

### NOTE ON FINITENESS FOR THE CAREFUL

Observe that $\mathcal{V}$ has two elements, $\mathcal{V}^2$ has four elements, and in general $\mathcal{V}^n$ has $2^n$ elements. A function going from $\mathcal{V}^n$ to $\mathcal{V}$ must make $2^n$ binary decisions (i.e. whether to set each function value to either $\top$ or $\bot$), so there are $2^{2^n}$ possible functions that could go from $\mathcal{V}^n$ to $\mathcal{V}$. We point this out as a means of saying that everything we have so far defined is *finite*.

We will soon work with arbitrary (i.e. potentially infinite) sets, though. Here's why we did not just start with infinite sets: intuitively, one thinks of a set as an unordered collection of objects with no repeats. However, this naive conception can lead to logical disaster.

For instance, we have the following argument, known as *Russell's Paradox*:

> Consider the set of all sets that are not elements of themselves – call this set $\Omega$. On the one hand, if $\Omega$ is an element of itself, then it is (by definition of $\Omega$) not an element of itself. On the other hand, if $\Omega$ is not an element of itself, then it belongs with all the other sets that aren't elements of themselves – namely, in $\Omega$. So $\Omega$ contains itself if and only if $\Omega$ does not contain itself – a contradiction.

We would like to exclude sets like $\Omega$ from all of mathematics. The working solution is called ZFC, a collection of axioms and axiom schemata that specify how sets ought to behave. An exploration of ZFC at this very moment would be a distraction.

Thus, we restrict our attention to finite sets – where our intuition is trustworthy. The nonnegative integers (as we shall call them) are essentially a part of the language layer, as in people know of them regardless of whether they care to learn mathematics. Even nonhuman animals have been observed to count.

### PROPOSITIONS ARE TRUE XOR FALSE

We must also rely on intuition for content on the notions of true and false, which mathematicians call *truth values* and accept as primitive.

**Definition 0.2.** A **proposition** is a gramatically correct declarative sentence that can be assigned exactly one truth value.

We often use the letters P, Q, R and S to denote propositions.

The following sections develop the nullary, unary, and binary truth functions.

## 0.0 NULLARY TRUTH FUNCTIONS

The set $\mathcal{V}^0 = \{()\}$ consists of exactly one 0-tuple (the only possible 0-tuple).

Thus we have two nullary truth functions:

$$\bot : \mathcal{V}^0 \to \mathcal{V} \text{ given by } () \mapsto \bot, \qquad \text{and} \qquad \top : \mathcal{V}^0 \to \mathcal{V} \text{ given by } () \mapsto \top.$$

That is, *the nullary truth functions are simply the truth values.*

## 0.1 UNARY TRUTH FUNCTIONS

The set $\mathcal{V}^1$ consists of two 1-tuples (corresponding to the two elements of $\mathcal{V}$).

To each 1-tuple there are two choices of output,

spawning a total of four unary truth functions.

(0) $\bot(P) = \bot$ (1) $\top(P) = \top$ (2) $\mathrm{id}(P) = P$ (3) $-(P) = -P$

The functions (0) and (1) are exactly the nullary truth functions from before.

We can describe functions (2) and (3) in tabular form:

| P | $-$P |
|---|---|
| $\bot$ | $\top$ |
| $\top$ | $\bot$ |

In (2), "id" stands for "identity function." This is standard notation.

The function (3) is called **negation**. It is a consequence of the table above that

$$P = -(-P),$$

a fact known as *double negation*.

## 0.2 BINARY TRUTH FUNCTIONS

The set $\mathcal{V}^2$ consists of four 2-tuples:

$$(\bot,\bot), \quad (\bot,\top), \quad (\top,\bot), \quad (\top,\top).$$

Denoting an element of $\mathcal{V}^2$ by $(P,Q)$, sixteen binary truth functions follow.[2]

(0) $\bot(P,Q) = \bot$
(1) $\top(P,Q) = \top$
(2) $=(P,Q) = (P = Q)$
(3) $\neq(P,Q) = (P \neq Q)$
(4) $P(P,Q) = P$
(5) $(-P)(P,Q) = -P$
(6) $Q(P,Q) = Q$
(7) $(-Q)(P,Q) = -Q$
(8) $\leq(P,Q) = (P \leq Q)$
(9) $\geq(P,Q) = (P \geq Q)$
(10) $<(P,Q) = (P < Q)$
(11) $>(P,Q) = (P > Q)$
(12) $\wedge(P,Q) = (P \wedge Q)$
(13) $\vee(P,Q) = (P \vee Q)$
(14) $\uparrow(P,Q) = (P \uparrow Q)$
(15) $\downarrow(P,Q) = (P \downarrow Q)$

We give their descriptions all at once:

| P | Q | $\bot$ | $\top$ | $P = Q$ | $P \neq Q$ | P | $-P$ | Q | $-Q$ |
|---|---|---|---|---|---|---|---|---|---|
| $\bot$ | $\bot$ | $\bot$ | $\top$ | $\top$ | $\bot$ | $\bot$ | $\top$ | $\bot$ | $\top$ |
| $\bot$ | $\top$ | $\bot$ | $\top$ | $\bot$ | $\top$ | $\bot$ | $\top$ | $\top$ | $\bot$ |
| $\top$ | $\bot$ | $\bot$ | $\top$ | $\bot$ | $\top$ | $\top$ | $\bot$ | $\bot$ | $\top$ |
| $\top$ | $\top$ | $\bot$ | $\top$ | $\top$ | $\bot$ | $\top$ | $\bot$ | $\top$ | $\bot$ |

| P | Q | $P \leq Q$ | $P \geq Q$ | $P < Q$ | $P > Q$ | $P \wedge Q$ | $P \vee Q$ | $P \uparrow Q$ | $P \downarrow Q$ |
|---|---|---|---|---|---|---|---|---|---|
| $\bot$ | $\bot$ | $\top$ | $\top$ | $\bot$ | $\bot$ | $\bot$ | $\bot$ | $\top$ | $\top$ |
| $\bot$ | $\top$ | $\top$ | $\bot$ | $\top$ | $\bot$ | $\bot$ | $\top$ | $\top$ | $\bot$ |
| $\top$ | $\bot$ | $\bot$ | $\top$ | $\bot$ | $\top$ | $\bot$ | $\top$ | $\top$ | $\bot$ |
| $\top$ | $\top$ | $\top$ | $\top$ | $\bot$ | $\bot$ | $\top$ | $\top$ | $\bot$ | $\bot$ |

Some comments on these functions follow.

- The functions (0) and (1) are exactly the nullary truth functions from before.
- Functions (2) **biconditional** and (3) **exclusive disjunction** have to do with whether the inputs agree:

$$(P \neq Q) = -(P = Q)$$

- Functions (4 - 7) are constructed by restricting to a single input and then applying a unary truth function.

[2] Note that while it is convenient to use prefix notation to define these truth functions, in practice one uses infix notation.

- Functions (8) **implication**, (9) **reverse implication**, (10) **negated reverse implication**, and (11) **negated implication** are all *asymmetric* and *transitive* binary truth functions.
- Functions (12) **conjunction** aka "and" and (13) **inclusive disjunction** aka "or (possibly both)" are related via *de Morgan's laws*:

  $$(-P) \vee (-Q) = -(P \wedge Q), \qquad (-P) \wedge (-Q) = -(P \vee Q)$$

  Furthermore, the following *distributive laws* hold:

  $$P \wedge (Q \vee R) = (P \wedge Q) \vee (P \wedge R), \qquad P \vee (Q \wedge R) = (P \vee Q) \wedge (P \vee R)$$

  We also have:

  $$(P \leq Q) = (-P \vee Q).$$

- Functions (14) **negated conjunction** aka "nand" and (15) **negated inclusive disjunction** aka "nor" are interesting because each individually can generate all the other binary truth functions:

  $$-P = P \uparrow P, \qquad P \wedge Q = -(P \uparrow Q), \qquad \text{etc.}$$

## 0.3 MORE ON CONDITIONAL STATEMENTS

In the proposition $P \leq Q$, we call P the **antecedent** and Q the **consequent**. The antecedent suffices for the consequent, whereas the consequent necessitates the antecedent. Amplifying,

> "P is a sufficient condition for Q" equals "Q is a necessary condition for P."

Be prepared to recognize the following forms of $P \leq Q$:

- "P implies Q" or "Q is implied by P"
- "if P then Q" or "Q if P"
- "only if Q then P" or "P only if Q"

The conditional $P \leq Q$ is the **converse** to $Q \leq P$ and the **contrapositive** of $-Q \leq -P$. The converse of the converse of a conditional is the original conditional, and the contrapositive of the contrapositive of a conditional is the original conditional.

## 0.4 LOGICAL DEDUCTION

This refers to the process of starting with a set of assumptions and arriving at a conclusion after a finite number of steps.

We care about truth functions and tabular proof because it is a quick way to get to the truth value of any proposition; we care about logical deduction because it is a microcosm not dissimilar to how mathematics actually functions.

Formally, a propositional calculus can be thought of as a set of propositions, a set of logical connectives, a set of axioms, and a set of laws of inference. Here is a common starting point:

(1) $\vdash (P \leq (Q \leq P))$

(2) $\vdash ((P \leq (Q \leq R)) \leq ((P \leq Q) \leq (P \leq R)))$

(3) $\vdash ((-P \leq -Q) \leq (Q \leq P))$

(MP) $P, (P \leq Q) \vdash Q$

The first three lines are *axioms* where P, Q, etc. can be any propositions. The remaining line is **modus ponens**, a *law of inference*. The $\vdash$ symbol is called a *turnstyle*, and it should be read as "given what's left of $\vdash$, we have what's right of $\vdash$."

The first move is to turn those axioms into inferences via (MP):

(1i) $P \vdash (Q \leq P)$

(2i) $(P \leq (Q \leq R)) \vdash ((P \leq Q) \leq (P \leq R))$

(3i) $(-P \leq -Q) \vdash (Q \leq P)$

We can actually further transform (3i) so that it becomes **modus tollens**:

(MT) $(-P \leq -Q), Q \vdash P.$

Now suppose we wanted to demonstrate that

(D1) $(P \leq Q), (P \leq (Q \leq R)) \vdash (P \leq R).$

We start by assuming both $P \leq Q$ and $P \leq (Q \leq R)$. We may then infer via (2i) that $(P \leq Q) \leq (P \leq R)$. Finally we get $P \leq R$ via (MP).

This is enough to demonstrate transitivity of implication, i.e. that

(TR) $(P \leq Q), (Q \leq R) \vdash (P \leq R).$

Easy – start by assuming both $P \leq Q$ and $Q \leq R$. Then by (1i), we get $P \leq (Q \leq R)$. Finally we get $P \leq R$ via (D1).

## 0.5 METHODS OF PROOF

There are roughly three ways one can approach a proof. We will use as our working example the fact that every integer is either even or odd. We can structure this in conditional form as follows:

$$n \text{ is an integer} \leq n \text{ is either even or odd.}$$

The simplest (but often not easiest) way to prove Q given P is to show $P \leq Q$ and then use modus ponens:

> *If* $n$ *is an integer, we may apply division with remainder, which states that for integers* $a$ *and* $b$ *there exists a unique integer quotient* $q$ *and remainder* $r$ *such that* $a = bq + r$ *and* $0 \leq r < b$. *The only two possiblilities for* $r$ *in this case are 0 and 1; hence,* $n$ *is either even or odd.*

This is called **direct proof**.

Another way to prove Q given P is to show $-Q \leq -P$ and then apply modus tollens:

> *Suppose* $n$ *is neither even nor odd. Then* $n+1$ *is neither even nor odd, so* $n(n+1)$ *is not necessarily divisible by 2. Since for every integer* $m$ *we must have* $2 \mid m(m+1)$, $n$ *is not necessarily an integer, i.e. it is not the case that* $n$ *must be an integer.*

This is known as **proof by contrapositive**.

A third way to prove Q given P is to show that if one had P but also $-Q$, then disaster ensues:

> *Suppose* $n$ *is neither even nor odd. Then neither* $n$ *nor* $n+1$ *is divisible by 2, so* $n(n+1)$ *is not divisible by 2. This contradicts the fact that every product of consecutive integers is divisible by 2.*
>
> *If* $nm$ *is a product of consecutive integers, then* $nm$ *is divisible by 2?*
>
> *Indeed: suppose* $nm$ *is a product of consecutive integers not divisible by 2. Then* $nm$ *is odd, and this entails* $n$ *and* $m$ *are odd. But then odd numbers are at least two apart, contradicting our assumption that* $n$ *and* $m$ *were consecutive.*

This is called **proof by contradiction**.

We will encounter all three methods of proof throughout our study of math.

## 0.6 USE OF THE $\leq$ SYMBOL

Readers with previous logic experience may be wondering whatever happened to the $\rightarrow$ or $\implies$ symbol. We now take a moment to explain.

**Definition 0.3.** A **bounded lattice** is a 6-tuple $(X, \wedge, \vee, \top, \bot, \leq)$

where $X$ is a set, $\vee$ (the *join* or **supremum**) and $\wedge$ (the *meet* or **infimum**)

are functions from $X^2$ to $X$, and $\top$ and $\bot$ are special elements of $X$.

These objects satisfy the following conditions:

$$x \wedge (y \wedge z) = (x \wedge y) \wedge z \quad \text{and} \quad x \vee (y \vee z) = (x \vee y) \vee z$$

$$x \wedge y = y \wedge x \quad \text{and} \quad x \vee y = y \vee x$$

$$x \wedge (x \vee y) = x \quad \text{and} \quad x \vee (x \wedge y) = x.$$

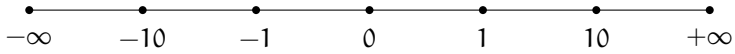
$$x \wedge \top = x \quad \text{and} \quad x \vee \bot = x.$$

We may further define $x \leq y$ to mean $x = x \wedge y$ (or equivalently $y = x \vee y$).

A bounded lattice with a notion of complement $-x$ is called a **boolean lattice**.

The logic we are introducing can be understood in terms of boolean lattices. This is why we use $X \leq Y$ instead of the more traditional $X \rightarrow Y$.

This can be visualized. Consider the number line $\mathbf{R}$, and add two endpoints, $-\infty$ placed to the left of all negative reals and $+\infty$ placed to the right of all positive reals:

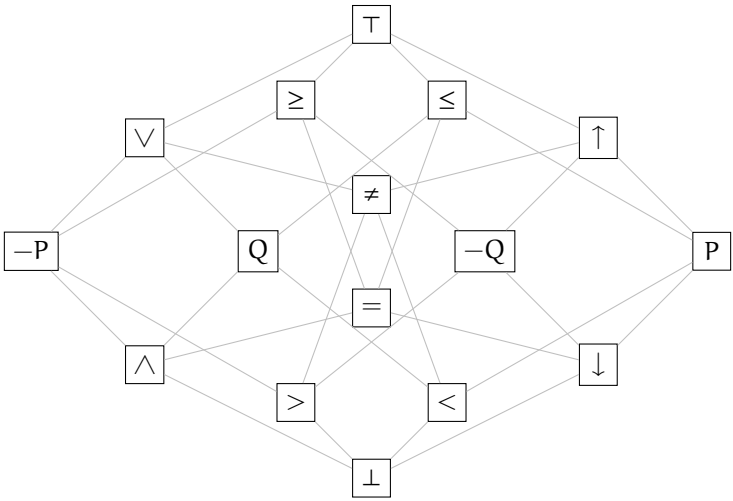



we thus form $\overline{\mathbf{R}}$, the extended real number line. Thus we can think of

- $+\infty$ as a $\top$ and $-\infty$ as a $\bot$,
- $\wedge$ as taking the minimum of two extended reals, $\vee$ as taking the maximum,
- and negation as rotating the number line 180 degrees about 0.

In other words, $\overline{\mathbf{R}}$ is *also* a bounded lattice, with $\mathcal{V}$ a boolean sublattice. (Note, however, that "$-$" does not logically complement elements in $\overline{\mathbf{R}}$.)

Another boolean lattice is the $n$-ary truth functions. Here we show $n = 2$.




## 0.7 EXERCISES AND REMARKS

1. Show that all sixteen binary truth functions can be built from $\uparrow$.
2. Deduce from the first two axioms and modus ponens that $\vdash (P \leq P)$.
3. Convince yourself that the $n$-ary operations $\mathcal{V}^n \to \mathcal{V}$ form a boolean lattice.

One of the surprising things about propositional calculus is it only needs one axiom to get off the ground:

$$(((((A \leq B) \leq (-C \leq -D)) \leq C) \leq E) \leq ((E \leq A) \leq (D \leq A)))$$

This is known as *Meredith's sole axiom*, and from it one can derive the axioms given in the section on logical deduction.

To learn more about (and get lots of practice with) logical deduction, I recommend the *Metamath Proof Explorer* webpage. There, one can flip through a large portion of zeroth and first order logic. I also recommend Tim Button's Cambridge version of *forallx* (by P.D. Magnus) as a general course in logic, as well as Peter Smith's *An Introduction to Formal Logic*.

# Chapter 1

# FIRST ORDER LOGIC

In this chapter we aim to develop a *formal language* atop which we may define a theory of sets. We will need a countably infinite set of objects $\Omega$.

**Definition 1.1.** A set is **countably infinite** if its elements can be sequenced.

For example, the number of primes is countably infinite, as are the number of nonnegative integers $\mathbf{Z}_{\geq 0}$, integers $\mathbf{Z}$, rationals $\mathbf{Q}$, and even algebraic numbers $\overline{\mathbf{Q}}$. See the exercises at the end of Chapter 3 for explicit bijections.

This is a fairly weak notion of infinity: for example, there are more real numbers in between 0 and 1 than there are elements in any countably infinite set. That is to say, $\mathbf{R}$ and certain subsets of it are in fact *uncountable*.

## WHAT IS $\Omega$?

One can think of $\Omega$ as a universe of propositions. We only need countably many propositions for now; by the time we start thinking about uncountably many objects at once, we will be well into set theory.

Still, this is the last time we will be careful in these notes, with a small exception regarding the Axiom of Countable Choice, about whether things are finite or even the small kind of infinite. For example, turn to Chapter 3 and you'll find an explanation of how a power set of a set is strictly larger than the original set, yet in Chapter 2 we take as an axiom that power sets always exist. The degree of infinity we thus require of mathematics is immense.

## 1.0 TERMS AND PREDICATES

The "words" in our formal language are called *terms*.

By a **term** we mean either:

- a **constant** $c$, i.e. some fixed object in $\Omega$, or
- a **variable** $x$, i.e. any object in $\Omega$, or
- a **function** $f : \Omega^n \to \Omega$.

We note that constants are nullary functions, and that the argument of a function could be another function. So we really should write $\Omega = \Omega_0$ for the universe of propositions, and then $\Omega_n$ for all the $n$-ary functions from $\Omega_0$ to $\Omega_0$. Keep in mind that variables can only refer to objects in $\Omega_0$ in our formal language.

By an $n$-ary **predicate** we mean a function $P : \Omega^n \to \mathcal{V}$. That is, an $n$-ary predicate takes in $n$ propositions and outputs a truth value. Predicates are the building blocks of formulas, which we discuss next.

## 1.1 FORMULAS AND QUANTIFIERS

The "sentences" in our formal language are called *formulas*.

By a **formula** we mean either:

- an $n$-ary predicate applied to an $n$-tuple of terms, or
- a negated formula, or two formulas joined by a binary truth function, or
- $\forall x P(x)$ or $\exists x P(x)$ where $x$ is a variable and $P(x)$ is a formula.

Here are some examples of formulas:

$$c_1, \quad \top, \quad P(c_1, c_2) \wedge Q, \quad \forall x R(x), \quad \exists x S(x) \vee P(c_3, c_4)$$

The symbols $\forall$ "for all" and $\exists$ "there exists" are called **quantifiers**.

The statement $\forall x P(x)$ may be read

"$P(x)$ is true for all $x$,"

whereas the statement $\exists x P(x)$ may be read

"there exists an $x$ for which $P(x)$ is true."

Many logical malapropisms can be attributed to accidentally swapping quantifiers.

## 1.2 NEGATION OF QUANTIFIERS

Note that

$$-(\forall x P(x)) = \exists x(-P)(x) \qquad \text{and} \qquad -(\exists y Q(y)) = \forall y(-Q)(y)$$

The above is basically de Morgan's laws:

think of $\forall$ as behaving like $\wedge$ and $\exists$ as behaving like $\vee$.

## 1.3 AN IMPORTANT EXAMPLE: $\varepsilon$-$\delta$

Here is an example where the quantifiers really matter. It looks forward a bit, but the example is so important that we preview it here. Reader, it is Cauchy's $\varepsilon$-$\delta$ definition of a limit. Remember it and it will serve you well.

**Definition 1.2.** Let $f : \mathbf{R} \to \mathbf{R}$.

We say that the **limit** of $f$ as $x$ approaches $c$ is $L$ if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \text{such that} \quad (|x - c| < \delta) \;\leq\; (|f(x) - L| < \varepsilon)$$

and write $\lim_{x \to c} f(x) = L$.

This can be formalized in terms of games: if Player A gives Player B an $\varepsilon_0$, Player B can then respond with a $\delta_0$, and if Player A then sets $\varepsilon_1 = \delta_0$ and then proposes $\varepsilon_1$ to Player B, then Player B produces a corresponding $\delta_1$ for Player A, and so on.

## 1.4 THE NON-LOGICAL $\in$ SYMBOL

We need one more symbol:[1]

$$\in$$

which denotes elementhood, i.e. $x \in X$ denotes the proposition

"$x$ is an element of $X$."

In our mathematical ontology we shall have three types: propositions, sets, and proper classes. Formally we will only have propositions and sets.

[1] The $\in$ symbol was orginally an $\epsilon$ (lower case Greek epsilon), but eventually became its own glyph.

### BITING THE APPLE

In the background will soon appear the class of all sets, an object so vastly infinite that Cantor himself regarded it as divine. With the adoption of sets into our ontology, we thus augment the universe of propositions.

For example, we can have a proposition $P_\alpha$ for each $\alpha \in \mathbf{R}$ that declares whether $\alpha$ is rational. So we have at least uncountably many propositions, and it is not hard to show that one can in fact do much worse.

If you worry about welcoming an entire zoo of infinities, you might do well to read some set theory. See the references at the end of chapter for more.

## 1.5 PROPER CLASSES

It is convenient to at least have a name for the class of all sets. We thus write

$$\mathsf{Set} = \{x : x = x\}$$

which, when taken alongside $\Omega$, we may call the

*classical universe of mathematical discouse*.

Informally we'll just call it "Set." It is a **proper class**, i.e. a class that is not a set.

The defintion is a reference to the three classical laws of thought:

- Excluded middle: *every proposition is either true or false.*
- Noncontradiction: *no proposition is both true and false.*
- Identity: *each object is equal to itself.*

The first two laws of thought are part of the defintion of proposition.

## 1.6 THE VOID

If we can characterize anything with the assumption $x = x$, we ought to be able to characterize nothing at all with the assumption $x \neq x$. We call

$$\varnothing = \{x : x \neq x\}$$

the **empty set**.

The empty set is sometimes an element of another set,

but no set (including the empty set) is an element of the empty set:

$$(x \in \varnothing) = \bot$$

There is exactly one function on $\varnothing$, the *empty function* which sends nothing nowhere.

## 1.7 THE NONNEGATIVE INTEGERS

Let's get ahead of ourselves for a moment.

The *binary union*

$$x \cup y$$

of two sets $x$ and $y$ consists of all $z$ such that either $z \in x$ or $z \in y$.

Define[2]

$$\begin{aligned}
0 &= \varnothing \\
1 &= 0 \cup \{0\} = \{0\} = \{\varnothing\} \\
2 &= 1 \cup \{1\} = \{0, 1\} = \{\varnothing, \{\varnothing\}\} \\
3 &= 2 \cup \{2\} = \{0, 1, 2\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} \\
4 &= 3 \cup \{3\} = \{0, 1, 2, 3\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}\} \\
&\vdots
\end{aligned}$$

Continuing this process ad infinitum, one obtains the **nonnegative integers**

$$\mathbf{Z}_{\geq 0} = \{0,\ 1,\ 2,\ 3,\ 4,\ \ldots\}$$

as well as the **successor function**

$$(\cdot)^{++} : \mathbf{Z}_{\geq 0} \to \mathbf{Z}_{\geq 0}$$

given by $n^{++} = n \cup \{n\}$. Thus $1^{++} = 2$, $2^{++} = 3$, and so on.

**Definition 1.3.** A **sequence** taking values in a set $X$ is a function $f : \mathbf{Z}_{\geq 0} \to X$.

[2] This next part is due to John von Neumann.

## 1.8 INDUCTION AND WELL-ORDERING

We also have the following schema (one axiom per proposition):

**Axiom 1.4** (Induction schema)**.** *If* $P(n)$ *satisfies*

$$P(0) \qquad \textit{and} \qquad \forall n \in \mathbf{Z}_{\geq 0}(P(n) \leq P(n^{++})),$$

*then* $\forall n \in \mathbf{Z}_{\geq 0}\, P(n)$.

In plain English, this is saying that if one has a property that holds at zero, and one can show that the property holding at $n$ implies the property holding at $n^{++}$, then the property holds for every nonnegative integer. We may restructure the axiom so that it is about sets instead of propositions, which is what we really want:

**Axiom 1.5** (Set Induction)**.** *If a subset* $X \subseteq \mathbf{Z}_{\geq 0}$ *satisfies* $0 \in X$ *and*

$$(n \in X) \;\leq\; (n^{++} \in X)$$

*for all* $n \in \mathbf{Z}_{\geq 0}$*, then* $X = \mathbf{Z}_{\geq 0}$.

These are logically equivalent to the following:

**Axiom 1.6** (Well-Ordering)**.** *Every nonempty subset* $X \subset Z_{\geq 0}$ *has a least element.*

Proof. Let $X$ be a subset of $\mathbf{Z}_{\geq 0}$ satisfying $0 \in X$ and

$$(n \in X) \;\leq\; (n^{++} \in X).$$

We proceed by contradiction: suppose $X \neq \mathbf{Z}_{\geq 0}$. Then there is a nonnegative integer not in $X$, i.e. $\mathbf{Z}_{\geq 0} \setminus X$ is nonempty. Then $\mathbf{Z}_{\geq 0} \setminus X$ has a least element $n$. Note that $n \neq 0$, since $0 \in X$. Thus, $n > 0$, and since $n$ is the least element not in $X$, $n-1$ must be in $X$. But by assumption, $(n-1)^{++} = n \in X$, contradicting our assumption that $n \notin X$. Thus proves that the well-ordering principle implies the principle of induction.

Conversely, consider a nonempty subset $Y \subset \mathbf{Z}_{\geq 0}$. If $Y$ has just one element, then that element is the least element of $Y$. Now suppose the well ordering principle is true for all subsets of $\mathbf{Z}_{\geq 0}$ with $n$ elements, and suppose $Y$ has $n^{++}$ elements. Take $y \in Y$ and let $z$ be the least element of $Y \setminus y$. Then $\min(\{y, z\})$ is the least element of $Y$. This proves that the principle of induction implies the well-ordering princple. ∎

## 1.9 EXERCISES AND REMARKS

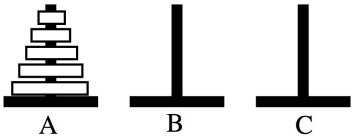
1. Is the class FinSet of all finite sets proper?

2. An element of a set is always another set. Can you see a potential hazard that we need to avoid? Hint: $x_0 \ni x_1 \ni x_2 \ni x_3 \ni \cdots$.

3. **Induction on trominos.** A *tromino* is a $2 \times 2$ square of $1 \times 1$ squares, with one corner square missing:

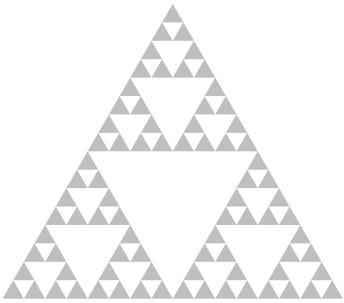
   Prove, via induction, that every $2^n \times 2^n$ grid of squares with exactly one square missing can be tiled with trominos.

4. **Towers of Hanoi.** Consider the following puzzle:




   The objective is to move the rings from one peg to another, with the constraint that a larger ring cannot be placed on top of a smaller ring.

   Convince yourself that the graph of states forms a partial Sierpinski triangle:



   It should now be immediate how to solve the puzzle. You could try to set this up as an induction proof; it would involve defining a couple of things.

5. (Beware: kinda tricky.)

   Let $F_n$ denote the $n$th Fibonacci number, where the Fibonacci numbers are defined so that $F_0 = F_1 = 1$, and $F_{n+2} = F_{n+1} + F_n$. Show that

   $$F_{n+1}^2 + F_n^2 = F_{2n+1}$$

   using induction.

I learned the tromino problem from Irena Swanson, and the Towers of Hanoi problem from Knuth's *Concrete Mathematics*.

I found the Fibonacci number problem in Larson's *Problem Solving Through Problems*. The book includes a solution, should you desire to view it.

# Chapter 2

# THE ZFC AXIOMS

It is now time to define sets. We deliver the axioms in unapologetic symbols, having prepared the reader with the previous two chapters.

The following treatment was inspired by a similar one in the book by Thomas Jech.

## 2.0 SETS ARE DEFINED BY THEIR ELEMENTS

**Axiom 2.1** (Extensionality)**.**

$$(x = y) = \forall z(z \in x = z \in y).$$

That is, two sets are equal if and only if they have the same elements.

## 2.1 UNORDERED PAIRS EXIST

**Axiom 2.2** (Pairing)**.**

$$\forall a, b \exists x \forall c(x \in c \leq x = a \vee x = b).$$

The pair defined above is unique by extensionality,

so we may now write $\{a, b\}$ for any sets $a$ and $b$.

In the case $a = b$, we may write $\{a\}$, the singleton containing just $a$.

Note that $a \neq \{a\}$.

Note also that $\{a, b\} = \{b, a\}$.

We define **ordered pairs** as follows:

$$(x, y) = \{\{x\}, \{x, y\}\}$$

## 2.2 POWER SETS EXIST

**Axiom 2.3** (Power set)**.**

$$\forall x \exists p \forall y ((y \in p) = \forall z ((z \in y) \leq (z \in x))).$$

For any set $x$, define the **power set** of $x$ to be the set $\mathcal{P}(x)$ of all subsets of $x$.

The **cardinality** of a set is basically the size of a set.

The power set of a set is strictly larger than the set itself.

In particular, this implies that $\mathcal{P}(\mathbf{Z}_{\geq 0})$ is uncountable.

**Theorem 2.4.** *Let $X$ be a set. Then $|\mathcal{P}(X)| > |X|$.*

Proof. We will show that any map $f : X \to \mathcal{P}(X)$ cannot be surjective. Suppose otherwise, and let $Y = \{x \in X : x \notin f(x)\}$. Then there exists $\xi \in X$ such that $f(\xi) = Y$. But by construction, $\xi \in Y$ if and only if $\xi \notin f(\xi) = Y$. This is a contradiction, so $f$ cannot be surjective.

On the other hand, $g : X \to \mathcal{P}(X)$ given by $g(x) = \{x\}$ is injective, so

$$|\mathcal{P}(X)| > |X|$$

as desired. ∎

## 2.3 ARBITRARY UNIONS EXIST

**Axiom 2.5** (Union)**.**

$$\forall x \exists y \forall z ((z \in y) = \exists w ((w \in x) \wedge (z \in w))$$

We denote $y$ by $\bigcup x$. In the case of $x = \{x_1, \ldots, x_N\}$ we write

$$\bigcup x = \bigcup_{n=1}^{N} x_n = x_1 \cup \cdots \cup x_N.$$

## 2.4 THE NONNEGATIVE INTEGERS EXIST

**Axiom 2.6** (Induction)**.**

$$\exists s(\varnothing \in s \wedge (\forall x \in s)(x^{++} \in s)).$$

This is actually the only thing that guarantees that any sets exist at all.

That is, at least one set exists because $\mathbf{Z}_{\geq 0}$ exists.[1]

## 2.5 ANY DEFINABLE SUBCLASS OF A SET IS A SET

**Axiom 2.7** (Separation schema)**.** *For* $\phi(z, p)$ *a formula, we have*

$$\forall x \forall p \exists y \forall z((z \in y) = (z \in x \wedge \phi(z, p))).$$

This basically says that the intersection of a set with a class forms another set.

From this, we get the existence of the empty set, and also the existence of intersections and complements.

We also get subsets, but in a specific way. Basically, one can only create a subset from an existing set. This is diffrent from the *comprehension schema* which falsely states that one can take subsets of any class.

The binary cartesian product can be written as subset of known sets:

$$x \times y \subset \mathcal{P}(\mathcal{P}(x \cup y))$$

A subset $r$ of $x \times y$ is called a **relation**.[2]

[1] The set $\mathbf{Z}_{\geq 0}$ is sometimes denoted $\mathbf{N}$ (for "natural numbers"). Some think the "natural numbers" ought to start at 1, others at 0. We avoid this ambiguity by calling this set the nonnegative integers.

The set $\mathbf{Z}_{\geq 0}$ is also sometimes denoted $\omega$, for "first infinite ordinal."

[2] For a relation $R$ we use $xRy$ to mean $(x, y) \in R$.

## 2.6 THE IMAGE OF A SET IS A SET

**Axiom 2.8** (Replacement schema)**.** *For* $\phi(x,y,p)$ *a formula, we have*

$$\forall x \forall y \forall z(\phi(x,y,p) \wedge \phi(x,z,p) \leq (y = z)) \leq \forall z \exists w \forall y(y \in w = (\exists x \in z)\phi(x,y,p)).$$

This says that the image of function whose domain is a set must be a set.

## 2.7 ALL SETS ARE WELL FOUNDED

**Axiom 2.9** (Regularity)**.**

$$\forall s((s \neq \varnothing) \leq (\exists x \in s)(s \cap x \neq \varnothing))$$

There are no infinite membership chains and no cyclic membership loops.

Every chain has an $\in$-minimal element.

## 2.8 THE AXIOM OF CHOICE

**Axiom 2.10** (Choice)**.** *Every family of nonempty sets has a choice function.*

**Definition 2.11.** A surjection

$$\pi : X \to Y$$

is said to **split** if there is some $\sigma : Y \to X$ such that $\pi \circ \sigma = \mathrm{id}_Y$.

Here is a weaker form of the axiom of choice:

---

There are three major kinds of relations:

- **Equivalence relations** are relations that are reflexive, symmetric, and transitive.
- **Partial orderings** are relations that are reflexive, antisymmetric, and transitive.
- **Functions** are possibly the most important relations. For a function $f : X \to Y$, every element $x \in X$ has some element $y \in Y$ such that $(x,y) \in f$, and $((x,a) = (x,b)) \leq (a = b)$.

**Axiom 2.12** (Countable choice)**.** *Let* $X$ *be a set,*

$$\pi : X \twoheadrightarrow \mathbf{Z}_{\geq 0}$$

*a surjection. Then* $\pi$ *splits.*

**Theorem 2.13.** *A countable union of countable sets is countable.*

## 2.9 EXERCISES AND REMARKS

1.

See *Set Theory* by Thomas Jech.

# Chapter 3

# RELATION AND NUMBER

This chapter begins to develop the actual mathematics.

## 3.0 HOMOGENEOUS BINARY RELATIONS

A relation $R \subseteq X \times Y$ is *homogeneous* if $X = Y$.

trivial universal identity

Three examples of homogeneous relations $x_1 R x_2$ are:

- The *trivial relation* $R = \varnothing$, which holds for no $(x_1, x_2) \in X^2$;
- The *universal relation* $R = X^2$, which holds for any $(x_1, x_2) \in X^2$;
- The *identity relation* $R = \mathrm{id}$, which holds when $x_1 = x_2$.

A homogeneous relation $R$ is:

- *reflexive* when $\mathrm{id} \subseteq R$
- *transitive* when $(xRy \wedge yRz) \leq xRz$ for all $x, y, z \in X$.

## 3.1 PARTIALLY ORDERED SETS

A homogeneous relation that is both reflexive and transitive is called a *preorder*.

- A preorder is *symmetric* if $xRy = yRx$;

  a symmetric preorder is also known as an **equivalence relation**.

- A preorder is *antisymmetric* if $((xRy \wedge yRx) \leq (x = y))$;

  an antisymmetric preorder is also known as a **partial order**.

**Definition 3.1.** A set equipped with a partial order is called a **poset**.

Posets are the simplest environment in which one can reason about order.

**Definition 3.2.** Let $(X, \preceq_X)$ be a poset, $S \subseteq X$.

We say $M \in X$ is an *upper bound* of $S$ (and that $S$ is *bounded above*)

if $s \preceq_X M$ for every $s \in S$.

We say $m \in X$ is a *lower bound* of $S$ (and that $S$ is *bounded below*)

if $m \preceq_X s$ for every $s \in S$.

If $S$ is both bounded above and bounded below, then we say $S$ is *bounded*.

We say $M \in X$ is the *meet* or *least upper bound* or **supremum** of $S$ if

- $M$ is an upper bound of $S$ and
- if for any upper bound $M'$ of $S$, we have $M \preceq_X M'$.

We say $m \in X$ is the *join* or *greatest lower bound* or **infimum** of $S$ if

- $m$ is a lower bound of $S$ and
- if for any lower bound $m'$ of $S$ we have $m' \preceq_X m$.

We denote the meet/supremum by $\bigvee S$ or $\sup S$

and the join/infimum by $\bigwedge S$ or $\inf S$.

**Definition 3.3.** A *lattice* is a poset stable under finite meets and joins.

A *complete lattice* is simply a poset stable under arbitary meets and joins.

## 3.2 FUNCTIONS AND INVERTIBILITY

Now we shift focus to relations $R \subseteq X \times Y$ where $X$ does not necessarily equal $Y$.

**Definition 3.4.** Let $R \subseteq X \times Y$ be a relation. We say $R$ is

- *left-total* if for all $x \in X$ there is some $y \in Y$ such that $xRy$
- *right-total* if for all $y \in Y$ there is some $x \in X$ such that $xRy$
- *left-unique* if for all $y \in Y$, $(xRy \wedge x'Ry) \leq (x = x')$
- *right-unique* if for all $x \in X$, $(xRy \wedge xRy') \leq (y = y')$

We then say that $R$ is

- a *partial function* if $R$ is right-unique but not necessarily left-total,
- a *multivalued function* if $R$ is left-total but not necessarily right-unique,
- a **function** if $R$ is both left-total and right-unique.[1]

Of course, this definition of function agrees with the one previously given.

We can also use the above four criteria to define when a function can be undone.

**Definition 3.5.** If $f$ is a function, then we say that $f$ is

- **injective** if $f$ is also left-unique,
- **surjective** if $f$ is also right-total,
- **bijective** or **invertible** if $f$ is also both left-unique and right-total.

Injective functions are also called injections, surjective functions are also called surjections, bijective functions are also called bijections.

[1] For emphasis, we sometimes say the function is *well-defined*.

**Definition 3.6.** Let $(X, \preceq_X)$ and $(Y, \preceq_Y)$ be partially ordered sets. A function

$$f : X \to Y$$

is **monotone** if for all $x, y \in X$ we have

$$(x \preceq_X y) \implies (f(x) \preceq_Y f(y)),$$

and **antitone** if for all $x, y \in X$ we have

$$(x \preceq_X y) \implies (f(y) \preceq_Y f(x)).$$

## 3.3 IMAGE AND PREIMAGE

One can send subsets forwards and backwards through functions.

**Definition 3.7.** Let $f : X \to Y$ be a function. There are functions

$$f_* : \mathcal{P}(X) \to \mathcal{P}(Y) \qquad \text{and} \qquad f^* : \mathcal{P}(Y) \to \mathcal{P}(X)$$

called **image** and **preimage**, respectively given by

$$f_*(A) = \{y \in Y : \exists x \in A \text{ such that } f(x) = y\} \quad \text{and} \quad f^*(B) = \{x \in X : f(x) \in B\}.$$

Note that image respects the order of composition whereas preimage reverses the order of composition:

$$(g \circ f)_* = g_* \circ f_* \qquad \text{and} \qquad (g \circ f)^* = f^* \circ g^*.$$

In slightly more ornate language, we say image is *covariant* and preimage is *contravariant*. Note that since the preimage of the codomain is the domain, we have $f_*(f^*(B)) = f_*(A)$. Furthermore,

**Proposition 3.8.** *For $f : A \to B$, we have $f^*(f_*(A)) = A$.*

Proof. Let $a \in A$. Then since $f$ is a function, there is a unique $b \in f_*(A)$ such that $f(a) = b$. But then $a \in f^*(f_*(A))$ by definition of preimage. Conversely, suppose $a' \in f^*(f_*(A))$. Then $f(a') = b'$ for some $b' \in f_*(A)$ by definition of preimage, hence $a' \in A$ since $f$ is a function. ∎

## 3.4 THE SCHRÖDER-BERNSTEIN THEOREM

Here we prove that two opposing injections between two sets is enough to establish a bijection between the sets. As we will see, this is, at its core, a result about fixed points of monotone functions on complete lattices.

**Theorem 3.9** (Knaster-Tarski)**.** *Let* $L$ *be a complete lattice and* $f : L \to L$ *a monotone function. Then*

$$\alpha = \bigvee \{x \in L : x \leq f(x)\}$$

*is a fixed point of* $f$*. Further,* $\alpha$ *is the greatest fixed point of* $f$*.*

Proof. Let $H = \{x \in L : x \leq f(x)\}$. For all $x \in H$ we have $x \leq \alpha$, so $x \leq f(x) \leq f(\alpha)$. Thus $f(\alpha)$ is an upper bound of $H$, so that $\alpha \leq f(\alpha)$. By monotonicity of $f$ we have $f(\alpha) \leq f(f(\alpha))$. So $f(\alpha) \in H$, i.e. $f(\alpha) \leq \alpha$. So $\alpha$ is a fixed point of $f$. If $f(\beta) = \beta$ then $\beta \in H$ and so $\beta \leq \alpha$. ∎

**Corollary 3.10** (Banach's Decomposition)**.** *Let* $X$ *and* $Y$ *be sets with* $f : X \to Y$ *and* $g : Y \to X$*. Then there exist disjoint subsets* $X_1$ *and* $X_2$ *of* $X$ *and* $Y_1$ *and* $Y_2$ *of* $Y$ *such that* $f(X_1) = Y_1$*,* $g(Y_2) = X_2$*,* $X = X_1 \sqcup X_2$ *and* $Y = Y_1 \sqcup Y_2$*.*

Proof. The key observation is that $\mathcal{P}(X)$ and $\mathcal{P}(Y)$ form complete lattices. For any set $S$, let $\alpha_S : \mathcal{P}(S) \to \mathcal{P}(S)$ denote complement, i.e. $\alpha_S(T) = S \setminus T$. Then define

$$\varphi : \mathcal{P}(X) \to \mathcal{P}(X) \quad \text{via} \quad S \mapsto \alpha_X \circ g_* \circ \alpha_Y \circ f_*(S).$$

Since $\varphi$ is the composition of two monotone functions and two antitone functions, it is itself monotone. Apply Knaster-Tarski to obtain a fixed point. ∎

**Theorem 3.11** (Schröder-Bernstein)**.** *Let* $X$ *and* $Y$ *be sets and suppose there exist injective maps* $f : X \to Y$ *and* $g : Y \to X$*. Then there is a bijective map* $h : X \to Y$*.*

Proof. Let $f$ and $g$ as in the Banach decomposition theorem be injective. Then we may set

$$h : X \to Y$$

to be defined as $f$ on $A$ and $g^{-1}$ on $X \setminus A$. This completes the proof. ∎

## 3.5 FUNCTIONS AND SET OPERATIONS

Here's how image and preimage behave with respect to union, intersection, and complement:

$$f_*\left(\bigcup_{\lambda\in\Lambda} X_\lambda\right) = \bigcup_{\lambda\in\Lambda} f_*(X_\lambda) \qquad f^*\left(\bigcup_{\lambda\in\Lambda} Y_\lambda\right) = \bigcup_{\lambda\in\Lambda} f^*(Y_\lambda)$$

$$f_*\left(\bigcap_{\lambda\in\Lambda} X_\lambda\right) \subseteq \bigcap_{\lambda\in\Lambda} f_*(X_\lambda) \qquad f^*\left(\bigcap_{\lambda\in\Lambda} Y_\lambda\right) = \bigcap_{\lambda\in\Lambda} f^*(Y_\lambda)$$

$$f_*(X^c) \subseteq (f_*(X))^c \qquad f^*(Y^c) = (f^*(Y))^c$$

Furthermore, we have equality in the intersection image case when $f$ is injective, and equality in the complement image case when $f$ is surjective.

## 3.6 ARITHMETIC ON THE NONNEGATIVE INTEGERS

Addition and multiplication are characterized on $\mathbf{Z}_{\geq 0}$ as follows.

**Theorem 3.12** (Characterization of arithmetic in $\mathbf{Z}_{\geq 0}$)**.** *We have:*

| | | | |
|---|---|---|---|
| $\forall x$ | $0 \neq x^{++}$ | $\forall x, y$ | $x + y^{++} = (x + y)^{++}$ |
| $\forall x, y$ | $(x^{++} = y^{++}) \leq (x = y)$ | $\forall x$ | $x \cdot 0 = 0$ |
| $\forall x$ | $x + 0 = x$ | $\forall x, y$ | $x \cdot y^{++} = x + x \cdot y$ |

Proof. These are proved by induction, I'll do them later. ∎

**Theorem 3.13.** *Let $X$ be a set, $a \in X$, $f : X \to X$. Then there is a unique function $F : \mathbf{Z}_{\geq 0} \to X$ such that $F(0) = a$ and $F(n^{++}) = f(F(n))$ for any natural number $n$.*

**Definition 3.14.** A **commutative monoid** is a set $M$ along with a binary operation

$$\star : M \times M \to M$$

such that:

- (Associativity) $a \star (b \star c) = (a \star b) \star c$ for all $a, b, c \in M$.
- (Identity) there exists $e \in M$ such that $e \star a = a = a \star e$ for all $a \in M$.
- (Commutativity) $a \star b = b \star a$ for all $a, b \in M$.

Hence, a monoid can be specified by naming a set, a binary operation, and an identity element.

**Definition 3.15.** A **commutative semiring** is a set $N$ equipped with two binary operations

$$+ : N \times N \to N \qquad \text{and} \qquad \cdot : N \times N \to N$$

such that:

- $(N, +, 0)$ and $(N, \cdot, 1)$ are commutative monoids,
- (Annihilation) $n \cdot 0 = 0$ for any $n \in N$,
- (Distributivity) $a \cdot (b + c) = a \cdot b + a \cdot c$ for any $a, b, c \in N$.

**Proposition 3.16.** $(\mathbf{Z}_{\geq 0}, +, \cdot, 0, 1)$ *forms a commuative semiring.*

## 3.7 INTEGERS AND RATIONAL NUMBERS

**Definition 3.17.** A *setoid* is a set $X$ equipped with an equivalence relation $\sim_X$. This $\sim_X$ relation *partitions* $X$ into disjoint equivalence classes. From the setoid $(X, \sim_X)$ we may form the **quotient set** $X/\sim_X$, which is the actual set of all equivalence classes of $X$.

This gives us $\mathbf{Z}$ and $\mathbf{Q}$.

**Definition 3.18.** Consider the nonnegative integers $\mathbf{Z}_{\geq 0}$ with addition and define the relation $\sim_1$ on $\mathbf{Z}_{\geq 0} \times \mathbf{Z}_{\geq 0}$ as follows:

$$((a, b) \sim_1 (c, d)) = (a + d = b + c).$$

The resulting quotient is the integers $\mathbf{Z}$.

**Definition 3.19.** A **commutative ring** is a commutative semiring in which every element has an additive inverse, i.e. a semiring $R$ is a ring when for every $r \in R$ there exists $r' \in R$ such that $r + r' = 0$.

A commutative nonzero ring that satisfies the *cancellation property*

$$(ac = bc \wedge c \neq 0) \leq (a = b),$$

is called a **domain**. A domain where every nonzero element has a multiplicative inverse is called a **field**.

We now introduce a strict partial order to the ring structure.

**Definition 3.20.** An **ordered ring** is a 6-tuple $(F, +, \cdot, 0, 1, <)$ such that:

- $(F, +, \cdot, 0, 1)$ forms a nonzero commutative ring
- for all $x \in F$, $x \not< x$
- for all $x, y, z \in F$, if $x < y$ and $y < z$ then $x < z$
- for all $x, y \in F$, one of the following hold:

$$x < y, \qquad y < x, \qquad x = y$$

- for all $x, y, z \in F$, if $x < y$ then $x + z < y + z$
- for all $x, y, z \in F$, if $x < y$ and $0 < z$ then $x \cdot z < y \cdot z$

That is, an ordered ring is a set in which one can add, subtract, multiply, and compare elements. An **ordered field** is an ordered ring in which one can divide by nonzero elements.

**Proposition 3.21.** $(\mathbf{Z}, +, \cdot, 0, 1, <)$ *forms an ordered domain.*

**Definition 3.22.** Consider the integers $\mathbf{Z}$ with multiplication and define the relation $\sim_2$ on $\mathbf{Z} \times \mathbf{Z}$ as follows:

$$((a, b) \sim_2 (c, d)) = (ad = bc).$$

The resulting quotient is the rationals $\mathbf{Q}$.

**Proposition 3.23.** $\mathbf{Q}$ *forms an ordered field.*

## 3.8 CAUCHY SEQUENCES IN AN ORDERED FIELD

The elements of an ordered field can be divided into positive, negative, and 0. This is enough to have an absolute value.

**Definition 3.24.** Let $F$ be an ordered field. If $\nu : F \to F$ is such that:

- for all $x \in F$, $\nu(x) \geq 0$
- for all $x \in F$, $(\nu(x) = 0) = (x = 0)$
- for all $x, y \in F$, $\nu(xy) = \nu(x)\nu(y)$
- for all $x, y \in F$, $\nu(x + y) \leq \nu(x) + \nu(y)$

then $\nu$ is called an **absolute value**.

The usual absolute value on $\mathbf{Q}$ is $|\cdot| : \mathbf{Q} \to \mathbf{Q}$, given by

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

It satisfies the properties above. We check the last one.

**Theorem 3.25** (Rational triangle inequality)**.** *For rational numbers* $x$ *and* $y$,

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

Proof. By cases.

- Suppose $x \geq 0$ and $y \geq 0$. Then $|x+y| = x+y = |x|+|y|$.
- Similarly, if $x < 0$ and $y < 0$, then

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

- Finally, suppose $x \geq 0$ but $y < 0$.
  - If $x+y \geq 0$, then $|x+y| = x+y < x-y = |x|+|y|$.
  - If $x+y < 0$, then $|x+y| = -(x+y) = -x-y < x-y = |x|+|y|$.
- Note that by symmetry we need not check when $x < 0$ and $y \geq 0$.

In all cases the inequality holds, as desired. ∎

We now discuss sequences in ordered fields. We focus on Cauchy sequences.

Throughout this section, let $\Gamma$ be an ordered field.

**Definition 3.26.** Denote by $\Gamma^+$ the positive elements of $\Gamma$.

A sequence $(x_n)_n \subseteq \Gamma$ is **Cauchy** if for any $\varepsilon \in \Gamma^+$ there is an $N \in \mathbf{Z}^+$ such that

$$n, n' \geq N \quad \text{implies} \quad |x_{n'} - x_n| < \varepsilon.$$

Denote by $\kappa(\Gamma)$ the set of all Cauchy sequences on $\Gamma$.

Because $\Gamma$ is an ordered field, we may add and multiply sequences with entries in $\Gamma$ by adding or multiplying their terms together.

**Proposition 3.27.** *$\kappa(\Gamma)$ is closed under addition.*

Proof. Let $(a_n)_n, (b_n)_n \in \kappa(\Gamma)$, and let $\varepsilon > 0$. Then $\varepsilon/2 > 0$ as well, so there exists $N_a$ such that for all $n, n' \geq N_a$ we have

$$|a_n - a_{n'}| < \varepsilon/2,$$

and an $N_b$ such that for all $n, n' \geq N_b$ we have

$$|b_n - b_{n'}| < \varepsilon/2.$$

Let $N = \max(\{N_a, N_b\})$. Then by the triangle inequality,

$$|(a_n + b_n) - (a_{n'} + b_{n'})| \leq |a_n - a_{n'}| + |b_n - b_{n'}| < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

so $(a_n + b_n)_n \in \kappa(\Gamma)$. ∎

Cauchy sequences are bounded.

**Proposition 3.28.** *If $(x_n)_n$ is Cauchy, then there is an $M \in \Gamma^+$ such that*

$$|x_n| < M$$

*for all nonnegative $n$.*

Proof. Let $\varepsilon = 1$. Then there exists an $N$ such that for all $m, n \geq N$ we have $|x_n - x_m| < 1$. Let $M = \max(\{|x_0|, \ldots, |x_N|\}) + 1$. Clearly $|x_n| < M$ for all $n \leq N$. Now suppose $n > N$. Then

$$|x_n| = |x_n - x_N + x_N| \leq |x_n - x_N| + |x_N| < M$$

so in all cases $(x_n)_n$ is bounded by $M$. ∎

The product of two Cauchy sequences is Cauchy.

**Proposition 3.29.** *$\kappa(\Gamma)$ is closed under multiplication.*

Proof. Let $(a_n)_n, (b_n)_n \in \kappa(\Gamma)$. By the previous proposition, there exist $M_a \in \Gamma^+$ bounding $(a_n)_n$ and $M_b \in \Gamma^+$ bounding $(b_n)_n$. Let $\varepsilon \in \Gamma^+$ so that $\lambda = \frac{\varepsilon}{M_a + M_b} \in \Gamma^+$ as well. Since $(a_n)_n$ is Cauchy, there exists $N_a \in \mathbf{Z}^+$ such that for all $n, n' \geq N_a$, we have $|a_n - a_{n'}| < \lambda$ and also $N_b \in \mathbf{Z}^+$ such that for all $n, n' \geq N_b$, we have $|b_n - b_{n'}| < \lambda$. Pick $N = \max(\{N_a, N_b\})$. Then

$$\begin{aligned} |a_n b_n - a_{n'} b_{n'}| &= |a_n b_n - a_n b_{n'} + a_n b_{n'} - a_{n'} b_{n'}| \\ &\leq |a_n||b_n - b_{n'}| + |a_n - a_{n'}||b_{n'}| \\ &< \lambda(M_a + M_b) = \varepsilon. \end{aligned}$$

Thus, $(a_n b_n)_n$ is Cauchy. ∎

**Definition 3.30.** Say that $(a_n)_n \in \kappa(\Gamma)$ **tends to zero** if for every $\varepsilon \in \Gamma^+$ there is some $N_\varepsilon \in \mathbf{Z}^+$ such that for all $n \geq N$, we have $|a_n| < \varepsilon$.

Say $(a_n)_n \sim_3 (b_n)_n$ if $(a_n - b_n)_n$ tends to zero.

**Definition 3.31.** $\mathbf{R} = \kappa(\mathbf{Q})/\sim_3$.

We also have the following:

**Definition 3.32.** Let $\mathbf{R}[x]$ denote the commutative ring of all polynomials with real coefficients. Define

$$\mathbf{C} := \frac{\mathbf{R}[x]}{(x^2 = -1)}.$$

This forms a field.

That is, complex numbers are polynomials in $i$ with real coefficients, where

$$i^2 = -1.$$

We also care about the the subsets

$\mathbf{C}^\times$ (nonzero complex numbers) and

$U_1$ (complex numbers of unit modulus)[2]

as well as the set $\widehat{\mathbf{C}}$ (complex numbers plus a point $\infty$).

We have the following diagram:

$$\begin{array}{ccccccccc} & & & & U_1 & \longrightarrow & \mathbf{C}^\times & \longrightarrow & \mathbf{C} & \longrightarrow & \widehat{\mathbf{C}} \\ & & & & & & & & \uparrow & & \\ \mathbf{Z}_{\geq 0} & \longrightarrow & \mathbf{Z} & \longrightarrow & \mathbf{Q} & \longrightarrow & \mathbf{R} & \longrightarrow & \overline{\mathbf{R}} & & \end{array}$$

where all arrows are understood to be inclusion maps.

[2]Also common is the notation $U(1)$.

## 3.9 EXERCISES AND REMARKS

Here we give interesting bijections between $\mathbf{Z}_{\geq 0}$, $\mathbf{Z}$, and $\mathbf{Q}$.

1. Is there a natural way to enumerate (i.e. to sequence) the integers?

   One way to do it is to note that every integer can be written as a sum of powers of $-2$ in a unique way. This is sometimes called *negative binary*.

   For example, $-1 = 11_{-2}$ and $8 = 11000_{-2}$.

   So we can get an enumeration of the integers by simply counting in base $-2$.

   Here are the first few terms of that enumeration:

   | $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
   |---|---|---|---|---|---|---|---|---|---|---|---|---|---|
   | $m_n$ | 0 | 1 | $-2$ | $-1$ | 4 | 5 | 2 | 3 | $-8$ | $-7$ | $-10$ | $-9$ | $-4$ |

   Note how the enumeration alternates between 1 positive number, then 2 negative numbers, then 4 positive numbers, and so on.

   This gives us a bijection $f : \mathbf{Z}_{\geq 0} \to \mathbf{Z}$.

2. Is there a natural way to enumerate the rationals?

   One way to do it is to use our bijection from the previous section combined with the fundamental theorem of arithmetic: send $n$ to $f(n)$, then map all of the exponents $e_i$ in the prime factorization of $f(n)$ to $f(e_i)$.

   Here are the first few terms of that enumeration:

   | $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
   |---|---|---|---|---|---|---|---|---|---|---|---|---|---|
   | $m_n$ | 0 | 1 | $-2$ | $-1$ | $\frac{1}{4}$ | 5 | 2 | 3 | $-\frac{1}{2}$ | $-7$ | $-10$ | $-\frac{1}{9}$ | $-\frac{1}{4}$ |

   This gives us a bijection $g : \mathbf{Z}_{\geq 0} \to \mathbf{Q}$.

I learned about Schröder-Bernstein via Knaster-Tarski through a handout by Joel Shapiro.

This book deals mostly with classical mathematics. There is also category theory, though. A brief introduction to the basics may help aid understanding certain concepts later on. If you do not wish to engage with categories, do not worry, as they will stay at the ends of the chapters, never to enter the main body of text.

**Definition 3.33.** A **category** $\mathcal{C}$ consists of:

- a class of *objects* $\mathcal{C}_0$,
- for any two objects $x, y$ a set of *arrows* $\mathcal{C}_1(x, y)$,[3]
- for any three objects $a, b, c$ an associative binary operation

$$\mathcal{C}_1(a, b) \times \mathcal{C}_1(b, c) \to \mathcal{C}_1(a, c)$$

  called *composition of arrows*,
- for any object $x$ an arrow $\mathrm{id}_x \in \mathcal{C}_1(x, x)$.

We'll also use the notation $\mathcal{C}_{\simeq}(x, y)$ for the set of isomorphisms between $x$ and $y$.

Here is an entire table of examples. The objects and arrows are things we will encounter, though the actual categories will stay in the background.

| $\mathcal{C}$ | $\mathcal{C}_0$ | $\mathcal{C}_1$ | $\mathcal{C}_{\simeq}$ |
|---|---|---|---|
| Set | sets | functions | bijections |
| Poset | posets | monotone fns | order isomorphisms |
| Top | topological spaces | continuous fns | homeomorphisms |
| Met | metric spaces | 1-Lipschitz fns | isometries |
| $\mathsf{Vect}_k$ | vector spaces over $k$ | $k$-linear maps | $k$-linear bijections |
| $\mathsf{Norm}_k$ | normed spaces over $k$ | 1-Lipschitz $k$-linear maps | $k$-linear isometries |
| Orth | inner product spaces over $\mathbf{R}$ | orthogonal maps | orthogonal maps |
| Unit | inner product spaces over $\mathbf{C}$ | unitary maps | unitary maps |

Other examples include:

- FinSet, the category of finite sets;
- FinPoset, the category of finite posets;
- $\mathsf{FinVect}_k$, the category of finite dimensional vector spaces over $k$;
- $\mathsf{FinNorm}_k$, FinOrth, FinUnit;
- CompMet, the category of complete metric spaces;
- Banach, the category of complete normed spaces (i.e. Banach spaces);
- Hilbert, the category of complete inner-product spaces (i.e. Hilbert spaces).

[3] Herein, $x$ is the *source* object and $y$ the *target* object.

# Chapter 4

# NEARNESS AND DISTANCE

In everyday life, a *ruler* is essentially some reference distance that we use to measure the length of other distances. We can mark a finite number of other reference points on that ruler to make it do the job of finitely many other rulers, but ultimately each ruler can only measure a limited number of distances.

We ignore this limitation in everyday life by settling for an approximation. If we want to measure lengths precisely all the time, though, things get interesting.

Associated to every ruler is a set $U$ such that if the length of the object being measured is in $U$, then this can be verified by holding the ruler up to the object and seeing that the ruler is entirely within the object. We'll call such a ruler $R_U$.

Now suppose you have two rulers, $R_U$ and $R_V$. To find out if if a length lies within $U \cup V$, one first holds up $R_U$ to the object and then $R_V$. If at least one ruler witnesses (fits within) the length, then the length lies within $U \cup V$. So we basically have a "phantom ruler" $R_{U \cup V}$.

Similarly, to find out if a length lies within $U \cap V$, one simply verifies that both $R_U$ and $R_V$ do the job. So $R_{U \cap V}$ is another "phantom ruler."

Now suppose you have a sequence of rulers $(R_{U_n})_{n \geq 0}$. To prove that some length lies in $\bigcup_{n \geq 0} U_n$, one only needs to provide a single witness ruler. But to prove that some length lies in $\bigcap_{n \geq 0} U_n$, one needs infinitely many witnesses, which makes for an infinitely long proof. Thus, $R_{\bigcup_n U_n}$ is a "phantom ruler," but $R_{\bigcap_n U_n}$ is not.

This is topology.

## 4.0 TOPOLOGICAL SPACES

**Definition 4.1.** A **topology** on a set $X$ consists of a set of *open sets* $\tau$ such that

1. $\varnothing \in \tau$ and $X \in \tau$
2. $\tau$ is stable under arbitrary unions:

$$((X_\lambda)_{\lambda\in\Lambda} \subseteq \tau) \leq \left(\bigcup_{\lambda\in\Lambda} X_\lambda \in \tau\right).$$

3. $\tau$ is stable under finite intersections:

$$\left((X_n)_{n=0}^{N} \subseteq \tau\right) \leq \left(\bigcap_{n=0}^{N} X_n \in \tau\right).$$

we call $(X,\tau)$ a **topological space**.

Every topological space also has a set of *closed sets* $\kappa$ such that

1. $\varnothing \in \kappa$ and $X \in \kappa$
2. $\kappa$ is stable under arbitrary intersections
3. $\kappa$ is stable under finite unions

The complement of an open set is a closed set: $(U \in \tau) = (X \setminus U \in \kappa)$.

Here are three examples of topological spaces.

- The coarsest (i.e. fewest open sets) example of a topology on $X$ is

  $$\tau = \{\varnothing, X\}.$$

  This is called the **trivial topology**.

- The finest (i.e. most open sets) example of a topology on $X$ is

  $$\tau = \mathcal{P}(X).$$

  This is called the **discrete topology**.

- The set of subsets $\{\varnothing, \{1\}, \{0,1\}\}$ forms a topology on the set 2 which is neither trivial nor discrete; this space is known as the **Sierpinski space**.

## 4.1 NEIGHBORHOODS AND POINTS

**Definition 4.2.** Let $X$ be a topological space, $S \subseteq X$.

A **neighborhood** of $S$ is a subset $V$ of $X$ containing an open set $U$ containing $S$:

$$S \subseteq U \subseteq V \subseteq X.$$

In particular, we are usually interested in the case when $S$ is just a single point.

Consider a subset $S$ of a topological space $X$.

- A point $x$ is an **interior point** of $S$ if $S$ is a neighborhood of $x$. The set of all interior points of $S$ is called the **interior** of $S$ and is denoted $S^\circ$.

- A point $x$ is a **boundary point** of $S$ if all neighborhoods of $x$ contain at least one point in $S$ and one point not in $S$. The set of all boundary points of $S$ is called the **boundary** of $S$ and is denoted $\partial S$.

- A point $x$ is a **limit point** of $S$ if all neighborhoods of $x$ contain at least one point of $S$ different from $x$ itself. Note that a limit point of $S$ does not have to be an element of $S$. The union of $S$ with the set of all limit points of $S$ is called the (topological) **closure** of $S$ and is denoted $\overline{S}$.

**Theorem 4.3.** *Let $S$ be a subset of a topological space $X$. Then $S$ is open if and only if $S = S^\circ$, and $S$ is closed if and only if $S = \overline{S}$.*

Proof. Suppose $S$ is open. Clearly we always have $S^\circ \subseteq S$, so it remains to show $S \subseteq S^\circ$. Let $x \in S$. Since $S$ is open, $S$ is a neighborhood of $x$:

$$x \in S \subseteq S \subseteq X.$$

So $x \in S^\circ$. Conversely, suppose $S \subseteq S^\circ$. Then every point of $S$ is an interior point, that is, for every point $x \in S$ there exists an open set $U_x$ such that $x \in U_x \subseteq S$. Then $\bigcup_{x \in S} U_x = S$, and since $S$ is a union of open sets, $S$ must itself be open.

Now suppose $S$ is closed. Clearly we always have $S \subseteq \overline{S}$, so it remains to show $\overline{S} \subseteq S$. Let $x \in \overline{S}$. Then either $x \in S$ or $x$ is a limit point of $S$. Suppose $x \notin S$ and $x$ is a limit point of $S$. If $S$ is closed, then $S^c$ is open. If $x \notin S$, then necessarily $x \in S^c$. Note that $S^c$ is a neighborhood of $x$ by the same trick as before. However, since $x$ is a limit point of $S$, we see $S^c$ must contain a point of $S$ (distinct from $x$), contradiction. So we must have $x \in S$.

Conversely, suppose every limit point of $S$ is itself in $S$. Let $x \in S^c$. Since $x$ is not in $S$, $x$ is not a limit point of $S$. Thus it is not the case that all neighborhoods of $x$ contain at least one point of $S$ different from $x$ itself. Thus there exists a neighborhood of $x$ that does not contain any points of $S$, which amounts to there being an open set $U$ such that $x \in U$ and such that $U$ is contained within $S^c$. Thus $S^c$ is a neighborhood of $x$, which means every point in $S^c$ is an interior point, which means $S^c$ is open, hence $S$ is closed. ∎

## 4.2 CONTINUOUS FUNCTIONS

**Definition 4.4.** Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be topological spaces.

A function $f : X \to Y$ is **continuous** if

$$(U \in \tau_Y) \leq (f^*(U) \in \tau_X),$$

that is, if the preimage of every open set is open.

Let $Y$ be a topological space.

- Consider the discrete space $(X, \mathcal{P}(X))$. Any function $f : X \to Y$ is continuous. Conversely, the only continuous functions $g : Y \to X$ are the locally constant functions, where a function $g$ is locally constant if every point has a neighborhood on which $g$ is constant.
- Now consider the trivial space $(S, \{\varnothing, S\})$. The only continuous functions $f : S \to Y$ are the constant functions, whereas every function $g : Y \to S$ is continuous.

**Proposition 4.5.** *Continuous functions are stable under composition.*

Proof. Let $f : X \to Y$ and $g : Y \to Z$ be continuous. Take $U_Z$ to be open in $Z$. Then since $g$ is continuous, $g^*(U_Z)$ is open in $Y$, so $f^* \circ g^*(U_Z) = (g \circ f)^*(U_Z)$ is open in $X$ by continuity of $f$. ∎

## 4.3 CONNECTEDNESS AND COMPACTNESS

**Definition 4.6.** A topological space $(X, \tau)$ is **connected** if there are no $X_1, X_2 \in \tau$ such that $X$ can be written as $X_1 \cup X_2$ with $X_1 \cap X_2 = \varnothing$.

The continuous image of a connected set is another connected set.

**Proposition 4.7.** *Let $f : X \to f_*(X)$ be continuous. If $X$ is connected then so is $f_*(X)$.*

Proof. By contraposition. If $f_*(X)$ is disconnected then there exist open sets $Y_1$ and $Y_2$ such that $f_*(X) = Y_1 \cup Y_2$ with $Y_1 \cap Y_2 = \varnothing$. Then

$$X = f^*(f_*(X)) = f^*(Y_1 \cup Y_2) = f^*(Y_1) \cup f^*(Y_2)$$

where $f^*(Y_1) \cap f^*(Y_2) = f^*(Y_1 \cap Y_2) = f^*(\varnothing) = \varnothing$, showing that $X$ is disconnected. Hence, if $X$ is connected, then $f_*(X)$ must be connected. ∎

**Definition 4.8.** A topological space $(X, \tau)$ is **compact** if for every open cover of $X$ there exists a finite subcover of $X$.

The continuous image of a compact set is another compact set.

**Proposition 4.9.** *Let* $f : X \to f_*(X)$ *be continuous. If* $X$ *is compact then so is* $f_*(X)$.

Proof. Suppose that $f$ is continuous and $X$ is compact, and let $\mathcal{U} = \{U_\alpha : \alpha \in I\}$ be an open cover of $f(X)$ (where $I$ is just some index set). Then each $U_\alpha$ is open, and $\bigcup_{\alpha\in I} U_\alpha = f(X)$. Since $f$ is continuous, each $f^{-1}(U_\alpha)$ is open, and

$$X = f^{-1}(f(X)) = f^{-1}(\bigcup_{\alpha\in I} U_\alpha) = \bigcup_{\alpha\in I} f^{-1}(U_\alpha),$$

i.e. $\mathcal{V} = \{f^{-1}(U_\alpha) : \alpha \in I\}$ is an open cover for $X$. Since $X$ is compact, we can reduce this to a finite subcover $\mathcal{V}' = \{f^{-1}(U_i) : 1 \le i \le n\}$ where $n \in \mathbf{Z}_{\ge 0}$. So each $f^{-1}(U_i)$ is open and $\bigcup_{i=1}^{n} f^{-1}(U_i) = X$. But then

$$f(X) = f(\bigcup_{i=1}^{n} f^{-1}(U_i)) = \bigcup_{i=1}^{n} f(f^{-1}(U_i)) = \bigcup_{i=1}^{n} U_i,$$

where the last equality follows since $f$ is surjective. So we have found a finite subcover for $f(X)$, which means $f(X)$ must also be compact. ∎

## 4.4 HAUSDORFF SPACES

Continuous bijection from a compact space to a Hausdorff space has continuous inverse (i.e. is a homeomorphism).

## 4.5 METRIC SPACES

**Definition 4.10.** Let $X$ be a set. A **metric** on $X$ is a function

$$d : X^2 \to \mathbf{R}$$

satisfying the following properties:[1]

1. $d(x,y) \geq 0$ with $d(x,y) = 0$ iff $x = y$
2. $d(x,y) = d(y,x)$
3. $d(x,z) \leq d(x,y) + d(y,z)$

we call the ordered pair $(X,d)$ a **metric space**.

Metric spaces are Hausdorff.

Previously we defined the notion of limit point of a set. Related to this notion is the limit of a sequence.

**Definition 4.11.** Let $X$ be a metric space. A sequence $(x_n)_n \subseteq X$ **converges to a limit** $x$ if for every $\varepsilon > 0$ there is an $N > 0$ such that

$$n \geq N \quad \text{implies} \quad d(x,x_n) < \varepsilon.$$

We write $\lim_{n\to\infty} x_n = x$ or simply $x_n \to x$.

[1] We're defining metrics to map into the reals (which is what should happen), but we haven't given a construction of **R** yet. When we do, we will do so via the "Cauchy sequence of Cauchy sequences" argument, and our metrics will be assumed to map into **Q**.

In metric spaces, limits are unique:

**Proposition 4.12.** *Let $X$ be a metric space. If $(x_n)_n \subseteq X$ converges to $x$ and $y$, then we must have $x = y$.*

Proof. Let $\varepsilon > 0$, so that $\varepsilon/2 > 0$ as well. Then there exists $N_x > 0$ such that

$$n \geq N_x \quad \text{implies} \quad d(x, x_n) < \varepsilon/2,$$

and $N_y$ such that

$$n \geq N_y \quad \text{implies} \quad d(y, x_n) < \varepsilon/2.$$

Pick $N = \max(\{N_x, N_y\})$. Then $n \geq N$ implies

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

so $x = y$. ∎

The metric respects convergence of sequences:

**Proposition 4.13.** *Let $X$ be a metric space with $(x_n)_n \subseteq X$, $(y_n)_n \subseteq X$, and $x, y \in X$ such that $x_n \to x$ and $y_n \to y$. Then $d(x_n, y_n) \to d(x, y)$.*

Proof. Let $\varepsilon > 0$, so that $\varepsilon/2 > 0$ as well. Then there exists $N_x > 0$ such that

$$n \geq N_x \quad \text{implies} \quad d(x, x_n) < \varepsilon/2,$$

and $N_y$ such that

$$n \geq N_y \quad \text{implies} \quad d(y, x_n) < \varepsilon/2.$$

Pick $N = \max(\{N_x, N_y\})$. Then by a previous proposition,

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

so $d(x_n, y_n) \to d(x, y)$, as desired. ∎

## 4.6 CAUCHY SEQUENCES

There are also sequences whose terms grow closer and closer together without necessarily reaching a limit.

**Definition 4.14.** Let $X$ be a metric space. A sequence $(x_n)_n \subseteq X$ is **Cauchy** if for every $\varepsilon > 0$ there is an $N > 0$ such that

$$n, m \geq N \quad \text{implies} \quad d(x_m, x_n) < \varepsilon.$$

Every converging sequence is Cauchy:

**Proposition 4.15.** *Let $X$ be a metric space and $(x_n)_n \subset X$ a sequence converging to $x \in X$. Then $(x_n)_n$ is Cauchy.*

Proof. Let $\varepsilon > 0$. Then $\varepsilon/2 > 0$ as well; so there exists an $N$ such that

$$d(x_n, x) < \varepsilon/2$$

for all $n \geq N$. But then for $m > n$,

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

which shows that $(x_n)_n$ is Cauchy. ∎

Not every Cauchy sequence converges, however. Consider the sequence

$$\left(\frac{\lfloor 10^n \sqrt{2} \rfloor}{10^n}\right)_n = (1,\ 1.4,\ 1.41,\ 1.414,\ \ldots)$$

in $\mathbf{Q}$. The sequence converges to $\sqrt{2}$, which is not in $\mathbf{Q}$.

## 4.7 COMPLETE METRIC SPACES

**Theorem 4.16.** *If a metric space completion exists, it is unique up to isometry.*

Proof. Let $X$ be a metric space and let $X^*, X^{**}$ be completions of $X$.

Every $x^* \in X^*$ has a Cauchy sequence $(x_n)_n \subset X$ such that $x_n \to x^*$ by completeness of $X^*$. But by completeness of $X^{**}$, we have $x_n \to x^{**}$ for some $x^{**} \in X^{**}$. Note that if

$y_n \to x^*$ then $y_n - x_n \to 0$, hence $y_n \to x^{**}$; that is, the map $f : X^* \to X^{**}$ given by $f(x^*) = x^{**}$ is well defined. Note that $f(x) = x$ for all $x \in X$.

Now suppose $x_n \to x^*, y_n \to y^*$ in $X^*$ and also $x_n \to x^{**}, y_n \to y^{**}$ in $X^{**}$. Then

$$\begin{aligned} d_*(x^*, y^*) &= \lim_{n\to\infty} d_*(x_n, y_n) \\ &= \lim_{n\to\infty} d(x_n, y_n) \\ &= \lim_{n\to\infty} d_{**}(x_n, y_n) = d_{**}(f(x^*), f(y^*)), \end{aligned}$$

so $f$ is an isometry. ∎

**Theorem 4.17** (Construction of $\mathbf{R}$)**.** *A completion of* $\mathbf{Q}$ *exists.*

Proof. Let $\mathbf{R} = \hat{\kappa}(\mathbf{Q})$. From previous work we know that $\mathbf{R}$ is an archimedean ordered field, which means that $\mathbf{Q} \subseteq \mathbf{R}$. We need to show that $\overline{\mathbf{Q}} = \mathbf{R}$ and that $\mathbf{R}$ is a complete metric space.

For elements $\alpha = (a_n)_n$, $\beta = (b_n)_n$ of $\kappa(\mathbf{Q})$, define

$$d_{\mathbf{R}}([\alpha], [\beta]) = [(d_{\mathbf{Q}}(a_n, b_n))_n].$$

By a previous proposition,

$$|d_{\mathbf{Q}}(a_m, b_m) - d_{\mathbf{Q}}(a_n, b_n)| \le d_{\mathbf{Q}}(a_m, a_n) + d_{\mathbf{Q}}(b_m, b_n),$$

i.e. the sequence of distances is Cauchy because $\alpha$ and $\beta$ are.

We claim that

$$(\alpha \sim \beta) = (d_{\mathbf{R}}([\alpha], [\beta]) = 0_{\mathbf{R}}).$$

Indeed, $\alpha \sim \beta$ if and only if for every $\varepsilon \in \mathbf{Q}^+$ there is some $N_\varepsilon \in \mathbf{Z}^+$ such that for all $n \ge N_\varepsilon$, $d_{\mathbf{Q}}(a_n, b_n) < \varepsilon$. But

$$(d_{\mathbf{Q}}(a_n, b_n) < \varepsilon) = (d_{\mathbf{Q}}(d_{\mathbf{Q}}(a_n, b_n), 0) < \varepsilon)$$

so the condition is equivalent to $d_{\mathbf{R}}([\alpha], [\beta]) = 0_{\mathbf{R}}$.

Another claim: the distance doesn't depend on equivalence class representative. This is easily shown if we know that $d_{\mathbf{R}}$ satisfies the triangle inequality, because then, if $\alpha \sim \alpha'$ and $\beta \sim \beta'$, we have

$$d_{\mathbf{R}}([\alpha], [\beta]) \le d_{\mathbf{R}}([\alpha], [\alpha']) + d_{\mathbf{R}}([\alpha'], [\beta']) + d_{\mathbf{R}}([\beta'], [\beta]) = d_{\mathbf{R}}([\alpha'], [\beta'])$$

$$d_{\mathbf{R}}([\alpha'],[\beta']) \leq d_{\mathbf{R}}([\alpha],[\alpha']) + d_{\mathbf{R}}([\alpha],[\beta]) + d_{\mathbf{R}}([\beta'],[\beta]) = d_{\mathbf{R}}([\alpha],[\beta])$$

showing that $d_{\mathbf{R}}([\alpha],[\beta]) = d_{\mathbf{R}}([\alpha'],[\beta'])$.

Let $\alpha = (a_n)_n$, $\beta = (b_n)_n$, $\gamma = (c_n)_n$ be in $\kappa(\mathbf{Q})$. Verifying the triangle inequality comes down to showing that

$$[(d_{\mathbf{Q}}(a_n, b_n))_n] \leq [(d_{\mathbf{Q}}(a_n, c_n) + d_{\mathbf{Q}}(c_n, b_n))_n].$$

Suppose equality does not hold. Then there exists an $\varepsilon \in \mathbf{Q}^+$ such that for all $N \in \mathbf{Z}^+$ there is some $n_0 \geq N$ such that

$$d_{\mathbf{Q}}(d_{\mathbf{Q}}(a_{n_0}, c_{n_0}) + d_{\mathbf{Q}}(c_{n_0}, b_{n_0}), d_{\mathbf{Q}}(a_{n_0}, b_{n_0})) \geq \varepsilon$$

We may use the $n_0$ indices to build subsequences such that there is some $\varepsilon \in \mathbf{Q}^+, M \in \mathbf{Z}^+$ such that for all $m \geq M$,

$$d_{\mathbf{Q}}(a_m, b_m) < d_{\mathbf{Q}}(a_m, c_m) + d_{\mathbf{Q}}(c_m, b_m) + \varepsilon.$$

So the triangle inequality holds, which implies that the distance is well-defined.

Since $d_{\mathbf{R}}$ is built from $d_{\mathbf{Q}}$, it is both symmetric and nonnegative. We just checked the triangle inequality, and the fact that distance zero implies points are equal is a consequence of how we defined the quotient space $\hat{\kappa}(\Gamma)$. So $d_{\mathbf{R}}$ is a metric, which makes $(\mathbf{R}, d_{\mathbf{R}})$ a metric space.

Since real numbers are equivalence classes of Cauchy sequences, any open neighborhood of $[\alpha] \in \mathbf{R}$ will contain the tail of every sequence $(a_n)_n \in [\alpha]$, i.e. every open neighorhood of every point in $\mathbf{R}$ contains at least one rational. So $\mathbf{Q}$ is dense in $\mathbf{R}$.

We must now show that $(\mathbf{R}, d_{\mathbf{R}})$ is complete.

Let $\Xi = (\xi_n)_n$ where $\xi_n = [(x_{(m,n)})_m]$ be a Cauchy sequence of real numbers. Then for every $\varepsilon \in \mathbf{Q}^+$ there is some $N_\varepsilon \in \mathbf{Z}^+$ such that for all $n, n' \geq N_\varepsilon$ there is some $M_{n,n'} \in \mathbf{Z}^+$ such that for all $m \geq M_{n,n'}$,

$$d_{\mathbf{Q}}(x_{(m,n)}, x_{(m,n')}) < \varepsilon.$$

Further, each $\xi_n$ is Cauchy, so for every $n \in \mathbf{Z}^+$ and $\varepsilon \in \mathbf{Q}^+$ there is some $M_{n,\varepsilon} \in \mathbf{Z}^+$ such that for $m, m' \geq M_{n,\varepsilon}$, we have

$$d_{\mathbf{Q}}(x_{(m,n)}, x_{(m',n)}) < \varepsilon.$$

Since we have equivalence classes, we may work with convenient (i.e. rapidly converging) representatives from each class. That is, we may assume that for all $m, m', n \in \mathbf{Z}^+$, we have

$$d_{\mathbf{Q}}(x_{(m,n)}, x_{(m',n)}) < 2^{-\min(\{m,m'\})}.$$

Let $\varepsilon \in \mathbf{Q}^+$, and pick $N = \max(\{N_{\varepsilon/2}, \lceil 1 - \log_2(\varepsilon) \rceil\})$. so that

$$2^{-N} < 2^{-(1-\log_2(\varepsilon))} = \varepsilon/2.$$

Then

$$d_{\mathbf{Q}}(x_{(n,n)}, x_{(n',n')}) \leq d_{\mathbf{Q}}(x_{(n,n)}, x_{(n,n')}) + d_{\mathbf{Q}}(x_{(n,n')}, x_{(n',n')}) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

so $\xi_\infty \in \mathbf{R}$. Further,

$$d_{\mathbf{Q}}(x_{(m,n)}, x_{(n,n)}) \leq d_{\mathbf{Q}}(x_{(m,n)}, x_{(m,m)}) + d_{\mathbf{Q}}(x_{(m,m)}, x_{(n,n)}) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

so $\xi_n \to \xi_\infty$. This shows that $\mathbf{R}$ is complete, which ends the proof. ∎

## 4.8 CONTRACTION MAPPINGS

**Definition 4.18.** Let $(X, d)$ be a metric space, $f : X \to X$.

We say $f$ forms a **contraction mapping** if there is some $\alpha \in [0, 1)$ such that

$$d(f(x), f(y)) \leq \alpha d(x, y)$$

for all $x, y \in X$.

**Theorem 4.19** (Banach fixed-point theorem)**.** *Let* $(X, d)$ *be a nonempty complete metric space with a contraction mapping* $f : X \to X$*. Then* $f$ *admits a unique fixed point* $x^* \in X$ *so that* $f(x^*) = x^*$*.*

Proof. ∎

## 4.9 EXERCISES AND REMARKS

1.

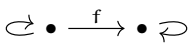
I first learned the "family of rulers" explanation from Dan Piponi on Math SE.

There is also Cat, the category of small categories,[2] wherein the arrows are **functors**. A functor is a map from one category to another that respects both the object structure and the arrow structure of each category. For example, the operation $(\cdot)^{\mathrm{op}}$ takes a category $\mathcal{C}$ and swaps the sources and targets of each arrow, creating the **opposite category** $\mathcal{C}^{\mathrm{op}}$. When this operation respects arrows, it is also a functor.

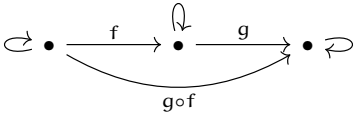
Nonnegative integers form small categories (i.e. elements of Cat), denoted in boldface.

The category **0** is empty, and **1** is the category with just one object and its identity,

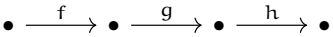
The category **2** has two objects, totally ordered with compositions:




The category **3** has three objects, totally ordered with compositions:




For clarity, we often omit compositions and identity arrows, so that **4** looks like




There are functors that connect these nonnegative integer categories: $\mathsf{Cat}_1(\mathbf{k}, \mathbf{n})$ has size $\binom{n}{k}$.

Note that while there is no natural[3] way to identify $\mathbf{n}$ with $\mathbf{n}^{\mathrm{op}}$, there *is* a natural way[4] to identify $\mathbf{n}$ with $(\mathbf{n}^{\mathrm{op}})^{\mathrm{op}}$.

Continuity is a functor from $\mathsf{Top}^{\mathrm{op}}$ to Poset.

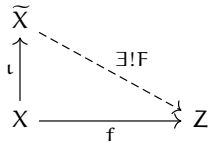
[2] A category is **small** when both $\mathcal{C}_0$ and $\mathcal{C}_1$ are both sets.

[3] The operation $\mathbf{n} \mapsto \mathbf{n}^{\mathrm{op}}$ in this case is not a functor, since it reverses order (functors between nonnegative integer categories must preserve order).

[4] This parallels the fact that there is no natural map from a finite dimensional vector space to its dual, there is a natural map from any finite dimensional vector space to its double dual.

The category Met of metric spaces has as its arrows the 1-Lipschitz maps, which makes the isomorphisms in Met the isometries.

Every metric space $X$ has a completion $\widetilde{X}$, and this completion is unique up to unique isometry:




That is, completions of metric spaces satisfy a **universal property**, and the completion functor from Met to CompMet is in a precise sense complementary to the inclusion functor going from CompMet into Met.[5]

The category Met is enriched over itself: each hom-set $\mathsf{Met}_1(a, b)$ can be thought of as a metric space under the sup metric.

[5] We say that the two functors form an **adjunction**.

# Chapter 5

# FORMS OF COMPLETENESS

The philosophy of this chapter stands opposite to that of how real analysis courses are usually taught, wherein the results that follow are presented in a disjoint manner, or otherwise ignored entirely. We present five equivalent forms of the completeness property of **R**. The point is to stress that when one can avail oneself of one of these, one can avail oneself of any of them, due to logical equivalence.

## 5.0 CAUCHY COMPLETENESS + ARCHIMEDEAN PROPERTY

As one might expect, the Cauchy contruction of the reals makes it easy to show Cauchy completeness. The Archimedean property is then a consequence of well-ordering.

**Theorem 5.1.** *Every Cauchy sequence of real numbers converges to some real limit.*

**Axiom 5.2** (Well-Ordering Principle)**.** *A nonempty set of positive integers has a minimum.*

**Corollary 5.3.** *There is no largest positive integer, i.e.* $\mathbf{Z}^+$ *is unbounded.*

Proof. Suppose there were some positive integer $N_\infty$ that was greater than or equal to all other positive integers. Consider $N_\infty - 1$. If there were some positive integer $M$ greater than $N_\infty - 1$, then $M + 1 > N_\infty - 1 + 1 = N_\infty$, contradiction, so $N_\infty - 1$ must also be greater than or equal to all other positive integers. But there was

nothing particularly special about $N_\infty - 1$, we could apply the same logic to $N_\infty - 2$, or $N_\infty - 3$, etc. In particular, we have produced a nonempty set of positive integers with no least element, contradiction. ∎

**Theorem 5.4** (Archimedean Property)**.** *Let* $\alpha, \beta \in \mathbf{R}^+$*. Then there is* $N \in \mathbf{Z}^+$ *such that*

$$N\alpha > \beta.$$

Proof. Suppose no such N exists. Then for every positive integer N, we have $N\alpha \leq \beta$, i.e. $\beta/\alpha$ is an upper bound for the positive integers. But $\mathbf{Z}^+$ is unbounded, so N must exist. ∎

## 5.1 SUPREMA AND INFIMA

This section covers the least upper bound property, which is the most common starting point for proving the results in this writeup. Many textbooks take the least upper bound property as an axiom; we will show it is equivalent to Cauchy completeness with the Archimedean property. Before we state the least upper bound property we must first define and characterize suprema and infima.

**Proposition 5.5.** *Let* $S \subseteq \mathbf{R}$ *be nonempty.*

*Then* $M = \sup S$ *if and only if* $M$ *is an upper bound of* $S$ *and for every* $\varepsilon > 0$ *there is some* $s \in S$ *such that* $M < s + \varepsilon$*.*

*Similarly,* $m = \inf S$ *if and only if* $m$ *is a lower bound of* $S$ *and for every* $\varepsilon > 0$ *there is some* $s \in S$ *such that* $s - \varepsilon < m$*.*

Proof. We'll prove the supremum version (the infimum version is nearly identical).

Suppose that for every upper bound $M'$ of S we have $M \leq M'$. Let $\varepsilon > 0$, and consider the interval $[M - \varepsilon/2, M]$. If there are no elements of S within this interval, then $M - \varepsilon/2$ is an upper bound of S, but this cannot be, since $M - \varepsilon/2 < M$. So pick some $s \in [M - \varepsilon/2, M]$. Then $M < s + \varepsilon$.

Conversely, suppose for every $\varepsilon > 0$ there is some $s \in S$ such that $M < s + \varepsilon$. Let $M'$ be an upper bound of S, and suppose $M > M'$. Then $M - M' > 0$, so there is some $s \in S$ such that $M < s + M - M'$, but this implies that $M' < s$ for some $s \in S$,

which contradicts the assumption that $M'$ was an upper bound of $S$. So we must have $M \leq M'$. ∎

Now let $S$ denote some nonempty set of real numbers.

**Theorem 5.6** (Least Upper Bound Property)**.** *If $S$ is bounded above, then* $\sup S$ *exists.*

This technique of proof will be used to prove the Theorem 2.3 under various assumptions.

Proof. (Kinda sketchy.) Take $s_0 \in S$ and $M_0$ an upper bound of $S$, and consider $x_0 = \frac{1}{2}(s_0 + M_0)$. If $x_0 \in S$, then set $s_1 = x_0, M_1 = M_0$. Continue on in this way, letting $x_1 = \frac{1}{2}(s_1 + M_1)$ and so on, and note that $y_i = s_i - M_i = 2^{-i}(s_0 - M_0)$ is a Cauchy sequence. By the archimedean property the sequence tends to 0, and since we are assuming that every Cauchy sequence converges to some real limit, $y_i \to 0$. That is, $s_i - M_i \to 0$.

Let $\alpha = \lim_i M_i = \lim_i s_i$. So if some $s \in S$ were such that $\alpha < s$, then $\lim_i M_i < s$, which contradicts the fact that the $M_i$ are all upper bounds of $S$. So $\alpha$ is an upper bound of $S$. Now let $\alpha'$ be another upper bound of $S$, and suppose $\alpha' < \alpha$. Then $\alpha' < \lim_i s_i$, so some element of $s_i$ is greater than $\alpha'$, contradicting the assumption that $\alpha'$ is an upper bound of $S$. So $\alpha \leq \alpha'$, i.e. $\alpha = \sup S$. ∎

This is actually equivalent to the following:

**Corollary 5.7** (Greatest Lower Bound Property)**.** *If $S$ is bounded below, then* $\inf S$ *exists.*

Proof. Consider the set $S'$ of lower bounds of $S$. It is nonempty and bounded above by some element of $S$, so $S'$ has a least upper bound $\alpha$. Since every upper bound of $S'$ is greater than or equal to $\alpha$, and every element of $S$ is an upper bound of $S'$, it follows that every element of $S$ is greater than $\alpha$, i.e. $\alpha$ is a lower bound of $S$. Now let $\alpha'$ be another lower bound of $S$ and suppose $\alpha < \alpha'$. Then since $\alpha' \in S'$, $\alpha$ is not an upper bound of $S'$, contradicting the assumtion that $\alpha$ was the least upper bound of $S'$. So we must have $\alpha' \leq \alpha$, i.e. $\alpha = \inf S$. ∎

Collecting the two results, we have:

**Theorem 5.8.** *Let $S \subset \mathbf{R}$ be nonempty and bounded. Then both* $\inf S$ *and* $\sup S$ *exist.*

## 5.2 DEMONSTRATING EQUIVALENCE

Theorem 2.3 is implied by Theorems 1.1 and 1.4 and implies Theorem 2.5. We now need to show that Theorem 2.5 implies Theorems 1.1 and 1.4.

**Theorem 5.9.** *Theorem 2.5 implies both Theorem 1.1 and Theorem 1.4.*

PROOF. Let $(x_n)_n$ be a Cauchy sequence of real numbers. Then $(x_n)_n$ is bounded, so $\sup_n x_n$ and $\inf_n x_n$ both exist. Extending this observation to $(x_{n+m})_n$ for some $m \in \mathbf{Z}^+$, $\sup_n x_{n+m} = \alpha_m$ and $\inf_n x_{n+m} = \beta_m$ all exist. Further, the sequences $(\alpha_m)_m$ and $(\beta_m)_m$ are nonempty bounded sets, so $\inf_m \alpha_m$ and $\sup_m \beta_m$ both exist. (To be transparent, we note that we are basically talking about liminf and limsup.) We claim that $\inf_m \alpha_m = \sup_m \beta_m$. Suppose they differed by $\varepsilon_0$. Then for every $N \in \mathbf{Z}^+$ we are then able to find $m, n \geq N$ such that $|x_m - x_n| \geq \varepsilon_0$, contradicting our assumption that $(x_n)_n$ was Cauchy. So $\inf_m \alpha_m = \lim_m \alpha_m = \lim_m \beta_m = \sup_m \beta_m$, and all of these equal $\lim_m x_m$ by the squeeze theorem. This proves Theorem 1.1.

Note that $\frac{1}{n} \to 0$: $(\frac{1}{n})_n$ is a bounded below by 0 strictly decreasing sequence of real numbers, and so converges to 0. Since $1/n$ gets arbitrarily small, $n$ gets arbitrarily large, showing that $\mathbf{Z}^+$ is unbounded. This proves Theorem 1.4. ∎

## 5.3 MONOTONE CONVERGENCE THEOREM

We will show that the following result is equivalent to the least upper bound property.

**Theorem 5.10** (Monotone Convergence Theorem)**.** *Let $(x_n)_n$ be a strictly decreasing bounded below sequence of real numbers (or, equivalently, a strictly increasing bounded above sequnence of real numbers). Then $(x_n)_n$ converges to some real limit.*

PROOF. Let $(x_n)_n$ be a strictly increasing bounded above sequence, and let $S = \{x_n : n \in \mathbf{Z}^+\}$. Then $S$ is a nonempty bounded above subset of real numbers. By Theorem 2.5, $S$ has a least upper bound $\sup S$. We claim that $x_n \to \sup S$. Let $\varepsilon > 0$. Then there is some $N \in \mathbf{Z}^+$ such that for all $n \geq N$, we have $x_n + \varepsilon > \sup S$. But since $\sup S$ is an upper bound of $S$, we have

$$|\sup S - x_n| = \sup S - x_n < \varepsilon.$$

The case where $(x_n)_n$ is strictly decreasing and bounded below is similar. ∎

**Theorem 5.11.** *Theorem 3.1 implies Theorem 2.5.*

Proof. Let S be a nonempty bounded above subset of $\mathbf{R}$. Pick $s_0 \in S$, and $M_0$ some upper bound of S, and consider $x_0 = \frac{1}{2}(s_0 + M_0)$. If $x_0 \in S$ then set $s_1 = x_0, M_1 = M_0$. Otherwise set $s_1 = s_0, M_1 = x_0$. Continue choosing $s_i$ and $M_i$ in this way so that $s_i$ is always in S and $M_i$ is always an upper bound of S. By construction, $s_i - M_i \leq 0$ for all $i$, and further $s_i - M_i = 2^{-i}(s_0 - M_0)$, i.e. $s_i - M_i$ is strictly increasing. By Theorem 3.1 $s_i - M_i$ converges to some real limit, and it is clear that this limit must be 0.

From here we set $\alpha = \lim_i s_i = \lim_i M_i$ and continue as in the proof of Theorem 2.3. ∎

## 5.4 BOLZANO-WEIERSTRASS THEOREM

**Theorem 5.12** (Bolzano-Weierstrass)**.** *Every bounded sequence has a convergent subsequence.*

**Theorem 5.13.** *Theorem 4.1 implies Theorem 3.1.*

## 5.5 INTERMEDIATE VALUE THEOREM

Finally, we show that the following result is equivalent to Theorem 4.1.

**Theorem 5.14** (Intermediate Value Theorem)**.** *Let* $f : \mathbf{R} \to \mathbf{R}$ *be a continuous function such that* $f(a) < 0$ *and* $f(b) > 0$ *where* $a < b$*. Then some* $c \in [a, b]$ *is a root of* f*.*

**Theorem 5.15.** *Theorem 5.1 implies Theorem 4.1.*

## 5.6 COMPLETENESS AS A PROPERTY

Just as metric spaces can be completed, one may also complete normed spaces and inner product spaces.

## 5.7 EXERCISES AND REMARKS

1. (A standard problem in an intro analysis course.) Consider the sequence

$$x_0 = 1, \qquad x_n = \frac{3(2x_{n-1}+1)}{x_{n-1}+5} \text{ for } n > 0.$$

   Show by induction that the sequence is monotone increasing, then show by induction that the sequence is bounded. Use the monotone convergence theorem to conclude that the sequence converges, and find its limit.

2. Now try $x_0 = 10$, $x_0 = 1000$, $x_0 = 0.001$, $x_0 = -10$.

   (Use a calculator or computer if you'd like.)

   Note that the limit found in the previous problem is a point of *stable equilibrium*. There is also a point of *unstable equilibrium* to this sequence as a function of its initial value – can you find it?

3. Prove that a continuous injective function $f : \mathbf{R} \to \mathbf{R}$ is monotone.

See the paper "Real Analysis in Reverse."

# Chapter 6

# FINITE DIMENSIONAL VECTOR SPACES

Here we cover basic linear algebra,

wherein the basic objects of study are *vector spaces*.

**Definition 6.1.** Let $k$ be either **R** or **C**. A **vector space** over $k$ is a set $X$ with

- a neutral element $0$ (called the *origin*)
- a function $+ : X^2 \to X$ given by $+(x, y) = x + y$ (called *vector addition*)
- a function $\sigma : k \times X \to X$ given by $\sigma(\alpha, x) = \alpha x$ (called *scalar multiplication*)

for which the following hold:

a1. $\forall x, y, z \in X$ we have $x + (y + z) = (x + y) + z$

a2. $\forall x \in X$ we have $0 + x = x$

a3. $\forall x \in X$ there is a unique $-x \in X$ such that $x + (-x) = 0$

a4. $\forall x, y \in X$ we have $x + y = y + x$

m1. $\forall x \in X$ we have $1x = x$

m2. $\forall \alpha, \beta \in k$ and $x \in X$ we have $(\alpha\beta)x = \alpha(\beta x)$

d. $\forall \alpha, \beta \in k$ and $x, y \in X$ we have $(\alpha + \beta)(x + y) = \alpha x + \alpha y + \beta x + \beta y$

That is, a vector space is an *additive group* with a compatible *scaling action*.

In this chapter we study vector spaces of finite dimension over a field $k$,

which will always be either $\mathbf{R}$ or $\mathbf{C}$.

## 6.0 SUBSPACES AND SPAN

**Definition 6.2.** Let $X$ be a vector space. A subset $Y \subseteq X$ is a **subspace** of $X$ if for any $(x_n)_{n=1}^N \subset Y$ we have that if $(\alpha_n)_{n=1}^N \subset k$ then $\sum_{n=1}^N \alpha_n x_n \in Y$.

That is, subspaces are stable under linear combinations.

Every subspace comes with an inclusion map $\iota$.

## 6.1 LINEAR INDEPENDENCE AND BASES

**Proposition 6.3** (Steinitz Exchange Lemma)**.** *Let $X$ be a vector space with*

$$Y = (y_m)_{m=1}^M \subseteq X \quad \textit{and} \quad Z = (z_n)_{n=1}^N \subseteq X$$

*such that $Y$ is linearly independent and $Z$ spans $X$.*

*Then $M \leq N$, and there is some $W \subset Z$ with $|W| = N - M$ such that $Y \cup W$ spans $X$.*

Proof. By induction. ∎

## 6.2 SUMS AND QUOTIENTS

**Definition 6.4.** Sum, direct sum of subspaces

**Definition 6.5.** quotient wrt a subspace

## 6.3 LINEAR MAPS

**Definition 6.6.** A function

$$T : X \to Y$$

is **linear** if

$$T\left(\sum_{n=1}^{N} \alpha_n x_n\right) = \sum_{n=1}^{N} \alpha_n T(x_n)$$

where $(\alpha_n)_{n=1}^N \subset k$ and $(x_n)_{n=1}^N \subset X$.

**Proposition 6.7.** *The set of linear maps from $X$ to $Y$ forms a vector space, denoted*

$$\mathcal{L}(X, Y).$$

**Proposition 6.8** (Subkernel property)**.** *If a linear map $T : X \to Y$ sends a subspace $Z$ of $X$ to zero, then there is a unique surjective linear map $\widetilde{T} : X/Z \to Y$.*

**Theorem 6.9** (Canonical decomposition)**.** *Every linear map decomposes as an injection followed by a bijection followed by a surjection.*

**Corollary 6.10** (Rank-nullity theorem)**.**

## 6.4 INJECTIVITY AND SURJECTIVITY IN FINITE DIMENSION

Let $X$ be a vector space with $0 < N = \dim X < \infty$,[1] and let $T : X \to X$.

**Theorem 6.11.** *$T$ is injective if and only if $T$ is surjective.*

Proof. Suppose $T$ is injective, i.e. suppose $\ker T = 0$. Let $x \in X$.

Consider the vectors $(T^n x)_{n=0}^N$. Since $\dim X = N$, this is a linearly dependent set,

[1] If $X = \{0\}$, the only map from $X$ to $X$ is the zero map, which sends 0 to 0 and is trivially bijective.

i.e. there exists a nonzero $(\alpha_n)_{n=0}^N$ such that $\sum_{n=0}^N \alpha_n T^n x = 0$.

In the case $\alpha_0 \neq 0$, we may divide by $\alpha_0$ and use linearity of $T$ to obtain:

$$x = T\left(-\frac{1}{\alpha_0}\sum_{n=1}^{N} \alpha_n T^{n-1} x\right).$$

In the case $\alpha_0 = 0$, we have $T\left(\sum_{n=1}^N \alpha_n T^{n-1} x\right) = \sum_{n=1}^N \alpha_n T^n x = 0$, i.e.

$$\sum_{n=1}^{N} \alpha_n T^{n-1} x \in \ker T.$$

But since $T$ is injective, this means $\sum_{n=1}^N \alpha_n T^{n-1} x = 0$, and so we're back where we started: we eventually encounter a nonzero $\alpha_n$, which we may divide by, and then use linearity of $T$ as before.

This proves $T$ is surjective.

Conversely, suppose $T$ is surjective, i.e. suppose $T(X) = X$. Let $x \in \ker T$.

Then $\mathrm{Span}(x) \subset \ker T$, so there exists a unique surjective linear map

$$\widetilde{T} : X/\mathrm{Span}(x) \to X,$$

which implies $\dim(X/\mathrm{Span}(x)) \geq N$. But we also have

$$\dim(X/\mathrm{Span}(x)) = N - \dim(\mathrm{Span}(x)) \leq N,$$

so $X/\mathrm{Span}(x)$ has dimension $N$. This implies $\mathrm{Span}(x) = 0$, hence $x = 0$.

This proves $T$ is injective. ∎

## 6.5 DUALS AND MULTILINEARITY

**Definition 6.12.** Let $\mathcal{L}(X_1, \ldots, X_N; Y)$ denote the space of **multilinear maps** from $\prod_{n=1}^N X_n$ to $Y$.

Define $\bigotimes_{n=1}^N X_n = \mathcal{L}(X_1', \ldots, X_N'; k)$.

## 6.6 EXERCISES AND REMARKS

These exercises will help you get aquainted with actually computing things about linear maps, using the language of *matrices*. Recall that we defined $n$ to be the set

$$\{0, \ldots, n-1\}$$

so that $|n| = n$.

**Definition 6.13.** An $m \times n$ **matrix** over a ring $R$ is a function $A : m \times n \to R$.

We write $A(i,j)$ as $a_{ij}$, and write the entire matrix as e.g.

$$\begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}$$

Matrices are a convenient way to write down the coefficients associated with a linear map – a representation.

1. **Classical logic gates.** Consider the field $\mathbf{F}_2$. Let

$$\bot = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad \top = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Then the identity gate is simply the matrix corresponding to the identity map:

$$G_{\text{id}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The negation gate is then

$$G_{\text{not}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

We also have the constant true and false gates

$$G_{\text{bot}} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad \text{and} \qquad G_{\text{top}} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

The binary operation "and" is then

$$G_{\text{and}} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

so that for example $\bot \otimes \bot = (1,0) \otimes (1,0) = (1,0,0,0)$

$$G_{\text{and}}(\top \otimes \top) = G_{\text{and}}(0,0,0,1) = (0,1) = \top.$$

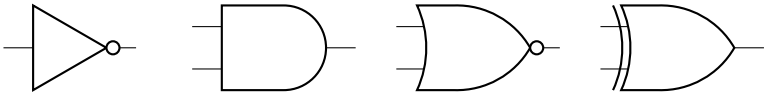
The binary operation "or (possibly both)" is then

$$G_{\text{or}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

so that

$$G_{\text{or}}(\top \otimes \bot) = G_{\text{or}}(0,0,1,0) = (0,1) = \top.$$

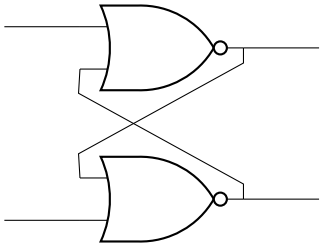
In this way, we may implement all classical logic gates:



Shown here are the NOT, AND, NOR, and XOR gates.

2. **Classical logic and computer memory.**

   Consider the simplified SR-latch:



   It consists of two NOR gates that feed back into each other.

   "SR" stands for "Set-Reset."

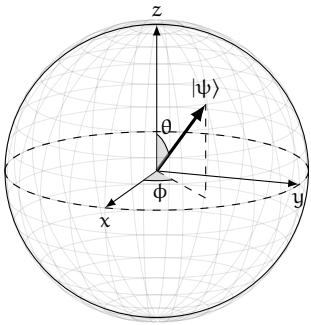
   This is the most basic form of computer memory.

   It holds one bit, plus its negation.

3. **Quantum logic gates: one qubit.** Instead of bits (binary digits), we now work with qubits, where a qubit is a complex linear combination

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \qquad |\alpha|^2 + |\beta|^2 = 1.$$

Defining $\alpha = \cos(\theta/2)$ and $\beta = e^{i\phi}\sin(\theta/2)$, we have




where $|\psi\rangle$ is represented by the point $(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$.

Behold the Pauli spin gates:

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

we also have the Hadamard gate and the S and T gates:

$$G_H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad G_S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}, \quad G_T = \begin{bmatrix} 1 & 0 \\ 0 & e^{\pi i/4} \end{bmatrix}$$

4. **Quantum logic gates: two or more qubits.** For two qubits, we write

$$\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$

where $|00\rangle$ is understood to mean $|0\rangle \otimes |0\rangle$, etc.

Here is a CNOT gate:

$$G_{\text{CNOT}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Here is a CZ gate:

$$G_{CZ} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Here is a SWAP gate:

$$G_{SWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

And finally the Toffoli aka CCNOT gate:

$$G_{\text{Toffoli}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

5. Constructing Bell states.

Vector spaces and linear maps form a category, $\mathsf{Vect}_k$.

Also of note is the subcategory of finite dimensional vector spaces, $\mathsf{FinVect}_k$.

We'll let $\Xi(X)$ denote the set of subspaces of $X$, ordered by inclusion. We have that $\Xi$ is actually a complete lattice: intersections are meets, sums are joins, the 0 subspace is the bottom element, and $X$ is the top element.

A corollary of the first isomorphism theorem is that

$$\Xi^{op}(X)$$

is actually the lattice of quotient spaces of $X$.

Thank you to tobiasBora on TeX StackExchange for the bloch sphere code.

# Chapter 7

# NORMS AND INNER PRODUCTS

A vector is commonly known as "a quantity with both magnitude and direction." Norms are a way to talk about lengths, inner products are a way to talk about angles.

**Definition 7.1.** Let $X$ be a vector space over a field $k$.

A **norm** on $X$ is a function $\|\cdot\| : X \to \mathbf{R}$ satisfying[1]

1. Positivity: $\|x\| > 0$ for all nonzero $x \in X$
2. Absolute homogeneity: $\|\alpha x\| = |\alpha| \|x\|$ for all $x \in X$, $\alpha \in k$
3. Subadditivity: $\left\|\sum_{n=1}^{N} x_n\right\| \leq \sum_{n=1}^{N} \|x_n\|$ for all $(x_n)_{n=1}^{N} \subset X$

We call $(X, \|\cdot\|)$ a **normed space**. If $X$ is complete, we call $X$ a **Banach space**.

**Proposition 7.2.** *In a normed space $(X, \|\cdot\|)$, the norm $\|\cdot\|$ is 1-Lipschitz.*

Proof. As a consequence of subadditivity, we have the reverse triangle inequality:

$$\big|\|x\| - \|y\|\big| \leq \|x - y\| \quad \forall x, y \in X.$$

This implies the result. ∎

[1] We note that $\|0\| = \|0x\| = |0| \|x\| = 0 \|x\| = 0$, so the claim

$$(x = 0) = (\|x\| = 0)$$

(a strengthening of positivity) holds. Also, $\|x\| \geq 0$ for all $x \in X$.

**Definition 7.3.** Let $X$ be a vector space over a field $k$.

An **inner product** on $X$ is a function $\langle \cdot, \cdot \rangle : X^2 \to k$ satisfying:

1. $\langle x, x \rangle > 0$ for all nonzero $x \in X$,
2. $\langle x, y \rangle = \overline{\langle y, x \rangle}$,
3. $\left\langle \sum_{n=1}^{N} \alpha_n x_n, y \right\rangle = \sum_{n=1}^{N} \alpha_n \langle x_n, y \rangle$.

## 7.0 THE CAUCHY-SCHWARZ INEQUALITY

**Theorem 7.4** (Cauchy-Schwarz)**.** *For $X$ an inner product space, $x, y \in X$:*

$$|\langle x, y \rangle| \leq ||x|| \, ||y||$$

*with equality when $x$ and $y$ are linearly dependent.*

Proof. (T. Brimnathwala, 2021)

Define

$$u = \frac{x}{||x||}, \quad v = \frac{y}{||y||}, \quad \gamma = \frac{\langle u, v \rangle}{|\langle u, v \rangle|}$$

so that if any of the denominators vanish then the result immediately follows.

Now let

$$p = \frac{u + \gamma v}{2}, \quad q = \frac{u - \gamma v}{2}.$$

One may check that $\langle p, q \rangle = 0$ and that $||p||^2 + ||q||^2 = 1$ (Exercise 1). These imply

$$|\langle u, v \rangle| = |1 - 2||q||^2| \leq 1 \tag{7.1}$$

with equality holding when either $||q||^2 = 0$ or $||q||^2 = 1$,

which is exactly when $u = \pm v$, or when $x = \alpha y$ for some $\alpha \in k$.

Multiply both sides of (7.1) by $||x|| \, ||y||$ to obtain the result. ∎

## 7.1 COMPACTNESS OF THE UNIT SPHERE

A normed space is finite dimensional if and only if the unit sphere is compact.

## 7.2 CONTINUOUS LINEAR MAPS

Recall that, for vector spaces $X$ and $Y$,

$\mathcal{L}(X,Y)$ denotes the vector space of linear maps from $X$ to $Y$.

Now if $X$ and $Y$ are normed spaces, we denote by $\mathcal{B}(X,Y)$

the normed space of continuous linear maps from $X$ to $Y$.

**Theorem 7.5.** *Let* $T \in \mathcal{L}(X,Y)$. *Then*

$$T \text{ is continuous} \quad = \quad T \text{ is continuous at } 0 \quad = \quad T \text{ is bounded.}$$

## 7.3 C*-ALGEBRAS

$\mathcal{B}(X)$ for $X$ a Hilbert space forms a C*-algebra.

$\mathcal{K}(X)$ for $X$ a Hilbert space forms a C*-algebra.

## 7.4 NEAREST POINT PROPERTY

**Definition 7.6.** Let $X$ be a vector space.

A **convex combination** is a linear combination $\sum_{n=1}^{N} \alpha_n x_n$

where each $\alpha_n \geq 0$ and $\sum_{n=1}^{N} \alpha_n = 1$.

A subset $S \subseteq X$ is **convex** if it is stable under convex combinations.

**Theorem 7.7** (Nearest point property)**.** *Let* $X$ *be a Hilbert space,*

*and let* $K$ *be a nonempty closed convex subset of* $X$.

*Then* $K$ *contains an element of smallest norm.*

*Moreover, given any* $x \in X$, *there is a unique* $k_0 \in K$ *such that*

$$\|x - k_0\| = \inf\{\|x - k\| : k \in K\}.$$

Proof. (Riesz) Uses the parallelogram law. ■

## 7.5 THE PROJECTION THEOREM

**Theorem 7.8** (Projection theorem)**.** *Let* $(X, \langle\cdot,\cdot\rangle)$ *be an inner product space,* $Z \subseteq X$ *a nonempty convex complete subset.*

1. *For any* $x \in X$ *there is a unique* $P(x) \in Z$ *such that*

$$\|x - P(x)\| = \inf_{z \in Z} \|x - z\|.$$

2. *The unique element* $P(x)$ *satisfies*

$$\mathrm{Re}\langle P(x) - x, z - P(x)\rangle \geq 0 \qquad \forall z \in Z.$$

   *Conversely, if* $y \in Z$ *satisfies*

$$\mathrm{Re}\langle y - x, z - y\rangle \geq 0 \qquad \forall z \in Z,$$

   *then* $y = P(x)$.

3. $P$ *is 1-Lipschitz.*

4. *Suppose* $Z$ *is a complete subspace of* $X$*. Then* $P(x)$ *satisfies*

$$\langle P(x) - x, z\rangle = 0 \qquad \forall z \in Z.$$

   *Conversely, if* $y \in Z$ *satisfies*

$$\langle y - x, z\rangle = 0 \qquad \forall z \in Z,$$

   *then* $y = P(x)$.

5. *The mapping* $P : X \to Z$ *is linear if and only if* $Z$ *is a subspace of* $X$*. If* $Z \neq \{0\}$ *then* $\|P\| = 1$.

## 7.6 ORTHOGONALITY AND DIRECT SUMS

**Proposition 7.9.** *For $X$ a Hilbert space and $Y$ a closed subspace of $X$,*

$$X = Y \oplus Y^{\perp}.$$

## 7.7 RIESZ REPRESENTATION THEOREM

**Theorem 7.10** (Riesz representation theorem)**.** *Let $X$ be a Hilbert space. Given a continuous linear function $\ell \in X'$, there exists one and only one vector $y_\ell \in X$ such that*

$$\ell(x) = \langle x, y_\ell \rangle \quad \forall x \in X.$$

## 7.8 HAHN-BANACH THEOREM

**Theorem 7.11** (Hahn-Banach)**.** *Let $X$ be a Hilbert space, $Y$ a subspace of $X$, $\ell \in Y'$. There is a unique $L \in X'$ such that $L(y) = \ell(y)$ for all $y \in Y$ and $\|L\|_{X'} = \|\ell\|_{Y'}$.*

## 7.9 EXERCISES AND REMARKS

1. In the proof of Cauchy-Schwarz, check that $\langle p, q \rangle = 0$ and $\|p\|^2 + \|q\|^2 = 1$.
2. *Bra-ket notation.* Instead of the mathematical

   $$\langle x, y \rangle$$

   where the conjugate-linearity is agreed to go in the *second* argument, physicists notate the inner product as

   $$\langle x \mid y \rangle$$

   where the conjugate-linearity is now agreed to go in the *first* argument.
3. *The Heisenberg Uncertainty Principle.* Let $\Omega$ and $\Lambda$ be Hermitian with commutator $[\Omega, \Lambda] = i\hbar$.

   $$(\Delta\Omega)(\Delta\Lambda) \geq \frac{\hbar}{2}.$$

Proof.

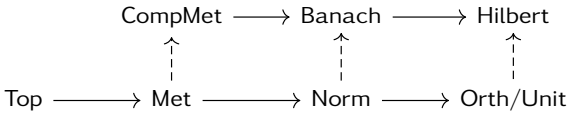
$$\begin{aligned}
(\Delta\Omega)^2(\Delta\Lambda)^2 &= \langle\psi \mid (\Omega - \langle\Omega\rangle)^2 \mid \psi\rangle\langle\psi \mid (\Lambda - \langle\Lambda\rangle)^2 \mid \psi\rangle \\
&= \langle\psi \mid \hat{\Omega}^2 \mid \psi\rangle\langle\psi \mid \hat{\Lambda}^2 \mid \psi\rangle \\
&= \langle\psi \mid \hat{\Omega}^\dagger\hat{\Omega} \mid \psi\rangle\langle\psi \mid \hat{\Lambda}^\dagger\hat{\Lambda} \mid \psi\rangle \\
&= \langle\hat{\Omega}\psi|\hat{\Omega}\psi\rangle\langle\hat{\Lambda}\psi|\hat{\Lambda}\psi\rangle \\
&= \|\Omega\psi\|^2\|\Lambda\psi\|^2 \\
&\geq |\langle\hat{\Omega}\psi \mid \hat{\Lambda}\psi\rangle|^2 \\
&= |\langle\psi \mid \hat{\Omega}\hat{\Lambda} \mid \psi\rangle|^2 \\
&= |\langle\psi \mid \tfrac{1}{2}[\Omega,\Lambda]_+ + \tfrac{1}{2}[\Omega,\Lambda] \mid \psi\rangle|^2 \\
&= |\tfrac{1}{2}\langle\psi \mid [\Omega,\Lambda]_+ \mid \psi\rangle + \tfrac{1}{2}\langle\psi \mid [\Omega,\Lambda] \mid \psi\rangle|^2
\end{aligned}$$

■

In the category Norm of normed spaces, the morphisms are the linear 1-Lipschitz maps, which means the isomorphisms are the linear isometries.

There are two inner product categories, which I'll call Orth and Unit. The objects in the category Orth are the real inner product spaces, whereas the objects in the category Unit are complex inner product spaces. The morphisms in both categories are the conformal linear isometries, which means the isos for Orth are the symmetic operators and the isos for Unit are the unitary operators.

Summarizing, we have the following diagram.




It does **not** commmute, e.g. Met is not "the pullback of Norm and CompMet."

# Chapter 8

# DIFFERENTIATION IN BANACH SPACE

## 8.0 LANDAU NOTATION

Let $X$ be a normed space, and let $U_X$ be an open neighborhood of $0$.

If $\varphi : U_X \to Y$, write:

$\varphi(u) = \varepsilon(u)$ if $\lim_{x\to 0} \|\varphi(x)\| = 0$,

$\varphi(u) = \mathcal{O}(u)$ if $\exists C < \infty$ and $r > 0$ such that $\|\varphi(x)\| \le C\|x\|$, $\forall x \in B_r(0)$,

$\varphi(u) = o(u)$ if $\varphi(u) = \varepsilon(u)\mathcal{O}(u)$, i.e. if $\lim\limits_{x\to 0} \dfrac{\|\varphi(x)\|}{\|x\|} = 0$.

## 8.1 THE FRECHET AND GATEAUX DERIVATIVES

**Definition 8.1.** Let $X$ and $Y$ be normed spaces,

$U_X \subseteq X$ open, $\varphi : U_X \to Y$, and $u \in U$.

If there is $F \in \mathcal{B}(X, Y)$ satisfying

$$\varphi(u + x) - \varphi(u) - Fx = o(x)$$

then $\varphi$ is said to be **(Frechet) differentiable** at $u$, where we write $F = \varphi'_u$.

There is an important distinction to make:

the "Frechet derivative at a point" is $\varphi'_u : X \to Y$, whereas

the "Frechet differential" is the map $\varphi' : U_X \to \mathcal{B}(X,Y)$, whereas

the "Frechet tangent map" is a map $D\varphi : X^2 \to Y^2$ given by

$$(u,x) \mapsto (\varphi(u), \varphi'_u x).$$

So even if $\varphi$ has continuous derivative at each point, it may not follow that $\varphi$ is continuously differentiable. Being continuously differentiable means that $\varphi'$ is continuous as a function of $u$ as well.

**Definition 8.2.** Let $X$ and $Y$ be normed spaces,

$U_X \subseteq X$ open, $\varphi : U_X \to Y$, and $u \in U$.

If there is a $G \in \mathcal{L}(X,Y)$ satisfying

$$\varphi(u + \alpha x) - \varphi(u) - \alpha G x = o(\alpha)$$

for every $x \in X$, then $\varphi$ is said to be **(Gateaux) differentiable** at $u$.

If $\varphi$ is Frechet differentiable at $u$, then $\varphi$ is Gateaux differentiable at $u$, and the two derivatives coincide. So we use the same notation for both.

Caution: Gateaux derivatives need not be bounded.

Note that any normed space $X$ has functions that are Gateaux differentiable at some point without being Frechet differentiable at that point, e.g. $\varphi(x) = \|x\|^2$ around 0.

We can say more in a Hilbert space, where Riesz representation allows us to say

$$\varphi'_u v = \langle \nabla\varphi(u), v \rangle$$

## 8.2 PROPERTIES OF THE DERIVATIVE

Differentiation is linear, product rule, differentiability implies continuity.

**Proposition 8.3.** *Let $X$ and $Y$ be Banach spaces with $U_X \subseteq X$ open, and*

$$\varphi : U_X \to Y$$

*a Frechet differentiable function. Then $\varphi$ is locally Lipschitz, i.e.*

$$\varphi(x) - \varphi(u) = \mathcal{O}(x-u)$$

*i.e. at any point $u \in U_X$.*

Proof. Let $u \in U_X$. Then by definition of the Frechet derivative,[1]

$$\varphi(x) - \varphi(u) = \varphi'_u(x-u) + o(x-u) = \mathcal{O}(x-u),$$

for any $x \in U_X$, as desired. ∎

## 8.3 THE CHAIN RULE

Let $X$, $Y$, and $Z$ be normed spaces with $U_X \subseteq X$ open and $U_Y \subseteq Y$ open,

and let $\phi : U_X \to Y$ and $\psi : U_Y \to Z$ be differentiable.

Then there exists $\psi \circ \phi : U_X \to Z$ whose derivative is given by

$$(\psi \circ \phi)'_u = \psi'_{\phi(u)} \circ \phi'_u$$

## 8.4 HIGHER ORDER DERIVATIVES AND TAYLOR'S THEOREM

We have

$$\phi(u+v) = \sum_{k=0}^{n} \frac{1}{k!} \phi_u^{(k)}(v,\dots,v) + R_n(v)$$

where $R_n(v) = o(\|v\|^n)$.

[1] Recall that the Frechet derivative is defined to be bounded.

## 8.5 THE INVERSE AND IMPLICIT FUNCTION THEOREMS

**Theorem 8.4** (Inverse function theorem)**.** *Let* $X$ *and* $Y$ *be Banach spaces and*

$$\varphi : U_X \to Y$$

*a* $C^1$ *function with* $U_X \subseteq X$ *open.*

*Further, assume* $\varphi_0'$ *is a bounded linear isomorphism of* $X$ *onto* $Y$*.*

*Then there exists an open neighborhood* $U_Y$ *of* $\varphi(0)$ *in* $Y$ *and a* $C^1$ *map*

$$\psi : U_Y \to X$$

*such that*

$$\varphi(\psi(y)) = y$$

*for all* $y \in U_Y$*.*

**Theorem 8.5** (Implicit function theorem)**.** *Let* $X$*,* $Y$*,* $Z$ *be Banach spaces with*

$$f : X \times Y \to Z$$

*a* $C^1$ *function. Suppose* $(x_0, y_0) \in X \times Y$ *with* $f(x_0, y_0) = 0$*.*

*If* $\phi : Y \to Z$ *given by* $y \mapsto f'_{(x_0,y_0)}(0, y)$ *is a Banach space isomorphism,*

*then there exist neighborhoods* $U_X$ *of* $x_0$ *and* $U_Y$ *of* $y_0$

*and a differentiable function* $g : U_X \to U_Y$ *such that*

$$f(x, g(x)) = 0 \quad \textit{and} \quad f(x, y) = 0 \quad \textit{iff} \quad y = g(x)$$

*for all* $(x, y) \in U_X \times U_Y$*.*

## 8.6 EXERCISES AND REMARKS

The chain rule as functoriality of the derivative.

The tangent map D given by $\varphi \mapsto D\varphi$ sends each Banach space to its Cartesian square.

For the second derivatives, we have:

$$\varphi'' : U \to \mathcal{B}(V, \mathcal{B}(V,W)) \quad \text{given by} \quad \varphi''(u)(v_1)(v_2) = \varphi''_u(v_1, v_2)$$

and

$$D^2\varphi((u,v),(w_1,w_2)) = ((\varphi(u), \varphi'_u(v)), (\varphi''_u(w_1,v) + \varphi'_u(w_2))$$

# Chapter 9

# MEASURE IN EUCLIDEAN SPACE

*What is mass?* Interestingly enough, measure theory seems to provide a fairly exact characterization of how mass ought to (classically) behave: the mass of a sum of components should be the sum of the masses of each component, mass should be nonnegative and monotone with respect to volume, etc.

Measure theory even provides a means of modeling all the strange things physicists do with mass, such as imagining point masses. One of the bizzare consequences of the Axiom of Choice is that there exist sets with no meaningful conception of mass (called non-measurable sets). Mathematicians, preferring abstraction whenever possible, refer to this generalized mass as *measure*.

## 9.0 BOREL SETS

**Definition 9.1.** Let $X$ be a set. A **$\sigma$-algebra** on $X$ is a subset

$$\Sigma \subseteq \mathcal{P}(X)$$

that is stable under complement and countable union.[1]

[1] Some observations:

Since $\varnothing$ is countable, $\bigcup \varnothing = \varnothing \in \Sigma$, i.e. the empty union (which forms the empty set) is in $\Sigma$.

Since $\Sigma$ is complement stable, $X \setminus \varnothing = X \in \Sigma$ as well.

By de Morgan's laws, $\Sigma$ must also be stable under countable intersection.

**Proposition 9.2.** *Let $X$ be a set. If $(\Sigma_\lambda)_{\lambda\in\Lambda}$ is a collection of $\sigma$-algebras on $X$, then*

$$\bigcap_{\lambda\in\Lambda}\Sigma_\lambda$$

*also forms a $\sigma$-algebra on $X$.*

For any set, there is a smallest $\sigma$-algebra generated by that set:

the intersection of all of the $\sigma$-algebras containing that set.

The **Borel sets** are simply elements of the Borel $\sigma$-algebra,

which is the algebra generated by the open sets.[2]

## 9.1 WHAT IS A MEASURE?

**Definition 9.3.** Let $X$ be a set and $\Sigma$ a $\sigma$-algebra over $X$. A set function

$$\mu : \Sigma \to [0,\infty]$$

is called a **measure** if the following hold:

1. $\mu(\varnothing) = 0$
2. For $(E_n)_{n\geq 1} \subset \Sigma$ pairwise disjoint,

$$\mu\left(\bigcup_{n\geq 1} E_n\right) = \sum_{n\geq 1}\mu(E_n).$$

A Borel measure (in general) is just a measure defined on the Borel sets of $X$. A Radon measure is a special kind of Borel measure; these are required to be locally finite and inner regular. The Dirac and Lebesgue measures are both examples of Radon measures.

This is a complete lattice where $X$ is the top element and $\varnothing$ is the bottom element.

[2] Note that the topology is part of the Borel $\sigma$-algebra.

## 9.2 LEBESGUE MEASURE

In the following let $B_i = B(r_i, x_i)$. These can be balls, boxes, diamonds... whatever you like. All norms in finite dimension are equivalent, after all.

See the exercises for how to actually compute $\mathrm{vol}(B_i)$.

**Definition 9.4.** Lebesgue outer measure is:

$$\lambda^*(E) = \inf\left\{\sum_{i\geq 0} \mathrm{vol}(B_i) : E \subset \bigcup_{i\geq 0} B_i,\ x_i \in E\right\}.$$

That is, $\lambda^*(E)$ is the smallest covering area of any open covering of $E$.

It is then natural to define a set $E$ to be **Lebesgue measurable** if for every $\varepsilon > 0$ there is some open set $U$ such that

$$E \subseteq U \quad \text{and} \quad \lambda^*(U \setminus E) < \varepsilon.$$

and the **Lebesgue measure** of $E$ to be $\lambda(E) = \lambda^*(E)$.

## 9.3 SETS OF MEASURE ZERO

Here we explain "almost everywhere."

A Lipschitz function is differentiable almost everywhere.

## 9.4 PRODUCT MEASURES AND FUBINI'S THEOREM

Let $(X, \Sigma_X)$ and $(Y, \Sigma_Y)$ be measurable spaces.

Define $\Sigma_{X\times Y}$ to be the σ-algebra generated by all sets of the form

$$A \times B$$

where $A \in \Sigma_X$, $B \in \Sigma_Y$.

Then if there exist measures $\mu_X$ on $X$ and $\mu_Y$ on $Y$ such that

$$\mu_{X\times Y}(A \times B) = \mu_X(A)\mu_Y(B),$$

we call $\mu_{X\times Y}$ a **product measure** and

$$(X\times Y,\ \Sigma_{X\times Y},\ \mu_{X\times Y})$$

a **product space** of $(X,\Sigma_X,\mu_X)$ and $(Y,\Sigma_Y,\mu_Y)$.

The product measure construction allows us to go from $[0,1]$ to $[0,1]^n$.

## 9.5 EXERCISES AND REMARKS

1. **The Gaussian integral.** Consider

$$I(\alpha)=\int_{-\infty}^{\infty}\exp(-\alpha x^2)\,dx,\qquad \alpha>0.$$

Write

$$\begin{aligned}I(\alpha)^2&=\left(\int_{-\infty}^{\infty}\exp(-\alpha x^2)\,dx\right)^2\\&=\left(\int_{-\infty}^{\infty}\exp(-\alpha x^2)\,dx\right)\left(\int_{-\infty}^{\infty}\exp(-\alpha x^2)\,dx\right).\end{aligned}$$

Note that the $x$ in $I(\alpha)$ is a dummy variable, so we may write[3]

$$\begin{aligned}I(\alpha)^2&=\left(\int_{-\infty}^{\infty}\exp(-\alpha x^2)\,dx\right)\left(\int_{-\infty}^{\infty}\exp(-\alpha y^2)\,dy\right)\\&=\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\exp(-\alpha x^2)\,dx\right)\exp(-\alpha y^2)\,dy\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(-\alpha x^2)\exp(-\alpha y^2)\,dx\,dy\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(-\alpha(x^2+y^2))\,dx\,dy\\&\overset{\star}{=}\int_0^{2\pi}\int_0^{\infty}\exp(-\alpha r^2)r\,dr\,d\theta.\end{aligned}$$

Since $d(\exp(-\alpha r^2))=-2\alpha r\exp(-\alpha r^2)\,dr$,

$$I(\alpha)^2=\int_0^{2\pi}\frac{1}{-2\alpha}\exp(-\alpha r^2)\Big|_{r=0}^{\infty}d\theta=\frac{\theta}{2\alpha}\Big|_{\theta=0}^{2\pi}=\frac{\pi}{\alpha}.$$

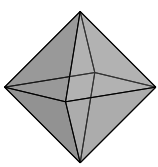
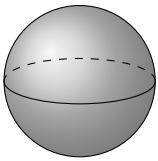
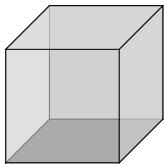
Hence, $I(\alpha)=\sqrt{\frac{\pi}{\alpha}}$.

[3]The starred equality is an important change of variable, see the next chapter for explanation.

2. **The Γ function.**

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$$

3. **Volume of the $n$-diamond, $n$-ball, and $n$-box.**

Previously we stated that it didn't matter in our definition of outer measure whether we covered with balls or boxes. From the standpoint of picking something that works in all dimensions, boxes are the more practical option, because the volume of the $n$-ball actually begins to *decrease* after $n = 5$.

| $n$ | $\frac{2^n}{n!}$ | $\frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$ | $2^n$ |
|---|---|---|---|
| 1 | 2 | 2 | 2 |
| 2 | 2 | $\pi$ | 4 |
| 3 | $\frac{4}{3}$ | $\frac{4\pi}{3} \simeq 4.19$ | 8 |
| 4 | $\frac{2}{3}$ | $\frac{\pi^2}{2} \simeq 4.94$ | 16 |
| 5 | $\frac{4}{15} \simeq 0.26$ | $\frac{8\pi^2}{15} \simeq 5.27$ | 32 |
| 6 | $\frac{4}{45} \simeq 0.08$ | $\frac{\pi^3}{6} \simeq 5.17$ | 64 |
| 7 | $\frac{8}{315} \simeq 0.03$ | $\frac{16\pi^3}{105} \simeq 4.73$ | 128 |
| 8 | $\frac{2}{315} \simeq 0.01$ | $\frac{\pi^4}{24} \simeq 4.06$ | 256 |

The table above has the volumes of the $n$-diamond, $n$-ball, and $n$-box. The box is enlarged by a power of 2, since these are really the unit balls in $\mathbf{R}^n$ with the 1-norm, 2-norm, and $\infty$-norm, respectively.

You may be wondering how we arrived at the formula for the $n$-ball, namely

$$\mathrm{vol}(B_1(0)) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}.$$

To do so, we use Gaussian integrals.

Our treatment of measure is a hybrid of the approach found in Fleming and Stein/Shakarchi.

# Chapter 10

# INTEGRATION IN HILBERT SPACE



$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega.$$



The Bochner integral is a limit of Lebesgue integrals. This integral satisfies a version of the Fundamental Theorem of Calculus on Banach spaces with the Radon-Nikodym Property – these include all reflexive Banach spaces, and in particular all Hilbert spaces. Thus we restrict our attention to integration on $\ell^2$.

Let $(X, \Sigma, \mu)$ be a measure space, $Y$ a Banach space.

Define a **simple function** to be any finite sum of the form

$$s(x) = \sum_{n=0}^{N} \chi_{E_n}(x) y_n$$

where the $E_n$ are disjoint members of the $\sigma$-algebra $\Sigma$, the $y_n$ are distinct elements of $Y$, and $\chi_E$ is the characteristic function of $E$. If $\mu(E_n)$ finite whenever $y_n$ nonzero, then $s$ is **integrable**, and the integral is defined by

$$\int_X \left[ \sum_{n=0}^{N} \chi_{E_n}(x) y_n \right] d\mu = \sum_{n=0}^{N} \mu(E_n) y_n$$

as it is for the Lebesgue integral.

A measurable function $f : X \to Y$ is **Bochner integrable** if there is a sequence of integrable simple functions $(s_n)_{n \geq 1}$ such that

$$\lim_{n \to \infty} \int_X \|f - s_n\|_Y \, d\mu = 0,$$

where the integral on the left is Lebesgue. We then define the **Bochner integral** via

$$\int_X f \, d\mu = \lim_{n \to \infty} \int_X s_n \, d\mu.$$

The sequence $\left( \int_X s_n \, d\mu \right)_{n \geq 1}$ is Cauchy, so the limit exists;

further, it is independent of the approximating sequence.

## 10.9 EXERCISES AND REMARKS

1. The change of variable in the Gaussian integral from the previous chapter can be justified as follows:

   Recall $x = r\cos\theta$, $y = r\sin\theta$.

   Differentiating implicitly,

$$\begin{aligned} dx &= d(r\cos\theta) \\ &= (dr)\cos\theta + r\,d(\cos\theta) \\ &= \cos\theta\,dr - r\sin\theta\,d\theta \end{aligned}$$

$$\begin{aligned} dy &= d(r\sin\theta) \\ &= (dr)\sin\theta + r\,d(\sin\theta) \\ &= \sin\theta\,dr + r\cos\theta\,d\theta. \end{aligned}$$

   Now substitute and multiply:

$$\begin{aligned} dx \wedge dy &= (\cos\theta\,dr - r\sin\theta\,d\theta) \wedge (\sin\theta\,dr + r\cos\theta\,d\theta) \\ &= \cos\theta\sin\theta\,dr \wedge dr + r\cos^2\theta\,dr \wedge d\theta \\ &- r\sin^2\theta\,d\theta \wedge dr - r^2\sin\theta\cos\theta\,d\theta \wedge d\theta \\ &= r(\cos^2\theta + \sin^2\theta)\,dr \wedge d\theta \\ &= r\,dr \wedge d\theta. \end{aligned}$$

   Thus, $dx\,dy = r\,dr\,d\theta$.

2. **Riemann integration in Common Lisp.**

   Lisp (short for "List Processing") is a flexible programming language that, among other things, can do basic calculus. Common Lisp is the most, well, common variant of Lisp. Reading some basic Lisp code implementing calculus can give insight into how calculus itself works.

   Here are some sample functions in Common Lisp, just to get a feel for the notation:

```
(defun sqx (x) (* x x))
(defun cubex (x) (* x x x))
```

The following constants will also be helpful:

```
(defparameter *dx* 0.01)
(defparameter *e* 2.71828)
```

Now, here is some Lisp code that does calculus.

```
(defmacro of (f x) `(apply ,f (list ,x)))

(defun deriv (f x)
  (/ (- (of f (+ x *dx*)) (of f x)) *dx*))
```

Study this definite integral function carefully: it is Riemann integration in a nutshell.

```
(defun avg (x y) (float (/ (+ x y) 2)))

(defun definteg (f a b n)
  (if (zerop n) (* (avg (of f a) (of f b)) (- b a))
                (+ (definteg f a (avg a b) (- n 1))
                   (definteg f (avg a b) b (- n 1)))))
```

As a bonus, we can compute small real values of the Gamma function:

```
(defun gamma (z)
  (definteg #'(lambda (x)
    (* (expt x (- z 1)) (expt *e* (* x -1)))) 0 (/ 1 *dx*) 10))
```

Try it out by downloading a Lisp (e.g. sbcl), writing the above commands to a file called `code.lisp`, and running the following:

```
sbcl --load "code.lisp"

(deriv #'gamma 7)
```

# Chapter 11

# EIGENVALUES AND EIGENVECTORS

Herein we cover the spectral theory of compact self-adjoint operators.

Throughout, $\Omega$, $\Lambda$, $\Theta$, etc. denote compact operators.

## 11.0 SPECTRUM OF AN OPERATOR

Let $X$ be a Banach space, $B(X)$ the algebra of bounded operators on $X$.

Define the spectrum of $\Omega \in B(X)$ as

$$\sigma(\Omega) = \{\omega \in \mathbf{C} : \Omega - \omega \text{ has no inverse in } B(X)\}.$$

Then $\omega \in \sigma(\Omega)$ in three cases:

1. $\Omega - \omega$ is not injective.

   Then $\omega \in \sigma_p(\Omega)$, the **point spectrum** of $\Omega$.

2. $\Omega - \omega$ is injective but does not have dense range.

   Then $\omega \in \sigma_c(\Omega)$, the **continuous spectrum** of $\Omega$.

3. $\Omega - \omega$ is injective and has dense range.

   Then $\omega \in \sigma_r(\Omega)$, the **residual spectrum** of $\Omega$.

Collecting everything,

$$\sigma(\Omega) = \sigma_p(\Omega) \sqcup \sigma_c(\Omega) \sqcup \sigma_r(\Omega).$$

## 11.1 PROJECTION-VALUED MEASURES

Let $(X, \Sigma)$ be a measurable space, $Y$ a Hilbert space.

A function $\mu : \Sigma \to B(Y)$ is a **projection valued measure** when the following hold:

1. For each $E \in \Sigma$, $\mu(E)$ is an orthogonal projection.
2. We have $\mu(\varnothing) = 0$ and $\mu(X) = \mathrm{id}$.
3. If $(E_n)_{n\geq 0} \subseteq \Sigma$ are pairwise disjoint, then
$$\mu\left(\bigcup_{n\geq 0} E_n\right)v = \sum_{n\geq 0} \mu(E_n)v$$
where the convergence of the sum is in the norm topology.
4. For $E_1, E_2 \in \Sigma$, we have $\mu(E_1 \cap E_2) = \mu(E_1)\mu(E_2)$.

## 11.2 SPECTRA OF COMPACT SELF-ADJOINT OPERATORS

Let $X$ be a Hilbert space, $\Omega : X \to X$ a compact operator.

Then there exists a unique atomic projection-valued measure $\mu^\Omega$ on the Borel sets of $\sigma(\Omega)$ such that
$$\Omega = \int_{\sigma(\Omega)} \omega \, d\mu^\Omega(\omega).$$
Said another way, there is an orthonormal basis $(e_\omega)_\omega$ for $X$ consisting of eigenvectors of $\Omega$, such that
$$\Omega = \sum_{\sigma(\Omega)} \omega \langle \cdot, e_\omega \rangle e_\omega.$$
Further when $\Omega$ is self-adjoint, $\sigma(\Omega) \subset \mathbf{R}$.

## 11.3 ALGEBRAIC AND GEOMETRIC MULTIPLICITY

## 11.4 SIMULTANEOUS DIAGONALIZATION

Two nondegenerate commuting operators may be diagonal with respect to a single eigenbasis.

## 11.5 EXERCISES AND REMARKS

# Chapter 12

# TEMPERED DISTRIBUTIONS

Fix a positive integer $N$ and let $L^p := L^p(\mathbf{R}^N, \mathbf{C})$ for $p \in [1, \infty]$.

Throughout, $dx$ denotes Lebesgue measure rescaled by $(2\pi)^{-N/2}$.

## 12.0 THE FOURIER TRANSFORM ON $L^1$

**Definition 12.1.** There exist functions

$$\mathcal{F} : L^1 \to L^\infty \quad \text{via} \quad f(x) \mapsto (\mathcal{F}f)(\xi) = \int_{\mathbf{R}^N} f(x) \exp(-i\langle x, \xi \rangle) dx$$

called the **Fourier transform**, and

$$\mathcal{F}^{-1} : L^1 \to L^\infty \quad \text{via} \quad g(\xi) \mapsto (\mathcal{F}^{-1}g)(x) = \int_{\mathbf{R}^N} g(\xi) \exp(i\langle x, \xi \rangle) d\xi$$

called the **Inverse Fourier Transform**.

## 12.1 THE SCHWARTZ SPACE

The Fourier transform has two opposing key takeaways:

*the smoother $f$ is, the faster $\mathcal{F}f$ decays; the faster $f$ decays, the smoother $\mathcal{F}f$ is.*

So, we move to a environment where we may assume smoothness and rapid decay.

**Definition 12.2.** The **Schwartz space** $\mathcal{S}$ consists of all $\varphi \in C^\infty$ such that

$$\varphi_u^{(n)}(v_1, \ldots, v_n) = \mathcal{O}(\|u\|^{-M}) \quad \text{as } x \to \infty$$

for all $n \geq 0$, $(v_k)_{k=1}^n \subset \mathbf{R}^N$, and $M > 0$,

where $\varphi_u^{(n)}$ is the $n$th Gateaux derivative at $u$.

All smooth functions with compact support are in $\mathcal{S}$.

Let's restrict the domain of $\mathcal{F}$ to the space $\mathcal{S}$.

For $\varphi \in \mathcal{S}$, we have the following inversion formula:

$$\mathcal{F}^{-1}\mathcal{F}\varphi = \varphi$$

## 12.2 THE FOURIER TRANSFORM ON $L^2$

It's an isometry!

## 12.3 THE DUAL OF THE SCHWARTZ SPACE

We call $\mathcal{S}'$ the space of **tempered distributions**.

## 12.4 THE FOURIER TRANSFORM ON $\mathcal{S}'$

Here is how to extend the Fourier transform to the entire dual space $\mathcal{S}'$:

$$\langle \mathcal{F}f, \varphi \rangle = \langle f, \mathcal{F}\varphi \rangle$$

for all $f \in \mathcal{S}'$ and $\phi \in \mathcal{S}$.

## 12.5 HERMITE FUNCTIONS AS EIGENFUNCTIONS

## 12.6 EXERCISES AND REMARKS

$\mathcal{S}$ is an example of a non-normable metric space.

Fun fact: Quantum Mechanics does not take place in Hilbert space.

Rather, it takes place on a *rigged/enriched Hilbert space*,

also known as a **Gelfand triple**. The classic one is

$$\mathcal{S} \subset \mathrm{L}^2 \subset \mathcal{S}'$$

These are continuous dense inclusions. There are tempered distributions that live in $\mathcal{S}'$ but not $\mathrm{L}^2$, such as the Dirac delta distribution $\delta$.

In fact, we have

$$\mathcal{F}1 = \delta$$

and this is equivalent to Fourier inversion for Schwartz functions at 0.

# RETROSPECT

We have just completed a gradual mathematical progression that started with propositional logic and set theory, and ended with spectral theory and distributions.

By no coincidence, the background knowledge gained through reading this book stations the reader so that they are ready to begin understanding the foundations of quantum mechanics and computation from a clean, theoretical viewpoint.

We sketch this now.

- Every quantum system can be thought of as a $C^*$-algebra of *observables* which we will denote $\mathcal{O}$. When $\mathcal{O}$ is commutative, we recover the setting of classical mechanics; this reflects the fact that classical observables can be measured simultaneously and in any order.
- Quantum phenomena occur when $\mathcal{O}$ is noncommutative: one usually takes $\mathcal{O}$ to be the self-adjoint (i.e. Hermitian) operators on a Hilbert space. Recall that the eigenvalues of these operators are real numbers; these correspond to physical measurements.
- A quantum state is a positive linear functional $\omega$ on $\mathcal{O}$, normalized so that $\omega(1) = 1$. Pure states arise from unit vectors $|\psi\rangle$, whereas mixed states correspond to density matrices $\rho$, aka positive operators with trace 1. Unitary operators evolve the state of the system through time.
- We may regard position and momentum observables as operator-valued distributions. Of course, position and momentum are Fourier conjugates; that is, each is the other's Fourier transform, up to a scaling factor.