Analysis & Its Foundations

from axiomatic set theory to C^*-algebras

A Monica Queen Exposition

Fall 2020/21

Abstract

This expository essay is aimed at introducing a variety of topics (in particular, *Linear Analysis* and *Mathematical Logic & the Foundations of Mathematics*), where some are connected, and others stand on their own. All discussed topics will however increase our understanding of the foundations of mathematics and analysis. The presentation is based on a lot of different sources of material (check references). A background in *Real & Complex Analysis* and *Linear Algebra* is essential, and while *Topological and Metric Spaces* is not essential, it would be useful.

Preface

This essay contains some easy exercises that the reader is recommended to attempt, to stay engaged in the material. But, mainly to keep the essay shorter, otherwise it would turn into a book.
In Chapters §2, §3, & §4, to lighten up the mood and let the reader enjoy the material, we will include many Latin phrases which will be italicised for a clear indication of such, and will usually be translated to English at first. A list of the Latin phrases used is in the Appendix C.
The chapters are *ordered* by Arabic numerals while the sections and subsections are *ordered* by Roman numerals. This is so the results do not get confused with the sections. So, 2.I is a section, while 4.13 is a result.

Acknowledgements

I would like to thank my supervisors *Paul* and *Fionntan* for their support and patience. Also my mentors *Amy* and *Eleanor*. And a special thanks goes out to my friend *Oghenetekevwe* for being there throughout all of this. Also, *Ermian* for forcing me to just submit this piece of work and move on to the other fun part (i.e., the thesis). Also, my friend *Freya*. And, last, but definitely not the least, my *brother*, *sister* and *parents*.

Contents

1 Introduction

The aims of this essay are two-fold. The *first* is to give a solid introduction to *Mathematical Logic & Set Theory*, in particular *First-Order Logic* and the *Zermelo-Fraenkel Axioms* with the *Axiom of Choice* (**ZFC**). Here, we will see constructions of the number systems in a purely set-theoretic way; namely, the natural numbers, the integers, the rationals, and the real numbers. This gives us a good base to discuss concepts of Analysis. The *second* aim is to give an introduction to *Linear Analysis*. These two topics should prepare the reader to study more advanced topics, such as *Model Theory* and *Operator Algebras*.

David Hilbert's goal was to find a formal language for all of mathematics. This is possible with the **ZFC** axiomatic system of set theory, which we will look into in Chapter §3.

Axiomatic set theory provides a conceptual foundation in mathematics. This allows us to be as rigorous as possible at defining things: to have a complete understanding of the foundations of mathematics. So we will get a brief understanding of how *all* of mathematics can, essentially, be formalised in terms of sets.

We do not assume any background in formal logic.

Outline

The aims of Chapters §2 and §3 is to understand the importance of set theory, and see how all the important number systems, namely $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$, can be constructed purely from sets (with respect to the axioms). The rough outline is as follows. Chapter §2 gives us an introduction into first-order logic. We will first see propositional logic, which is just the study of combining logical statements, and predicate logic, which is an extension of propositional logic with the use of variables and quantifiers. In the end, we will see some formal proof systems, including natural deduction systems. Chapter §3 is on the **ZFC** set theory. After introducing the axioms, we will show how the natural numbers can be naturally (in the sense of the axioms) constructed from the empty set. From this, we can construct integers, and hence the rationals. The construction of the real numbers is given by Cauchy sequences, and so will be a *nice* application to the completion of normed vector spaces found in Chapter §5. This concludes the first part of the essay, namely, *Logic & the Foundations of Mathematics*.

The first part sets the scene for the next chapters (§5 – §8.I), i.e., Functional (Linear) Analysis, which is essentially calculus (analysis) on vector spaces. In this part, we will explore results from a first course on functional (linear) analysis. In Chapter §5, we will will introduce normed vector spaces (which will consequentially include results from topological and metric spaces). In Chapter §6, we will explore results from the spaces of continuous linear maps. We will then introduce Hilbert spaces in Chapter §7. And, finally, to conclude this part of study, we will see Banach algebras and spectral theory in Chapter §8. To conclude this part of the essay, we will see some results on C^**-algebras* in Chapter §8.I.

2 First-Order Logic

To precisely state the axioms of set theory, we need a formal language: the language of first-order logic. This is to avoid any ambiguity that arises in natural languages. We will not go deep into formal logic and proof theory. The interested reader is directed to [2], [6], [17], [23], and [29].

This chapter is based on the presentations given in [2], [5], [6], [17], [23], [29], and [31].

Logic, primarily, first-order logic, is essentially the study of arguments, and is the language of set theory (along with the set membership relation $\in$).

2.I Propositional logic

Propositional logic, sometimes called propositional calculus, is the study of logical statements with logical connectives (defined below). So, we can think of this as a way to describe combining statements and sentences.

Definition 2.1. A *proposition* p is a variable (i.e., a statement or formal expression) that has a *truth value*, i.e., true (in latin: *verum*) denoted as $\top$ or T, or false (in latin: *falsum*) denoted as $\bot$ or F.

Remark. It is important to note that, as logicians and mathematicians, it is our responsibility to ensure the *validity* of the arguments, but not our responsibility to ensure the truth of the premises. For instance, the sentence,

> If all humans can fly, and I am a human. Therefore, I can fly.

is logically valid. However, the premise that all humans can fly is obviously not true.

A *formal language* consists of an *alphabet* (i.e., a collection of symbols), *syntax* (i.e., what is in the language and how we can form formulas), and *semantics* (i.e., the interpretations of formulas). We will denote the *language of propositional logic* as $\mathcal{L}_{\text{Prop}}$.

In order to discuss the syntax and semantics of our language we need a *metalanguage*, i.e., a language that describes the objects in our language. For propositional logic and predicate logic, we will use English as our metalanguage.

The *alphabet* of propositional logic consists of the following symbols,

- propositional variables $a, b, \ldots$
- the *unary operator*: *negation*, $\neg$
- the *binary connectives (operators)* $\Rightarrow, \wedge, \vee, \Leftrightarrow$, which stand for *implication*, *conjunction*, *disjunction*, and *equivalence*, respectively
- *punctuation marks*: the left and right parentheses, (and), commas ,, and periods .
- *nullary constants*: $\top$ and $\bot$

To remove any ambiguity, the *alphabet* of the language consists of propositional variables and the following symbols,

$$\neg \ \Rightarrow \ \wedge \ \vee \ \Leftrightarrow \ \top \ \bot \ (\ ) \ , \ .$$

The *syntax* of propositional logic is defined by one of the following,

(i) every propositional variable is a *formula*

(ii) $\bot$ and $\top$ are *formulas*

(iii) if A and B are *formulas*, then so are all $\neg A$, $(A \Rightarrow B)$, $(A \wedge B)$, $(A \vee B)$, and $(A \Leftrightarrow B)$.

The *binding strength* of the logical operators are as follows:

$$\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$$

Also note that the connectives are read from right to left (see below example).

Example 2.2. We can build new propositions (formulas) using logical operators. Let p, q and r all be propositions.

So then p and q are also formulas. So are $\neg p$ and $(p \Rightarrow q)$. And so is $(((p \Rightarrow q) \wedge (\neg p))$. And so on ....

Because of the binding strength of the connectives, we can omit the brackets to make reading the proposition easier. For instance $(((p \Rightarrow q) \wedge (\neg p))$ can be rewritten as $(p \Rightarrow q) \wedge \neg p$.
However, if we were given the following formula $p \Rightarrow q \vee r \Rightarrow p$, then it is an abbreviation of $p \Rightarrow (q \vee r) \Rightarrow p$, since $\vee$ binds stronger than $\Rightarrow$. And this is an abbreviation of $p \Rightarrow ((q \vee r) \Rightarrow p)$ since the connectives are read from right to left.

So the above example shows that we can ignore the parentheses, if the formula does not become ambiguous.

Now to talk about the *semantics* of propositional logic, we need to define what the logical operators really "mean." We do this with *truth tables*, i.e., we list out all the possible truth values of each propositional variable in a propositional formula. So we define a *valuation* (or *truth assignment*) on the language $\mathcal{L}_{\text{Prop}}$ that assigns a truth value of T (*verum*) or F (*falsum*) to the propositional variable. A valuation is uniquely determined by their propositional variables. For example, given propositions p and q, if the valuation of p is T and the valuation of q is F, then the valuation of $p \vee q$ is T. We can write $[\![p]\!]$ for the valuation of p. Considering the truth tables below, each row shows the valuation of each propositional variable and the valuation of the formula under each connective.

Remark. The implication operator is defined in a way that represents the Latin phrase: *ex falso sequitur quodlibet*, i.e., "from falsehood, anything follows." This means that, for any valuation in our language, the valuation of the formula p implies q is always *verum* whenever the valuation of p is *falsum*. *De facto*, the truth value of an implication, say $p \Rightarrow q$, is *falsum* if and only if p (called the *antecedent*) is *verum* and q (called the *consequent*) is *falsum*, otherwise it is *verum*.

The truth table of the binary operators is as follows.

p	q	$p \wedge q$	$p \vee q$	$p \Rightarrow q$	$p \Leftrightarrow q$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	F	T	T	F
F	F	F	F	T	T

The truth table of the unary operators and the constants is as follows.

p	$\neg p$	$\top$	$\bot$
T	F	T	F
F	T	T	F

A *tautology* t is a proposition which is always *verum* (i.e., has a true value for all valuations), denoted as $\vDash t$. You can usually show that that a proposition is a tautology by showing the truth table. We say that the proposition is *valid* if it is a tautology. A *contradiction* is a proposition which is always *falsum*.

Given propositions p, q. We say q is a *logical consequence* of p, and write it as $p \vDash q$, if for

any valuation we have that if p is *verum* then q is *verum*.

Proposition 2.3. *Given propositions p and q. We have $p \vDash q$ if and only if $\vDash (p \Rightarrow q)$.*

Proof.

($\Rightarrow$) Suppose $p \vDash q$. Then for any valuation, we have that if p is *verum* then q is *verum*. So then the valuation of $(p \Rightarrow q)$ is also *verum*. Now if we have that if the valuation of p is *falsum*, then *ex falso sequitur quodlibet*, so the valuation of $(p \Rightarrow q)$ is *verum*. And so we have the following truth table

p	q	$p \Rightarrow q$
T	T	T
F		T

So, $\vDash (p \Rightarrow q)$.

($\Leftarrow$) Now suppose $\vDash (p \Rightarrow q)$, so for any valuation, we have $(p \Rightarrow q)$ is *verum*. So if the valuation of p is *verum*, then the valuation of q is also *verum* by the definition of the implication operator. So $p \vDash q$.

□

So, then we can show q is a logical consequence of p by showing $p \Rightarrow q$ is a tautology.

Now two propositions p and q are said to be *logically equivalent* if $(p \vDash q)$ and $(q \vDash p)$, written as $p \vDash\dashv q$, i.e., for every valuation, the valuation of p if and only if the valuation of q (in other words, has the same truth value for all valuations).

Proposition 2.4. *The proposition p is logically equivalent to the proposition q if and only if $\vDash (p \Leftrightarrow q)$.*

Proof.

($\Rightarrow$) Suppose p and q are logically equivalent. Then for any valuation we have that the truth value of p is the same as the truth value of q, and so we get the truth table,

p	q	$p \Leftrightarrow q$
T	T	T
F	F	T

So, $\vDash (p \Leftrightarrow q)$.

($\Leftarrow$) Now suppose $\vDash (p \Leftrightarrow q)$. Then for any valuation, we have that the valuation of p if and only if the valuation of q since the valuation of $(p \Leftrightarrow q)$ is *verum*. So, we have p and q are logically equivalent.

□

So to show that propositions p and q are logically equivalent, we can just show p if and only if q is a tautology.

Remark. We could have used Proposition 2.3 to show the above. It follows that given propositions p and q, the proposition $p \Leftrightarrow q$ is equivalent to $(p \Rightarrow q) \wedge (q \Rightarrow p)$.

Now, sometimes, in order to prove a statement, we will need to prove the *contrapositive* statement, i.e., $p \Rightarrow q$ is logically equivalent to $\neg q \Rightarrow \neg p$ (our next result).

Proposition 2.5. *Given propositions p and q. We have $\vDash (p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$, in other words, $p \Rightarrow q$ is logically equivalent to $\neg q \Rightarrow \neg p$.*

Proof. We show this with the following truth table. We see that the columns for p implies q is the same as its contrapositive statement.

p	q	$\neg p$	$\neg q$	$p \Rightarrow q$	$\neg q \Rightarrow \neg p$
T	T	F	F	T	T
T	F	F	T	F	F
F	T	T	F	T	T
F	F	T	T	T	T

We can also add another column to the table, namely, $(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$, and we can easily see that it is a tautology, and so the result follows. □

Proposition 2.6. *We can define the connectives $\neg, \vee, \wedge$ with only $\Rightarrow$ and $\bot$. In particular, for propositions p and q,*

(i) $\neg p \vDash\dashv (p \Rightarrow \bot)$,

(ii) $(p \wedge q) \vDash\dashv \neg(p \Rightarrow \neg q)$,

(iii) $(p \vee q) \vDash\dashv (\neg p \Rightarrow q)$.

Proof.

p	q	$\neg p$	$p \Rightarrow \bot$	$p \wedge q$	$\neg q$	$p \Rightarrow (\neg q)$	$\neg(p \Rightarrow (\neg q))$	$p \vee q$	$(\neg p) \Rightarrow q$
T	T	F	F	T	F	F	T	T	T
T	F	F	F	F	T	T	F	T	T
F	T	T	T	F	F	T	F	T	T
F	F	T	T	F	T	T	F	F	F

□

From the above result, we can see that we can restrict our definition of the alphabet of $\mathcal{L}_{\text{Prop}}$ to consist of $\bot$ and $\Rightarrow$ instead of $\neg, \Leftrightarrow, \wedge$ and $\vee$.

The following is left as an exercise.

Proposition 2.7. *Given propositions p and q. We have*

(i) $\neg(p \vee q) \vDash\dashv (\neg p \wedge \neg q)$

(ii) $\neg(\neg p) \vDash\dashv p$

(iii) $((p \Rightarrow q) \vee (q \Rightarrow r)) \vDash\dashv (p \vee \neg p)$ □

2.II Predicate logic (first-order logic)

Predicate logic, also known as predicate calculus or first-order logic, is an extension of propositional logic such that it includes variables and quantified variables.

The *alphabet* of the *language of predicate logic* $\mathcal{L}_{\text{Pred}}$ consists of variable symbols $x, y, \ldots$, *predicate* (*relation*) symbols $P, Q, \ldots, p, q, \ldots$, the binary relation *equals* $=$, the connectives $\neg \wedge \vee \Rightarrow$, the constants $\top \bot$, punctuation marks ( ) , . :, and the *universal and existential quantifiers* $\forall$ and $\exists$.

Essentially, a *predicate* is a statement with variables and has its truth value dependent on its variables.

Given a formula φ. We define the *universal quantifier* $\forall$ which stands for "for all." For example, $\forall x : \varphi$ is read as, "for all x (in the universe of discourse) such that φ is true." We also define the *existential quantifier* $\exists$ which stands for "there exists." For example, $\exists x : \varphi$ is read as, "there exists an x (in the universe of discourse) such that φ is true."

The *syntax* of this language is defined by the following

(i) any variable is a *term*

(ii) for any predicate P, and variables $x, y, \ldots$, we have $P(x, y, \ldots)$ is also a *term*

(iii) for any terms $a, b, \ldots$ and predicate P, we have $P(a, b, \ldots)$ is a *formula*

(iv) for terms p, q, $p = q$ is a *formula*

(v) if φ and ϕ are *formulas*, then so are $(\neg\varphi), (\varphi \Rightarrow \phi), (\varphi \wedge \phi), (\varphi \vee \phi)$

(vi) if φ is a *formula* and x is a variable, then $\forall x : \varphi$ and $\exists x : \varphi$ are *formulas*

(vii) $\top$ and $\bot$ are *formulas*

We say that a variable in a formula is *bound* if it in the scope of a quantifier. If the variable is not bound, then it is *free*.

The *binding strength* of the operators are now

$$\neg, \forall x, \exists x, \wedge, \vee, \Rightarrow$$

Substitution in predicate logic means replacing every occurrence of a variable in a formula by a term, i.e., for a formula φ, $\varphi[t/x]$ means that we need to substitute every free occurrence of variable x with the term t in φ. For example, if we have the formula φ given by $x = y$, then the substitution $\varphi[t/x] = (t = y)$.

However, substitution of a term t for a variable x in a formula φ is only allowed if t is *free for* x in φ. We say that a term t is *free for* a variable x if and only if the substituted variables in $\varphi[t/x]$ are not bound by a quantifier. In other words, if the variables in the term do not become bound after substitution. In particular, for variables x, y, terms t, u, v and $a, \ldots, b$, formulas φ, ϕ, and predicate P,

$(\phi \odot \varphi)[t/x]$ is defined as $\phi[t/x] \odot \varphi[t/x]$ where $\odot$ is any of $\wedge, \vee, \Rightarrow$
$(\neg\phi)[t/x]$ is defined as $\neg(\phi[t/x])$
$(\forall y : \varphi)[t/x]$ is $\begin{cases} \forall y : \varphi \text{ if } x \text{ and } y \text{ are equal,} \\ \forall y : \varphi[t/x] \text{ if } x \text{ and } y \text{ are not equal} \end{cases}$ similarly for $\exists$
$(u = v)[t/x]$ is defined as $u[t/x] = v[t/x]$
$P(a, \ldots, b)[t/x]$ is $P(a[t/x], \ldots, b[t/x])$
$\bot[t/x]$ is simply $\bot$ similarly for $\top$
$z[t/x]$ is defined as $\begin{cases} z \text{ if } x \text{ and } z \text{ are not equal} \\ t \text{ if } x \text{ and } y \text{ are equal} \end{cases}$

Informally, the *semantics* of this language is just some *interpretation* (denoted as $[\![x]\!]$ for some formula x) of the formulas. For example, we want to say $\forall x : \varphi$ is *verum* only if the variable being quantified over is in the universe of discourse. So, essentially, the interpretation of

$[\![x]\!]$ is just the valuation of x
$[\![\bot]\!] \Leftrightarrow \bot$
$[\![t = u]\!]$ $\begin{cases} \Leftrightarrow \top \text{ if } [\![t]\!] \Leftrightarrow [\![u]\!] \\ \Leftrightarrow \bot \text{ if } [\![t]\!] \not\Leftrightarrow [\![u]\!], \text{ note that } \not\Leftrightarrow \text{ is an abbreviation of } \neg([\![t]\!] \Leftrightarrow [\![u]\!]) \end{cases}$
$[\![P(a, \ldots, b)]\!]$ $\begin{cases} \Leftrightarrow \top \text{ if } a, \ldots, b \text{ is in the universe of discourse} \\ \Leftrightarrow \bot \text{ otherwise} \end{cases}$
$[\![\forall x : \phi]\!]$ $\begin{cases} \Leftrightarrow \top \text{ if } [\![\phi[t/x]]\!] \text{ for every } t \text{ in the universe of discourse} \\ \Leftrightarrow \bot \text{ otherwise} \end{cases}$

For a formal definition of semantics and models, the reader is directed to [5], [31], and [35].

So, essentially, we need to understand the interpretation of the quantifiers. In essence, if we have $\forall x : p(x)$, we need to check that for any arbitrary substitution $p[t/x]$ is true. And if we have $\exists x : p(x)$, then we need at least one substitution to hold. So, $\forall x : p(x)$ would need the following to be true $p[a/x] \wedge p[b/x] \wedge \cdots$ for all terms in the universe of discourse, and $\exists x : p(x)$ would need the following to be true $p[a/x] \vee p[b/x] \vee \cdots$ for all terms in the universe of discourse. We can also form a truth table, just as before. An example of a possible truth table would be the following,

$p(a)$	$p(b)$	$\forall x : p(x)$	$\exists x : p(x)$
T	T	T	T
T	F	F	T
F	T	F	T
F	F	F	F

Note that we do not need to check that every substitution is satisfied as long as we pick our terms to be arbitrary.

Analogous to propositional logic, we can define logical consequences and equivalences with respect to the interpretation. We can also define contradictions and tautologies similarly. So we can define the existential quantifier with the universal quantifier, and vice versa. So $\exists x : P(x)$ is equivalent to $\neg(\forall x : \neg P(x))$ (see the below truth table). And, likewise, $\forall x : P(x)$ is equivalent to $\neg(\exists x : \neg P(x))$ (exercise).

$P[a/x]$	$P[b/x]$	$\neg P[a/x]$	$\neg P[b/x]$	$\forall x : \neg P(x)$	$\exists x : P(x)$	$\neg(\forall x : \neg P(x))$
T	T	F	F	F	T	T
T	F	F	T	F	T	T
F	T	T	F	F	T	T
F	F	T	T	T	F	F

Note that we can combine quantifiers in a formula, or even have more than one variable in a quantifier. Also note that the order here matters, i.e., given the predicate $P(x, y)$, the formula $\exists x : \forall y : P(x, y)$ is generally not equivalent to the formula $\forall y : \exists x : P(x, y)$.

2.III Proof theory

An *axiomatic system*, also known as a *theory*, is any collection of *axioms* (i.e., formulas that we take to be true), $\alpha, \beta, \ldots, \zeta$. We can use these axioms to logically derive other formulas. A *formal proof* of a formula p in an axiomatic system $\alpha, \beta, \ldots, \zeta$, is a succession of formulas $q, \ldots, t$ such that $p = t$, and that for each $q, \ldots, t$ we have one of the following (call the chosen formula s).

$$s \text{ is an axiom.} \tag{PI}$$

$$s \text{ is a tautology.} \tag{PII}$$

$$\exists c, a : s \text{ comes after } c \text{ and } a, (c \wedge a \Rightarrow t) \text{ is true.} \tag{PIII}$$

(PIII) is a rule that corresponds to the Latin phrase ***m****odus ponendo* ***p****onens* which translates to, "mode that by affirming affirms." This rule essentially means, for propositions p, q, if $p \Rightarrow q$ and we have p, then we get (*deduce*) q. We will usually denote this as either (mp) or the Latin phrase itself.

We say the axiomatic system $\alpha, \beta, \ldots, \zeta$ (we call this S) *proves* the formula p, if there exists a proof of p in S, and is denoted as $S \vdash p$. In propositional logic, the axiomatic system is the empty set, i.e., if proposition p can be proved in propositional logic, then $\emptyset \vdash p$. We call a formula a *theorem* of an axiomatic system, if the axiomatic system proves the formula. We say an axiomatic system S is *consistent*, if $S \vdash \bot$. We can, equivalently, say that it is consistent if contradicting formulas (i.e., α and $\neg\alpha$) are not in the list of axioms.

An example of this classical proof system (also known as the *Hilbert system*) is as follows.

Example 2.8. We will show that from $p \Rightarrow (q \Rightarrow r)$ and $p \Rightarrow q$, we can prove $p \Rightarrow r$. *Id est*, $(p \Rightarrow (q \Rightarrow r)), (p \Rightarrow q) \vdash (p \Rightarrow r)$.

First we want to show $(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$ is a tautology. This is left as an exercise. And so, we then have the following proof.

1.	$p \Rightarrow (q \Rightarrow r)$	axiom
2.	$p \Rightarrow q$	axiom
3.	$(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$	tautology
4.	$(p \Rightarrow q) \Rightarrow (p \Rightarrow r)$	*modus ponendo ponens* 1,3
5.	$p \Rightarrow r$	*modus ponendo ponens* 2,4

And thus completes the formal proof. □

Remark. Note that in our proofs, we will usually call the axioms, *premises*.

Also note that to make things easier, instead of having to show tautologies constantly (which are of the same form, except with different variables), we can form a collection of axioms. For instance, in the above example, we said $(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$ is a tautology. *De facto*, we can say $(\varphi \Rightarrow (\phi \Rightarrow \psi)) \Rightarrow ((\varphi \Rightarrow \phi) \Rightarrow (\varphi \Rightarrow \psi))$, where φ, ϕ

and ψ are *meta-variables* (so can be replaced with any other variable in our language) is a tautology (and hence can be used as an axiom in the Hilbert style proof).

Before going deeper into other systems of proof theory, we will talk about soundness and completeness.

2.IV Soundness and Completeness

Here, we will see a connection between the semantics and syntax of logic. The proofs are omitted, but can be found in [5], [29], and [31].

Lemma 2.9 (Soundness)**.** *If our theory, $\alpha, \beta, \ldots, \zeta$, denoted by S, proves ϕ, then ϕ is a logical consequence of S. That is, $S \vdash \phi \Rightarrow S \vDash \phi$. In other words, all conclusions that are derivable from true premises are true.* □

Lemma 2.10 (Completeness)**.** *If a formula ϕ is a logical consequence of our theory, $\alpha, \beta, \ldots, \zeta$, denoted by S, then S proves ϕ. That is, $S \vDash \phi \Rightarrow S \vdash \phi$.* □

Theorem 2.11. *Let S denote our theory $a_1, a_2, \ldots, a_n$, and let p be a formula. Then $S \vDash p$ if and only if $S \vdash p$.* □

Both the Hilbert system and Natural deduction (also known as Fitch's system) are sound and complete for propositional and predicate logic.

2.V Natural deduction

In this section we will learn about natural deduction.

This is based on the presentations given in [6] and [29].

Natural deduction is a method of proof where we "naturally deduce" the arguments. This method of formal logical proof corresponds closely to our non-formal mathematical proof. The rules are called *inference rules*, and are written via proof trees, so what goes above the line are called *premises* and below the line is the *conclusion*.

Example 2.12.

$$\frac{\text{humans can fly} \quad \text{I am human}}{\text{I can fly}}$$

The premises are, "humans can fly," and "I am human." So we conclude that, "I can fly."

If A and B are true, then we can conclude $A \wedge B$ is also true. We call this rule the *conjunction introduction rule*, and denote it as $\wedge$-int. Similarly, if we have A, then we can conclude $A \vee B$ or $B \vee A$, and we call this rule the *disjunction introduction rule*, $\vee$-int. We also have the *conjunction elimination rule*, where if we have $A \wedge B$, then we conclude A and B (we write them as two separate rules, one for the left and one for the right).

$$\frac{A \quad B}{A \wedge B}\,\wedge\text{-int}$$

$$\frac{A}{A \vee B}\,\vee\text{-int} \qquad \frac{B}{A \vee B}\,\vee\text{-int}$$

$$\frac{A \wedge B}{A}\,\wedge\text{-elim} \qquad \frac{A \wedge B}{B}\,\wedge\text{-elim}$$

Another rule is called the *implication elimination rule*, if we have A and $A \Rightarrow B$, then *modus ponendo ponens* and we conclude B.

$$\frac{A \quad A \Rightarrow B}{B} \Rightarrow\text{-elim}$$

Remark. Note that we will sometimes regard the above rule as $\Rightarrow$-elim (for instance, in proof trees for readability), but usually we will write it as *modus ponendo ponens*.

We also have the *implication introduction rule*: if we have A and conclude C, then we clearly have $A \to C$. Also, if we assume A (which we write as $[A]$) and conclude C, then we have $A \to C$.

$$\frac{\begin{array}{c} A \\ \vdots \\ C \end{array}}{A \Rightarrow C} \Rightarrow\text{-int} \qquad \frac{\begin{array}{c} [A] \\ \vdots \\ C \end{array}}{A \Rightarrow C} \Rightarrow\text{-int}$$

This is because if our assumption is incorrect then *ex falso sequitur quodlibet*.

Example 2.13. An example of how a proof tree might look like is the following

$$\dfrac{\dfrac{\dfrac{[A] \quad B}{A \wedge B} \wedge\text{-int} \quad A \wedge B \Rightarrow C}{C} \Rightarrow\text{-elim}}{A \Rightarrow C} \Rightarrow\text{-int}$$

Which we can translate to a sequential proof.

1.	B	premise
2.	$A \wedge B \Rightarrow C$	premise
3.	A	assumption
4.	$A \wedge B$	$\wedge$-int 1,3
5.	C	*modus ponendo ponens* 2,4
6.	$A \Rightarrow C$	$\Rightarrow$-int 3–5

Now, suppose that we have $A_i \vee \cdots \vee A_j$ and we proved C from the assumption of each $A_i, \ldots, A_j$, then we can conclude C. This is the *disjunction elimination rule*.

$$\frac{A_i \vee \cdots \vee A_j \qquad \begin{array}{c}[A_i] \\ \vdots \\ C\end{array} \quad \cdots \quad \begin{array}{c}[A_j] \\ \vdots \\ C\end{array}}{C} \vee\text{-elim}$$

Since *ex falso sequitur quodlibet*, we also have the *falsum elimination rule*, i.e.,

$$\frac{\bot}{A} \bot\text{-elim}$$

We also have the *contradictum (contradiction) rule*,

$$\frac{A \quad \neg A}{\bot} \textit{contradictum}$$

Another rule is one that follows the Latin phrase *reductio ad absurdum*, "reduction to absurdity." So if we can prove $\bot$ from $\neg A$ (equivalently, $A \to \bot$), then we can conclude A is true. This argument is known as proof by contradiction. For proof trees, we will denote this as $\neg$-elim since it 'eliminates the negation.'

$$\dfrac{\begin{array}{c}[A \to \bot] \\ \vdots \\ \bot\end{array}}{A}\ \neg\text{-elim}$$

The following example is taken from [6].

Proposition 2.14. $\vdash A \Rightarrow (B \Rightarrow C) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$

Proof.

1.	$A \Rightarrow (B \Rightarrow C)$	assumption
2.	$A \Rightarrow B$	assumption
3.	A	assumption
4.	$A \Rightarrow (B \Rightarrow C)$	copied from 1
5.	$B \Rightarrow C$	*modus ponendo ponens* 3,4
6.	$A \Rightarrow B$	copied from 2
7.	B	*modus ponendo ponens* 3,6
8.	C	*modus ponendo ponens* 5,7
9.	$A \Rightarrow C$	$\Rightarrow$-int 3–8
10.	$(A \Rightarrow B) \Rightarrow (A \Rightarrow C)$	$\Rightarrow$-int 2–9
11.	$A \Rightarrow (B \Rightarrow C) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$	$\Rightarrow$-int 1–10

□

Proposition 2.15. $\vdash A \Rightarrow ((((A \Rightarrow B) \Rightarrow B) \Rightarrow C) \Rightarrow C)$

Proof.

1.	A	assumption
2.	$((A \Rightarrow B) \Rightarrow B) \Rightarrow C$	assumption
3.	$A \Rightarrow B$	assumption
4.	B	*modus ponendo ponens* 1,3
5.	$(A \Rightarrow B) \Rightarrow B$	$\Rightarrow$-int 3–4
6.	C	*modus ponendo ponens* 2,5
7.	$(((A \Rightarrow B) \Rightarrow B) \Rightarrow C) \Rightarrow C$	$\Rightarrow$-int 2–6
8.	$A \Rightarrow ((((A \Rightarrow B) \Rightarrow B) \Rightarrow C) \Rightarrow C)$	$\Rightarrow$-int 1–7

□

Proposition 2.16. $\vdash A \Rightarrow (A \Rightarrow A)$

Proof.

1.	A	hypothesis
2.	A	copied from 1
3.	A	copied from 1
4.	$A \Rightarrow A$	$\Rightarrow$-int 2–3
5.	$A \Rightarrow (A \Rightarrow A)$	$\Rightarrow$-int 1–4

□

The next result shows that the disjunction is commutative. It is easy to check that conjunctions are also commutative, and is left as an exercise.

Proposition 2.17. $p \vee q \vdash q \vee p$

Proof.

1.	$p \vee q$	premise
2.	p	assumption
3.	$q \vee p$	$\vee$-int 2
4.	q	assumption
5.	$q \vee p$	$\vee$-int 4
6.	$q \vee p$	$\vee$-elim 2–3, 4–5

□

This next result is the opposite of *modus ponendo ponens*, and is also known by its Latin phrase *modus tollendo tollens*, this translates to, "mode of denying denies." Which means that if we have $p \Rightarrow q$ and $\neg q$, then we can conclude $\neg p$.

Proposition 2.18. $p \Rightarrow q, \neg q \vdash \neg p$

Proof.

1.	$p \Rightarrow q$	premise
2.	$\neg q$	premise
3.	p	assumption
4.	q	*modus ponendo ponens* on 1,3
5.	$\bot$	*contradictum* 2,4
6.	$p \Rightarrow \bot$	$\Rightarrow$-int 3–5
7.	$\neg p$	6 rewritten

□

This next result is known as the *law of excluded middle*, or in Latin: *tertium non datur*, which translates to, "the third is not given."

Proposition 2.19. $\vdash p \vee \neg p$

Proof.

1.	$\neg(p \vee \neg p)$	assumption
2.	p	assumption
3.	$p \vee \neg p$	$\vee$-int 2
4.	$\bot$	*contradictum* 1,3
5.	$p \Rightarrow \bot$	$\Rightarrow$ -int 2–4
6.	$\neg p$	5 rewritten
7.	$p \vee \neg p$	$\vee$-int 6
8.	$\bot$	*contradictum* 1,7
9.	$p \vee \neg p$	*reductio ad absurdum* 1–8

□

Proposition 2.20. $\neg(p \wedge q) \dashv\vdash \neg p \vee \neg q$

Proof. For $\vdash$:

1.	$\neg(p \wedge q)$	premise
2.	$\neg p \vee \neg q \Rightarrow \bot$	assumption
3.	$\neg(\neg p \vee \neg q)$	2 rewritten
4.	$p \Rightarrow \bot$	assumption
5.	$\neg p$	4 rewritten
6.	$\neg p \vee \neg q$	$\vee$-int 5
7.	$\bot$	*contradictum* 3,6
8.	p	*reductio ad absurdum* 4–7
9.	$q \Rightarrow \bot$	assumption
10.	$\neg q$	9 rewritten
11.	$\neg p \vee \neg q$	$\vee$-int 10
12.	$\bot$	*contradictum* 3,11
13.	q	*reductio ad absurdum* 9–12
14.	$p \wedge q$	$\wedge$-int 8,13
15.	$\bot$	*contradictum* 3,14
16.	$\neg p \vee \neg q$	*reductio ad absurdum* 2–15

For $\dashv$:

1.	$\neg p \vee \neg q$	premise
2.	$p \wedge q$	assumption
3.	p	$\wedge$-elim 2
4.	q	$\wedge$-elim 2
5.	$\neg p$	assumption
6.	$\bot$	*contradictum* 2,5
7.	$\neg q$	assumption
8.	$\bot$	*contradictum* 3,7
9.	$\bot$	$\vee$-elim 1,5–6,7–8
10.	$p \wedge q \Rightarrow \bot$	$\Rightarrow$-int 2–9
11.	$\neg(p \wedge q)$	10 rewritten

□

Proposition 2.21. $p \Rightarrow q \vdash \neg q \Rightarrow \neg p$ *and* $\neg q \Rightarrow \neg p \vdash p \Rightarrow q$

Proof. For $p \Rightarrow q \vdash \neg q \Rightarrow \neg p$, we have,

1.	$p \Rightarrow q$	premise
2.	$\neg q$	assumption
3.	$\neg p$	*modus tollendo tollens* on 1,2
4.	$\neg q \Rightarrow \neg p$	$\Rightarrow$-int 2–3

And for $\neg q \Rightarrow \neg p \vdash p \Rightarrow q$, we have,

1.	$\neg q \Rightarrow \neg p$	premise
2.	p	assumption
3.	$\neg p$	assumption
4.	$\bot$	*contradictum* 2,3
5.	$\neg p \Rightarrow \bot$	$\Rightarrow$-int 3–4
6.	$\neg\neg p$	5 rewritten
7.	$\neg\neg q$	*modus tollendo tollens* 1,6
8.	$q \Rightarrow \bot$	assumption
9.	$\neg q$	8 rewritten
10.	$\bot$	*contradictum* 7,9
11.	q	*reductio ad absurdum* 8–10
12.	$p \Rightarrow q$	$\Rightarrow$-int 2–11

□

Remark. Note that in the preceding result, we used *modus tollendo tollens* from Proposition 2.18. We can use any derived result to further simplify a proof.

When we have $A \vdash B$ and $B \vdash A$, just as seen from the previous result, we say that they are *provably equivalent*, and is written as $A \dashv\vdash B$. It is a theorem that for formulas p, q, we have q is a logical consequence of p if and only if p proves q, which we have already seen in the previous section (Theorem 2.11).

The rule for *equality introduction* is that any term (i.e., variables or constants) is equal to itself,

$$\frac{}{t = t} = \text{-int}$$

We also have the *equality elimination rule*,

$$\frac{s = t \quad \varphi[s/x]}{\varphi[t/x]} = \text{-elim}$$

The following result shows that the equality predicate is reflexive (*ipso facto* from equality introduction), symmetric and transitive.

Proposition 2.22. *$s = t \vdash t = s$ and $s = t, t = u \vdash s = u$*

Proof. For $s = t \vdash t = s$ we have,

1.	$s = t$	premise
2.	$s = s$	$=$-int
3.	$t = s$	$=$-elim 1,2, $\varphi(x) := (x = s)$

For $s = t, t = u \vdash s = u$ we have,

1.	$s = t$	premise
2.	$t = u$	premise
3.	$s = u$	$=$-elim 1,2, $\varphi(x) := (s = x)$

□

Now, we have the *elimination inference rule* for $\forall$ and the *introduction inference rule* for $\exists$, which are clear. If we have $\varphi[t/x]$, then there exists some x such that φ. And if we have φ for all x, then $\phi[t/x]$ for some t in the universe of discourse.

$$\frac{\varphi[t/x]}{\exists x : \varphi}\ \exists\text{-int} \qquad \frac{\forall x : \varphi}{\varphi[t/x]}\ \forall\text{-elim}$$

The *introduction rule* for $\forall$ is valid when t is arbitrary. So, to prove ϕ for all x, we will show $\phi[t/x]$ for some arbitrary t.

$$\frac{\left[\begin{array}{c} t \\ \\ \vdots \\ \varphi[t/x] \end{array}\right]}{\forall x : \varphi}\ \forall\text{-int}$$

Now suppose we know that there exists some x such that φ, and we proved p from $\varphi[t/x]$ for some t, then we get p. This is the *elimination rule* for $\exists$.

$$\dfrac{\exists x : \varphi \qquad \left[\begin{array}{c} t \\ \varphi[t/x] \\ \vdots \\ p \end{array}\right]}{p}\ \exists\text{-elim}$$

Proposition 2.23. $\exists x : \neg\varphi \dashv\vdash \neg\forall x : \varphi$

Proof. For $\vdash$:

1.		$\exists x : \neg\varphi$	premise
2.		$\forall x : \varphi$	assumption
3.	t	$\neg\varphi[t/x]$	assumption
4.		$\varphi[t/x]$	$\forall$-elim 2
5.		$\bot$	*contradictum* 3,4
6.		$\bot$	$\exists$-elim 1,3–5
7.		$\forall x : \varphi \Rightarrow \bot$	$\Rightarrow$ -int 2–6
8.		$\neg\forall x : \varphi$	7 rewritten

For $\dashv$:

1.		$\neg\forall x : \varphi$	premise
2.		$\exists x : \neg\varphi \Rightarrow \bot$	assumption
3.		$\neg\exists x : \neg\varphi$	2 rewritten
4.	t		
5.		$\varphi[t/x] \Rightarrow \bot$	assumption
6.		$\neg\varphi[t/x]$	5 rewritten
7.		$\exists x : \neg\varphi$	$\exists$-int 6
8.		$\bot$	*contradictum* 3,7
9.		$\varphi[t/x]$	*reductio ad absurdum* 5–8
10.		$\forall x : \varphi$	$\forall$-int 4–9
11.		$\bot$	*contradictum* 1,10
12.		$\exists x : \neg\varphi$	*reductio ad absurdum* 2–11

□

Definition 2.24. Rules for a connective $\eth$ are *harmonious* if, for statements ϕ and ψ, $\phi \eth \psi$ is the strongest assertion you can deduce from the assumptions of $\eth$-int, and $\phi \eth \psi$ is the weakest thing that implies the conclusion of the $\eth$-elim rule.

To show harmony, we need to show that the introduction and elimination rule do not change our information, but only provide us with a different structure.

Proposition 2.25. *The introduction and elimination rules for* $\vee, \wedge, \Rightarrow, \forall, \exists$ *and* $=$ *are harmonious.*

Proof. For $\wedge$: we need to show that when we have ϕ and ψ, the introduction rule gives us $\phi \wedge \psi$, and the elimination rule on this gives us both ϕ and ψ (what we started with).

$$\dfrac{\dfrac{\phi \quad \psi}{\phi \wedge \psi}\,\wedge\text{-int}}{\psi}\,\wedge\text{-elim} \qquad \dfrac{\dfrac{\phi \quad \psi}{\phi \wedge \psi}\,\wedge\text{-int}}{\phi}\,\wedge\text{-elim}$$

Likewise, we also have that if we apply the elimination rule first, then the introduction rule we get back what we started with.

$$\dfrac{\dfrac{\phi \wedge \psi}{\phi}\,\wedge\text{-elim} \quad \dfrac{\phi \wedge \psi}{\psi}\,\wedge\text{-elim}}{\psi \wedge \psi}\,\wedge\text{-int}$$

Now, for $\vee$, we first need to deduce $\phi \vee \psi$ from either ϕ or ψ, and then we deduce φ by deducing it from both ϕ and ψ. Similarly, if we start with $\phi \vee \psi$, then by introduction, we can hypothetically conclude $\phi \vee \psi$ from both ϕ and ψ, and finally by elimination we get back what we started with.

$$\dfrac{\dfrac{\phi}{\phi \vee \psi}\,\vee\text{-int} \qquad \begin{matrix}[\phi] \\ \vdots \\ \varphi\end{matrix} \qquad \begin{matrix}[\psi] \\ \vdots \\ \varphi\end{matrix}}{\varphi}\,\vee\text{-elim} \qquad \dfrac{\phi \vee \psi \quad \dfrac{[\phi]}{\phi \vee \psi}\,\vee\text{-int} \quad \dfrac{[\psi]}{\phi \vee \psi}\,\vee\text{-int}}{\phi \vee \psi}\,\vee\text{-elim}$$

For $\Rightarrow$: we start with $\phi \Rightarrow \psi$, and assume ϕ, to conclude ψ, and then use the introduction rule again to get back what we started with. Similarly, if we have ϕ and prove ψ from assuming ϕ, then we get ψ.

$$\dfrac{\dfrac{\phi \Rightarrow \psi \quad [\phi]}{\psi}\,\Rightarrow\text{-elim}}{\phi \Rightarrow \psi}\,\Rightarrow\text{-int} \qquad \dfrac{\begin{matrix}[\phi] \\ \vdots \\ \psi\end{matrix}}{\phi \to \psi}\,\Rightarrow\text{-int} \quad \phi}{\psi}\,\Rightarrow\text{-elim}$$

For $\forall$,

$$\dfrac{\dfrac{\phi[t/x]}{\forall x : \phi}\,\forall\text{-int}}{\phi[t/x]}\,\forall\text{-elim} \qquad \dfrac{\dfrac{\forall x : \phi}{\phi[t/x]}\,\forall\text{-elim}}{\forall x : \phi}\,\forall\text{-int}$$

For $\exists$,

$$\dfrac{\dfrac{\phi[t/x]}{\exists x : \phi}\,\exists\text{-int} \qquad \left[\begin{matrix} a & \phi[a/x] \\ & \vdots \\ & \psi \end{matrix}\right]}{\psi}\,\exists\text{-elim} \qquad \dfrac{\exists x : \phi \quad \dfrac{\phi[t/x]}{\exists x : \phi}\,\exists\text{-int}}{\exists x : \phi}\,\exists\text{-elim}$$

Lastly, for $=$,

$$\dfrac{\dfrac{}{t = t}\,=\text{-int} \quad \phi[t/x]}{\phi[t/x]}\,=\text{-elim} \qquad \dfrac{s = t \quad \dfrac{}{s = s}\,=\text{-int}}{s = t}\,=\text{-elim}$$

□

A nice exercise to do is to prove the following.

Proposition 2.26.

(i) $\neg(\varphi \wedge \phi) \dashv\vdash \neg\varphi \vee \neg\phi$

(ii) $\neg\exists x : \varphi \dashv\vdash \forall x : \neg\varphi$

(iii) $\forall x : \forall y : \varphi \dashv\vdash \forall y : \forall x : \varphi$

(iv) $\exists x : \exists y : \varphi \dashv\vdash \exists y : \exists x : \varphi$

(v) if x is not free in ϕ, then $(\forall x : \varphi) \wedge \phi \dashv\vdash \forall x : (\varphi \wedge \phi)$

(vi) if x is not free in ϕ, then $(\exists x : \varphi) \Rightarrow \phi \dashv\vdash \forall x : (\varphi \Rightarrow \phi)$ □

The above result, namely, parts *(iii)* and *(iv)*, we see $\forall x : \forall y$ can be written as $\forall x, y$, likewise $\exists x : \exists y$ can be written as $\exists x, y$.

This concludes this brief chapter on first-order logic.

3 Axiomatic Set Theory

The aim of this chapter is to discuss the axioms of **ZFC**.

This chapter is based on the presentations given in [5], [11], [12], [25], [32], and [36].

3.I Motivation

Intuitively, a *set* is any mathematical object. A set can be equal to another set, be contained in another set, and can contain other sets. We have that either set x is in set y or set x is not in set y, written as $x \in y$ and $x \notin y$, respectively.

In Naive set theory, we have an axiom scheme known as *unrestricted comprehension*, i.e., we can define a set for all x such that $\phi(x)$ (which we write as $\{x : \phi(x)\}$). This brings up inconsistencies in our system. We have the following contradictory example, known as *Russell's paradox*. Let A denote $\{x : x \notin x\}$. Then then this would mean $A \in A$ if and only if $A \notin A$.

These inconsistencies motivate the idea of formalising set theory.

3.II The language of set theory

Formally, a *set* is not defined, apart from what the axioms say. So, we will define a *class* as any collection of objects. A *proper class* is a class which is not a set.

The language of predicate logic $\mathcal{L}_{\text{Prod}}$, together with the set membership relation symbol $\in$, is the *formal language of set theory* $\mathcal{L}$.

Our intended interpretation of the *set membership* relation symbol $\in$, say for the sentence $x \in y$, is if indeed the set x is a member (element) of set y. The negation of this relation (i.e., $\neg(x \in y)$) is $x \notin y$, and means x is not a member of y. We can also say $y \ni x$ when $x \in y$.

The *subset* relation symbol $\subseteq$, written as $x \subseteq y$, is an abbreviation of $\forall a : (a \in x \Rightarrow a \in y)$. The negation of this is $\not\subseteq$. And, obviously, $y \supseteq x$ when $x \subseteq y$. We can think of $x \subseteq y$ as, "x is contained in y" or "x is a subset of y."

In summary, the *alphabet* of $\mathcal{L}$ consists of: predicate symbols, variables, constant symbols, and the following

$$\wedge \ \vee \ \neg \ \Rightarrow \Leftrightarrow \ \forall \ \exists \ (\ ) \ , \ = \in$$

The *syntax* is given by one of the following (this is a slightly different formulation than we did in the previous section, albeit captures the same idea).

(i) Given variables x and y, then $x \in y$ and $x = y$ are *formulas*.

(ii) If φ and ϕ are *formulas*, then so are all of $(\neg\varphi), (\varphi \wedge \phi), (\varphi \vee \phi), (\varphi \Rightarrow \phi)$ and $(\varphi \Leftrightarrow \phi)$.

(iii) $\forall x : \varphi$ and $\exists x : \varphi$ are *formulas*, where x is a variable and φ is a *formula*.

The universe of discourse of the quantified variables in this language is the universe of sets (which we will see soon is not a set).

For a more formal definition of the language, see [21].

3.III The axioms of ZFC & the fundamentals of sets

We shall start with the existence of a set.

Axiom 1 - Empty set. $\exists x : \forall y : (y \notin x)$
There exists a set that contains no elements.

Axiom 2 - Extensionality. $\forall x : \forall y : (x = y) \Leftrightarrow ((x \subseteq y) \wedge (y \subseteq x))$
Two sets are equal if and only if they have the same elements.

Remark. Most texts do not begin with abbreviating the subset relation and dive straight into the axiom of extensionality, in this case, they would write it as $\forall x : \forall y : ((x = y) \Leftrightarrow \forall z : (z \in x \Leftrightarrow z \in y))$.

This axiom will inform us when an object is not a set. So, if the statement $x \in y$ is not a formula, then x and y are not both sets. Thus, we will only deal with sets from this point forward.

And now we can show that there is *de facto* only one empty set, which we denote by $\emptyset$. We start with some results that will help us.

Lemma 3.1. *Given formulas p and r. We have the tautology $((p \Rightarrow r) \wedge p) \Rightarrow r$.*

Proof. Denote $p \Rightarrow r$ by q and $q \wedge p$ by s, so we have the following truth table.

p	r	q	s	$s \Rightarrow r$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

□

Lemma 3.2. *Given formulas x, y and z. Where x and z are both empty sets, i.e. $\forall y : y \notin z$ and $\forall y : y \notin x$. We have the following tautology: $(y \notin x) \Rightarrow (x \subseteq z)$; consequentially, so is $(y \notin z) \Rightarrow (z \subseteq x)$.*

Proof. For this to be a tautology, we need $x \subseteq z$ to be *verum*, which is rewritten as $\forall y : (y \in x) \Rightarrow (y \in z)$. We already know $y \in x$ is *falsum* since $\forall y : y \notin x$ is *verum*. And so *ex falso sequitur quodlibet*. □

Note that the *unique existential quantifier* can be defined via the existential and universal (or only existential, and only universal). So $\exists! x : P(x)$ is equivalent to $\exists x : \forall y : P(y) \Leftrightarrow x = y$.

Theorem 3.3. *The empty set is unique, i.e., there exists only one empty set: $\emptyset$. This can be formally written as (Axioms 1,2)* $\vdash \exists! x : \forall y : (y \notin x)$.

Proof. Suppose the contrary, and let x and x' be empty sets such that $x \neq x'$. So by the axiom this means, $\forall y : y \notin x$ and $\forall y : y \notin x'$. We give a sketch of a semi-formal proof.

1.	$\neg\exists! x : \forall y : (y \notin x)$	assumption
2.	$\forall y : y \notin x$	Axiom 1
3.	$\forall y : y \notin x'$	Axiom 1
4.	$x \neq x'$	from 1,2,3
5.	$(y \notin x) \Rightarrow (x \subseteq x')$	tautology by lemma
6.	$(y \notin x') \Rightarrow (x' \subseteq x)$	tautology by lemma
7.	$x \subseteq x'$	lemma and *modus ponendo ponens* 2,5
8.	$x' \subseteq x$	lemma and *modus ponendo ponens* 3,6
9.	$x = x'$	from 7,8
10.	$\perp$	*contradictum* 4,9
11.	$\exists! x : \forall y : (y \notin x)$	*reductio ad absurdum* 1–10

Thus, there can only be one empty set, $\emptyset$. □

Remark. A non-formal (in the sense of logic) sketch of the above proof would be to first assume the contrary, and let x and x' both be empty sets such that $x \neq x'$. Which by Axiom 2 would mean $\forall y : y \notin x$ and $\forall y : y \notin x'$. And so by the subset abbreviation, we would have $x \subseteq x'$ and $x' \subseteq x$. Finally, by Axiom 2, we would then have $x = x'$, and so *reductio ad absurdum*, thus completing the proof.

Axiom 3 - Pair set. $\forall x : \forall y : \exists m : \forall u : (u \in m) \Leftrightarrow ((u = x) \vee (u = y))$
Given sets x and y. There exists a set whose elements are exactly x and y.

We will write this set, m, as $\{x, y\}$. Here, $\{, \}$ are added informally to make our language more readable. It is very important to note the brackets $\{\}$, since obviously set x is not equal to $\{x\}$.

It is clear from Axiom 2, if $x = y$, then $\{x, y\} = \{x\}$. So, if we have set x, then we know $\{x\}$ is also a set. The following result shows the uniqueness of the pair set and is left as an exercise.

Proposition 3.4. *Given sets x and y. We have the pair set $\{x, y\}$ equal to $\{y, x\}$.* □

Given sets x and y, we can define an *ordered pair* (x, y), with the set $\{\{x\}, \{x, y\}\}$. This is a set since by Axiom 3, we have $\{x\}$ and $\{x, y\}$ are sets, and so it follows from using Axiom 3 again. Ordered pairs are useful because $(x, y) \neq (y, x)$, which we can see from the following proposition.

Proposition 3.5. *Given sets x, y, z, w, if $(x, y) = (z, w)$, then $x = z$ and $y = w$.*

Proof. We rewrite the ordered pairs (x, y) and (z, w) as $\{\{x\}, \{x, y\}\}$ and $\{\{z\}, \{z, w\}\}$, respectively. If $x = y$, then we would have $\{\{x\}, \{x, x\}\} = \{\{x\}, \{x\}\} = \{\{x\}\}$. And so by Axiom 2, we have that both $\{z\} = \{x\}$ and $\{z, w\} = \{x\}$, which, again by Axiom 2, would mean $z = x = w = y$. So, now assume $x \neq y$ and $z \neq w$. By Axiom 2, we have $\{x\} = \{z\}$ which means $x = z$; otherwise, we would have $\{x\} = \{z, w\}$ which would contradict our assumption of $z \neq w$. We also have $\{x, y\} = \{z, y\} = \{z, w\}$, which means $y = w$. □

By Axiom 1, we have $\emptyset$, and by repeatedly using Axiom 3, we get $\{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}, \ldots$ *ad infinitum*. So, thus far, our theory (i.e., Axioms 1, 2, and 3) shows the existence of infinitely many sets. However, we have only shown the existence of sets with 'zero', 'one' and 'two' elements, where the *number words* 'zero', 'one' and 'two' indicate the *number* of elements in a set. More precisely, 'zero' elements means that there does not exist any elements, i.e., $\emptyset$. Set m has 'one' element if and only if $\exists x : (m = \{x\})$, this is also known as a *singleton* set. Finally, set m has 'two' elements if and only if $\exists x, y : (m = \{x, y\}) \wedge (x \neq y)$.

Remark. Note that we can form ordered triples, say (x, y, z), by setting it as $((x, y), z)$ which is a set by Axiom 3. And similarly, ordered quadruples, ordered quintuples, and so on.

Axiom 4 - Union set. $\forall x : \exists u : \forall y : ((y \in u) \Leftrightarrow (\exists s : (y \in s) \wedge (s \in x)))$
Given a set x. There exists a set u, whose elements are the elements of the elements of x. We denote this set u as, $\bigcup x$.

Given sets a, b, c. By Axiom 3 (pair set), we have $\{a, b\}$ and $\{c\}$, and so we also have $\{\{a, b\}, \{c\}\}$. So by Axiom 4 (union set), we get $\bigcup\{\{a, b\}, \{c\}\} = \{a, b, c\}$. Also, by Axiom 4, we have $\bigcup\{a, b\}$, which contains all the elements in a and b, we denote this as $a \cup b$, just as normal.

So now we can show that there exists a set with *more than* 'two' elements from Axiom 4 (union set). We can have 'three' elements by essentially unionising the set that has 'two' elements, call it m, with $\{m\}$. We will see later that this is known as the successor.

Remark. For better readability, we have for set y, the formula $\forall x \in y : P(x)$ is an abbreviation of $\forall x : (x \in y \Rightarrow P(x))$. And similarly we have $\exists x \in y : P(x)$ is an abbreviation of $\exists x : (x \in y) \wedge P(x)$.

Axiom 5 - Replacement schema.

$$\forall x : [\forall y, z, u : ((\phi(y, z) \wedge \phi(y, u)) \Rightarrow (z = u)) \Rightarrow (\exists m : \forall t : (t \in m) \Leftrightarrow (\exists s \in x : \phi(s, t)))]$$

Given a set x and a formula $\phi(y, t)$, where for any y there exists at most one t such that $\phi(y, t)$. Then there exists a set m of all those t, written as $\{t : \exists y \in x : \phi(y, t)\}$.

Essentially, this axiom says that the range of a function exists (see below for precise definitions). It is clear that this set m is unique from Axiom 2 (exercise).

The following result is known as the *separation axiom schema* or the *restricted comprehension axiom schema*.

Proposition 3.6. *Given a formula $P(y)$ and a set x, we then have the set of all $t \in x$ such that $P(t)$, i.e., $\{t \in x : P(t)\}$.*

Proof. Let x be a set, $P(y)$ be a formula, and $\phi(y,t)$ denote the formula $y = t \wedge P(t)$. Then we have that for any y we have $y = t$ if $P(t)$ holds. By Axiom 5, we have the set $\{t : \exists y \in x : \phi(y,t)\}$, which we can rewrite as $\{t \in x : P(t)\}$. □

Remark. If we assume that the universe of sets is not empty (some texts add this as an axiom, namely "Axiom 0" and it is stated by $\exists x : x = x$), then we can see that Axiom 1 is implied from the other axioms now. Where from Axiom 5 and Proposition 3.6, we can see $\exists x : \forall y : (y \notin x)$. And so by Axiom 2, we have $\exists! x : \forall y : (y \notin x)$.

Note that Axiom 5 and Proposition 3.6 (comprehension schema) are axiom *schemes*, so they provides an axiom for each formula. Thus we have an *infinite* number of axioms.

Now given a set x. We have the *intersection* of x, denoted by $\bigcap x$, and is defined as $\{a \in \bigcup x : \phi(a,x)\}$, where $\phi(a,x) := \forall b \in x : a \in b$. This is a set by Proposition 3.6. We say that the sets a and b are *disjoint* if $a, b \in x$ and $\bigcap x = \emptyset$. We can also define the *complement* of set $x \subseteq y$, as $y \setminus x := \{u \in y : u \notin x\}$, which is a set by Proposition 3.6. So, given sets x and y, we have the intersection, $x \cap y := \{a \in x : a \in y\}$.

Now we can see that the set of all sets does not exist, i.e., the universe of sets. Since if we suppose there does exist one, then $U = \{x : x = x\}$ is the universe of sets. So, following *Russel's paradox*, we can let $A = \{x \in U : x \notin x\}$, which is a set by Proposition 3.6. So we know $A \in U$, but then we get $A \notin A$ if and only if $A \in A$. And so *reductio ad absurdum*. Thus, the universe of sets is a proper class.

Axiom 6 - Power set. $\forall x : \exists y : \forall a : (a \in y) \Leftrightarrow (a \subseteq x)$
Given a set x, there exists a set y, denoted as $\mathcal{P}(x)$, whose elements are the subsets of x.

Example 3.7. $\mathcal{P}(\emptyset) = \{\emptyset\}$
$\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$
$\mathcal{P}(\{\emptyset, \{\emptyset\}\}) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$

Given sets X and Y, we can define the *Cartesian product* set $X \times Y$, which satisfies $X \times Y \subseteq \mathcal{P}(\mathcal{P}(\bigcup\{X, Y\}))$. Which as we know is the set of ordered pairs of X and Y, i.e., $(x,y) = \{\{x\}, \{x,y\}\}$ where $x \in X$ and $y \in Y$. More precisely, we can write,

$$X \times Y := \left\{ z \in \mathcal{P}\left(\mathcal{P}\left(\bigcup\{X,Y\}\right)\right) : p(X,Y) \right\},$$

where $p(X,Y) := \exists u : \exists v \left((u \in X) \wedge (v \in Y) \wedge (z = (u,v))\right)$. This is obviously a set, by Axioms 3, 4, 5, and 6.

Given sets A and B, a *relation* R between A and B is $R \subseteq A \times B$, i.e., it is a relation if it only consists of ordered pairs. A relation R on A is $R \subseteq A \times A$. A relation $f \subseteq A \times B$ is a *map*, or *function*, from A to B, if for each $a \in A$, there exists a unique $b \in B$ such that $(a,b) \in f$. We will write the map as $f : A \to B$, given by $a \mapsto f(a)$, where $f(a) = b$ such that $(a,b) \in f$.

Precisely, given a set f, it is a *function* from X to Y if $f \subseteq X \times Y$ and

$$\forall a : (a \in X) \Rightarrow (\exists b : (b \in Y) \wedge ((a,b) \in f) \wedge ((\forall z : (a,z) \in f) \Rightarrow (z = b))).$$

We define the *domain* and *range* of any set f respectively as

$$\text{domain of } f = \left\{ a \in \bigcup\bigcup f : \exists b : (a,b) \in f \right\},$$

$$\text{range of } f = \left\{ b \in \bigcup\bigcup f : \exists a : (a,b) \in f \right\}.$$

Using these definitions, we can instead define f to be a *function* if and only if f is a relation and $\forall x \in \text{domain of } f : \exists! y : (x,y) \in f$. And so we denote the unique y by $f(x)$. If we have $f \subseteq X \times Y$ and the domain of f is equal to X and the range of f is a subset of Y, then we can write the map as $f : X \to Y$.

Now the *image* of A under f, $\text{im}_f(A)$ (usually written as $f[A]$) is the set $\{f(a) : a \in A\}$. Now, we say that the map is *injective* if $\forall a, b \in A : f(a) = f(b) \Rightarrow a = b$. Or *surjective* if $f(A) = B$ (or range of f is B). Or *bijective* if it is both injective and surjective.

Given maps $f : A \to B$ and $g : B \to C$, we have the *composition map* $g \circ f : A \to C$, defined by $a \mapsto g(f(a))$. It is clear that this is a set and a map, more precisely we have,

$$g \circ f = \{(a,c) \in A \times C : \exists b \in B : ((a,b) \in f) \wedge ((b,c) \in g)\}.$$

It is easy to see that composition of maps is associative, and is left as an exercise, i.e., $h \circ (g \circ f) = (h \circ g) \circ f$ for $f : A \to B$, $g : B \to C$ and $h : C \to D$.

Now given a map $f : A \to B$. If f is a bijection, then we can define the *inverse map* f^{-1}, such that $f^{-1} \circ f = \text{id}_A$ and $f \circ f^{-1} = \text{id}_B$, where $\text{id}_A : A \to A$, $a \mapsto a$ and $\text{id}_B : B \to B$, $b \mapsto b$ and are known as the *identity map*. It is easy to see that the inverse map exists if and only if the map is bijective.

We can also define the *preimage* of a map $f : A \to B$ for $C \subseteq B$, $f^{-1}(C) = \{a \in A : f(a) \in C\}$.

We can also define the set of all functions on set X into set Y by $Y^X := \{f \in \mathcal{P}(X \times Y) : (f \text{ is a map}) \wedge (\forall x \in X : \exists y \in Y : (x,y) \in f)\}$.

Definition 3.8. Given a set A and a relation R,

(i) R is *reflexive* on A if and only if $\forall a \in A : aRa$.

(ii) R is *symmetric* on A if and only if $\forall a, b \in A : aRb \Leftrightarrow bRa$

(iii) R is *antisymmetric* on A if and only if $\forall a, b \in A : (aRb \wedge bRa) \Rightarrow (a = b)$

(iv) R is *asymmetric* on A if and only if $\forall a, b \in A : aRb \not\Rightarrow bRa$

(v) R is *transitive* on A if and only if $\forall a, b, c \in A : (aRb \wedge bRc) \Rightarrow aRc$

(vi) R satisfies *linearility* (or *trichotomy*) on A if and only if $\forall x, y \in A : (xRy) \vee (yRx) \vee (x = y)$

We call a relation R an *equivalence relation* on set A if and only if R is reflexive, symmetric and transitive on A. We can then define the *equivalence class* of $a \in A$ as $[a] = \{b \in A : aRb\}$. We can also then define the *quotient set* of A by R as, $A/R := \{[a] : a \in A\}$. Exercise: show that the quotient set is indeed a set. It is also easy to see that the equivalence class of two elements, say $[a]$ and $[b]$, are either related (and so are equal) or they are disjoint, i.e., either $[x] = [y]$ or $[x] \cap [y] = \emptyset$.

A relation R on a set A is a *weak partial order* on A if and only if R is reflexive, transitive, and antisymmetric on A. We say that it is a *strict partial order* if R is not reflexive, transitive, and asymmetric on A. We say that R is a *weak total order* if either xRy or yRx for any $x, y \in A$. We say that R is a *strict total order* if for any $x, y \in A$ such that $x \neq y$, then either xRy or yRx.

Remark. Observe that in our definition above, we could have said that the relation R is a strict partial order if and only if R is not reflexive and transitive on A. This is because these two properties imply xRy and yRx cannot both hold for all $x, y \in A$. We see this by assuming that for $x, y \in A$, let xRy and yRx. Then by transitivity, we see xRx, a contradiction to the not reflexive property.

Axiom 7 - Infinity. $\exists x : (\emptyset \in x) \wedge (\forall y : (y \in x) \Rightarrow (y \cup \{y\} \in x))$
Given a set y. There exists a set that contains the empty set, y, and $y \cup \{y\}$.

We can define the natural numbers using this axiom (see next chapter).

Axiom 8 - Choice. $\forall x : P(x) \Rightarrow \exists y : \forall a \in x : \exists! b \in a : a \in y$, where $P(x) \Leftrightarrow (\exists a : (a \in x) \wedge (\forall a : \forall b : (a \in x) \wedge (b \in x)) \Rightarrow \bigcap\{a, b\} = \emptyset)$
Given a set x, whose elements are non-empty and mutually disjoint. There exists a set y which contains exactly one element of each element of x.

We will discuss this axiom a bit further in the next chapter (§4.I.II).

Axiom 9 - Foundation. $\forall x : \exists a : (a \in x) \Rightarrow (\exists y \in x : \bigcap\{x, y\} = \emptyset)$
Every non-empty set has a minimal element. More precisely, every non-empty set x contains an element y that has none of its elements in common with x.

This axiom tells us that there exists no set that contains itself as an element. So, given set x, we have $x \notin x$. Otherwise, we would have $x \in x$ and $x \in \{x\}$, which is a contradiction to Axiom 9, since $x \cap \{x\} = x \neq \emptyset$.

4 Construction of Number Systems

4.I Construction of $\mathbb{N}$.

In this chapter we will construct the number systems $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ from a purely set-theoretic way.

This is based on the presentations given in [5], [9], [11], [12], [15], [21], [25], and [32].

Remark. The original Zermelo axioms had the following axiom for infinity

$$\exists x : (\emptyset \in x) \wedge (\forall y : (y \in x) \Rightarrow (\{y\} \in x))$$

Id est, given a set y, there exists a set that contains the empty set, y, and $\{y\}$. And so via this definition, we would have the set of natural numbers $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \ldots\}$, so each number represented by Arabic numerals would be defined as $0 = \emptyset, 1 = \{\emptyset\}, 2 = \{\{\emptyset\}\}, \ldots$. This is known as *Zermelo's ordinals*. This, however, was not a good definition since the set membership $\in$ is not transitive via this definition of natural numbers. For instance, $\emptyset \in \{\emptyset\}$ and $\{\emptyset\} \in \{\{\emptyset\}\}$, however $\emptyset \notin \{\{\emptyset\}\}$.

From the infinity axiom, there exists a set x such that $\emptyset \in x$, and if $y \in x$, then $y \cup \{y\} \in x$. We call any set that has these properties of the infinity axiom, an *inductive* set. And it turns out that if y and x are both sets that have the inductive property, then the set $y \cap x$ also has the inductive property. This is because if we have $\emptyset \in x$ and $\emptyset \in y$, then $\emptyset \in y \cap x$. And so if we have $z \in y \cap x$, then by the definition of intersection, we have $z \in y$ and $z \in x$, and so by the inductive property we have $z \cup \{z\} \in y$ and $z \cup \{z\} \in x$, and hence we get $z \cup \{z\} \in y \cap x$.

Definition 4.1. Given any inductive set x (from the infinity axiom), we define the set of *natural numbers* as $\mathbb{N} := \bigcap\{z \in \mathcal{P}(x) : z \text{ is inductive}\}$, i.e., the intersection of all inductive subsets of any inductive set.

We can see that this set $\mathbb{N}$ is *de facto* inductive. This follows from the above, $y \cap x$ is inductive if both y and x are inductive, and so $\emptyset$ is in $\mathbb{N}$ since it has to be in every inductive set, and so has to be in the intersection of all subsets of an inductive set x. And if we have $y \in \mathbb{N}$, then we have $y \in z$ for all $z \subseteq y$ such that z is inductive. And thus for any z, we get $y \cup \{y\} \in z$, and so will be in the intersection $\mathbb{N}$.

We call $a \in \mathbb{N}$, a *natural number*. Now we have $(\emptyset \in \mathbb{N})$, $(\emptyset \cup \{\emptyset\} = \{\emptyset\} \in \mathbb{N}), \ldots$. So $\mathbb{N} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots\}$. Then we can define the *numbers* via Arabic numerals, i.e., $0 := \emptyset$, $1 := \{\emptyset\}$, $2 := \{\emptyset, \{\emptyset\}\}$, and so on.

We can then define the *successor* of a set in $\mathbb{N}$ as $S : \mathbb{N} \to \mathbb{N}$, $n \mapsto n \cup \{n\}$. So, $S(0) = S(\emptyset) = \emptyset \cup \{\emptyset\} = \{\emptyset\} = 1$, $S(1) = S(\{\emptyset\}) = \{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\} = 2$, and so on. It is easy to see if $S(m) = S(n)$ then $m = n$, since if $m \cup \{m\} = n \cup \{n\}$, then either $m \in n$ or $m \in \{n\}$ (which means $m = n$) and $n \in m$ or $n \in \{m\}$ (i.e., $n = m$). And if we assume $m \neq n$, then $m \in n$ and $n \in m$, which is a contradiction.

We can also define the *predecessor* of a set in $\mathbb{N}$ as $P : \mathbb{N} \setminus \{\emptyset\} \to \mathbb{N}$ the inverse successor map given by $n \cup \{n\} \mapsto n$. Or we can also say that the predecessor is the map $n \mapsto m$, such that $m \in n$ and $S(m) = n$. So, $P(3) = P(\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}) = P(\{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\}) = \{\emptyset, \{\emptyset\}\} = 2$, and likewise $P(2) = 1$ and so on.

This next result is known as *definition by recursion* (see [12] for proof).

Theorem 4.2. *Given a set A, a map $g : A \to A$, and $a \in A$. Then there exists a unique map $f : \mathbb{N} \to A$ such that $f(0) = a$ and $f(S(n)) = g(f(n))$ for all $n \in \mathbb{N}$.* □

So, now we can *recursively* define S^n as $S \circ S^{P(n)}$ if $n \in \mathbb{N} \setminus \{\emptyset\}$, and $S^{\emptyset} := \mathrm{id}_{\mathbb{N}}$.

For any subset $z \subseteq \mathbb{N}$, if z has the inductive property, then $\mathbb{N} \subseteq z$ since it is the intersection of all inductive sets, and so $z = \mathbb{N}$. This provides a basis for "*proofs by induction.*"

Now we define the relation $<$ on $\mathbb{N}$. So for any $a, b \in \mathbb{N}$, we say $a < b$ if and only if $a \in b$. For instance, $1 < 3$ since $\{\emptyset\} \in \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$. We also define the relation $\leq$, which for $a \leq b$ is just $(a < b) \vee (a = b)$. The relation $<$, is transitive, not reflexive and is linear. We will show this by induction. But, first we will start with a result that will help us.

Remark. For the sake of completeness, if $a < b$ then $b > a$, likewise if $a \leq b$ then $b \geq a$.

Lemma 4.3. *For any $a, b \in \mathbb{N}$, if $a < b$, then $S(a) < S(b)$.*

Proof. We want to show that the set $z = \{b \in \mathbb{N} : (a < b) \Rightarrow (S(a) < S(b))\} \subseteq \mathbb{N}$ for $a \in \mathbb{N}$ is inductive, i.e., proof by induction on b. Clearly, $\emptyset \in z$, since $a \in \emptyset$ is false, and so *ex falso sequitur quodlibet*, and so $S(a) < S(\emptyset)$. Now if $b \in z$, then $(a < b) \Rightarrow (S(a) < S(b))$. So we want to show $(a < S(b)) \Rightarrow (S(a) < S(S(b))$. And since $a < b$, then $a < S(b)$ so $a \in S(b)$, and so either $a \in b$ or $a \in \{b\}$. If $a \in \{b\}$, then $a = b$, and so $S(a) = S(b) \in S(S(b))$. Now if $a \in b$, then we know $S(a) \in S(b)$, and so clearly $S(a) \in S(S(b))$. So, the result then follows. □

Remark. Note that in the preceding (and the following) result, we used $<$ interchangeably with $\in$. This is to emphasise the definition.

Proposition 4.4. *The relation* $<$ *is not reflexive, transitive and linear on the set of natural numbers, i.e.,* $\in$ *is not reflexive, transitive and linear on* $\mathbb{N}$.

Proof. The not-reflexive property follows from the axiom of foundation, where we saw that no set can contain itself as an element. And so for any $a \in \mathbb{N}$, we have $a \not< a$ since $a \notin a$.

For the transitive property, we want to show that for all $c \in \mathbb{N}$, if we have $a < b < c$, then we have $a < c$, where $a, b \in \mathbb{N}$. So, we want to show the following set is inductive, $z = \{c \in \mathbb{N} : (a < b < c) \Rightarrow (a < c)\} \subseteq \mathbb{N}$. Clearly, $\emptyset \in z$ since $a < b < \emptyset$ is false and so *ex falso sequitur quodlibet*. If we have $c \in z$, then $a < b < c$ and so $a < c$. So, then if $a < b < S(c)$, then we have $b \in S(c)$, which by the definition of unions, we have $b \in c$ or $b \in \{c\}$. If it is the latter, then $b = c$ (and the result follows), otherwise $b \in c$, and we already know $a < c$ if $a < b < c$, so we have $S(c) \in z$. Thus ends the proof by induction on $\mathbb{N}$, and hence shows the transitive property of the relation.

For the linearality of $<$, we will also want to use proof by induction on $\mathbb{N}$. So, we want to show that for all $a \in \mathbb{N}$ we will have $a < b$ or $a = b$ or $b < a$, where $b \in \mathbb{N}$. So, we let $z = \{a \in \mathbb{N} : (a < b) \vee (a = b) \vee (b < a)\} \subseteq \mathbb{N}$. Clearly, we have $\emptyset \in z$ since $\emptyset < b$ (since $\emptyset$ contains all other natural numbers) or $b = \emptyset$, and so the result then follows. If we have $a \in z$, then we have that either $a < b$ or $b < a$ or $a = b$. So, if $b < a$, then $b < S(a)$ and so $S(a) \in z$. If $a = b$, then $b < S(a)$, and so the result follows. Now if $a < b$, then $b \neq \emptyset$, and so $b = S(e) = e \cup \{e\}$ for some $e \in \mathbb{N}$. And so $a < e \cup \{e\}$, which means $a \in e \cup \{e\}$, so either $a \in e$ or $a \in \{e\}$. If $a \in \{e\}$, then $a = e$, and so $S(a) = S(e) = b$. If $a \in e$, then by the Lemma above, we have $S(a) < S(e) = b$. And so the result then follows. □

Using the above definitions, we can then define *addition* on $\mathbb{N}$ as $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, by $(m, n) \mapsto +(m, n) = S^n(m)$. We will write it as the usual notation $m+n$. So, for instance, $1 + 1 = +(1, 1) = +(\{\emptyset\}, \{\emptyset\}) = S^{\{\emptyset\}}(\{\emptyset\}) = S \circ S^{P(\{\emptyset\})}(\{\emptyset\}) = S(S^{P(\emptyset \cup \{\emptyset\})}(\{\emptyset\})) = S(S^{\emptyset}(\{\emptyset\})) = S(\mathrm{id}_{\mathbb{N}}(\{\emptyset\})) = S(\{\emptyset\}) = \{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\} = 2$. Thus we have shown from a purely set theoretic method $1 + 1 = 2$. From the successor definition we can easily see $m + \emptyset = S^{\emptyset}(m) = \mathrm{id}_{\mathbb{N}}(m) = m$ for any $m \in \mathbb{N}$. Similarly, we also see $\emptyset + m = m$ (exercise). So $\emptyset = 0$ is a neutral element under addition on $\mathbb{N}$.

Example 4.5.

$$
\begin{aligned}
2 + 2 = +(2, 2) = +(\{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}) &= S^{\{\emptyset, \{\emptyset\}\}}(\{\emptyset, \{\emptyset\}\}) \\
&= S \circ S^{P(\{\emptyset, \{\emptyset\}\})}(\{\emptyset, \{\emptyset\}\}) \\
&= S(S^{P(\{\emptyset\} \cup \{\{\emptyset\}\})}(\{\emptyset, \{\emptyset\}\})) \\
&= S(S^{\{\emptyset\}}(\{\emptyset, \{\emptyset\}\})) \\
&= S(S \circ S^{P(\{\emptyset\})}(\{\emptyset, \{\emptyset\}\})) \\
&= S(S(S^{P(\emptyset \cup \{\emptyset\})}(\{\emptyset, \{\emptyset\}\}))) \\
&= S(S(S^{\emptyset}(\{\emptyset, \{\emptyset\}\}))) \\
&= S(S(\mathrm{id}_{\mathbb{N}}(\{\emptyset, \{\emptyset\}\}))) \\
&= S(S(\{\emptyset, \{\emptyset\}\})) \\
&= S(\{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\}) \\
&= S(\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\})
\end{aligned}
$$

$$\begin{aligned}&= \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \cup \{\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}\\ &= \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}\\ &= 4\end{aligned}$$

From this point forward, we will use the numbers we defined, i.e., $0, 1, 2, 3, \ldots$, instead of $\emptyset, \{\emptyset\}, \ldots$, for better readability and comprehension.

Using addition, we can also define *multiplication* on $\mathbb{N}$ as $\cdot : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, by $a \cdot b = a + (a \cdot P(b))$ and $a \cdot 0 = 0$. For example, $4 \cdot 3 = 4 + (4 \cdot 2) = 4 + (4 + (4 \cdot 1)) = 4 + (4 + (4 + (4 \cdot 0))) = 4 + (4 + (4 + (0))) = 4 + (4 + 4) = 4 + (S^4(4)) = 4 + 8 = S^4(8) = 12$ (we leave the details of addition to the reader). This is well-defined. We can clearly see that for any m, $m \cdot 1 = m + (m \cdot 0) = m$. So, 1 is the neutral element under multiplication. But there are no multiplicative inverses, obviously.

We will now show associativity of addition, and we leave the rest of the following as an exercise.

Proposition 4.6. *For all $a, b, c \in \mathbb{N}$, we have that addition is associative and commutative, i.e., $a + (b + c) = (a + b) + c$, and $a + b = b + a$. We also have that multiplication is associative and commutative, i.e., $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ and $a \cdot b = b \cdot a$. We also have the distributive property, i.e., $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.*

Proof. We want to show $z = \{c \in \mathbb{N} : a + (b + c) = (a + b) + c\} \subseteq \mathbb{N}$ for $a, b \in \mathbb{N}$ is inductive. So, we have $\emptyset \in z$, since $a + (b + \emptyset) = a + S^{\emptyset}(b) = a + \mathrm{id}_{\mathbb{N}}(b) = a + b$, and $(a + b) + \emptyset = S^{\emptyset}(a + b) = \mathrm{id}_{\mathbb{N}}(a + b) = a + b$, and the result follows. Now if $c \in z$, then $a + (b + c) = (a + b) + c$. We want to show $a + (b + S(c)) = (a + b) + S(c)$. So,

$$\begin{aligned}a + (b + S(c)) = a + S^{S(c)}(b) &= a + S(S^c(b)) = a + S(b + c)\\ &= S^{S(b+c)}(a) = S(S^{b+c}(a))\\ &= S(a + (b + c))\\ &= S((a + b) + c), \text{ by assumption}\\ &= S(S^c(a + b)) = S^{S(c)}(a + b)\\ &= (a + b) + S(c)\end{aligned}$$

Thus completes the proof by induction. □

We also define *exponentials* on $\mathbb{N}$ as $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$, by $a^b = a^{P(b)} \cdot a$ and $a^0 = 1$.

Example 4.7.

$$\begin{aligned}2^2 &= 2^{P(2)} \cdot 2\\ &= 2^1 \cdot 2 = (2^{P(1)} \cdot 2) \cdot 2\\ &= (2^0 \cdot 2) \cdot 2\\ &= (1 \cdot 2) \cdot 2 = 2 \cdot 2\\ &= 2 + (2 \cdot P(2)) = 2 + (2 \cdot 1) = 2 + 2\\ &= 4, \text{ from above example}\end{aligned}$$

The following is left as an exercise.

Proposition 4.8. *For all* $a, b, c \in \mathbb{N}$*, we have* $a^{b+c} = a^b \cdot a^c$*,* $(a^b)^c = a^{bc}$*, and* $(a \cdot b)^c = a^c \cdot b^c$*.* □

One who is familiar with the *Peano Axioms* (see below), the non-set-theoretical method of forming the natural numbers, can see that we have already proven that our constructed set $\mathbb{N}$ satisfies all of the axioms.

Remark. Peano Arithmetic is a first-order axiomatic system, its language is the *language of arithmetic*, where it is essentially the first-order logic (predicate logic) with the constant 0, the relations $+, \cdot$ and the predicate function S such that $S(n) = n + 1$. The universe of discourse in this language is $\mathbb{N}$. The axioms are as follows.

(i) $\exists 0 \in \mathbb{N}$
there exists 0 in the natural numbers

(ii) $\forall x : S(x) \neq 0$
there is no natural number with 0 as its successor

(iii) $\forall x, y : (S(x) = S(y)) \Rightarrow (x = y)$

(iv) if n is a natural number, then $S(n)$ is a natural number

(v) if we have a set y such that 0 is in y and for every natural number n, if n is in y then $S(n)$ is in y (i.e., it is inductive)

If we consider the union of (**ZFC** axioms without the infinity axiom) and (the negation of the infinity axiom, i.e., there does not exist a set that is inductive), then this theory, call it **TZ**, and the Peano Arithmetic theory, call it **PA**, are basically the same. That is **TZ** proves any sentence if and only if **PA** proves the translation of that sentence (from set theory to arithmetic). And **PA** proves any sentence if and only if **TZ** proves the translation of that sentence (from arithmetic to set theory) ([25]).

4.I.I Brief on Cardinals

Here, we will define the notion of *finite* and *countability*. But we will not include any proofs. The curious reader is directed to [12].

Given sets A and B. They are *equinumerous* if there exists a bijection from A to B. We write $|A| = |B|$ when A and B are equinumerous. Intuitively, we can think of this as the *size* of the set, i.e., the number of elements. We call $|A|$ the *cardinality* of A. We say A has cardinality less than or equal to B if there exists an injection between A to B, and write it as $|A| \leq |B|$. Finally, if $|A| \leq |B|$, but $|A| \neq |B|$, then $|A| < |B|$. We also have the result that if $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$. This result is called the *Schroder-Bernstein theorem*.

Recall that the number of elements in a natural number n is *de facto* n. So we say a set A is *finite*, if there exists some natural number n such that $|A| = n$.

Is the set of natural numbers $\mathbb{N}$ *infinite*? Yes. We can first show that the cardinality for any natural number n is not equal to any *proper subset* (i.e., a subset without equality) of n. And so we can see a bijection from $\mathbb{N}$ to $S(\mathbb{N})$, i.e., the successor of all natural numbers. And clearly $S(\mathbb{N})$ is not equal to $\mathbb{N}$. Thus we have an infinite set.

If the sets A and B are disjoint then $|A| + |B| = |A \cup B|$. We can also define multiplication on cardinals, i.e. $|A| \cdot |B| = |A \times B|$. These operations are well-defined (exercise).

We say a set C is *countable* if $|C| = |\mathbb{N}|$. We have that a subset of a countable set is countable. We also have the union of a finite number of countable sets is countable. We also have that the Cartesian product of a finite number of countable sets is countable. We call $\aleph_0$, pronounced as *aleph-nought*, the cardinal number of any countable set. So $\aleph_0 = |\mathbb{N}|$.

4.I.II Axiom of Choice (AC) and its equivalences

Given that the relation $\leq$ is a weak total order on A. We say a *chain* in a partially ordered set A is a totally ordered subset of A, i.e., $B \subseteq A$ is a chain if $\forall a, b \in B : (a \leq b) \vee (b \leq a)$. And we say $u \in A$ is an *upper bound* of chain B if for all $a \in B$ we have $a \leq u$.

Theorem 4.9. *Working in **ZFC**, we have the following equivalences.*

(i) The Axiom of Choice (Axiom 8)

(ii) Zorn's Lemma: Let P be a partially ordered set. Suppose that every totally ordered subset has an upper bound. Then the set P has a maximal element.

(iii) Well-Ordering Principle: Every set can be well-ordered.

(iv) Every vector space has a basis. □

You can find the proof of the theorem in [12], [15], and [32].

4.II Construction of $\mathbb{Z}$.

The natural numbers $\mathbb{N}$ do not contain any additive inverses. Non-set theoretically, one might say that the natural numbers is a subset of non-negative integers. This however is not a rigorous definition since you can define the negative integers as additive inverses of the natural numbers.

In order to define the integers, we need to define a notion of *subtraction*. So we define a relation R on $\mathbb{N} \times \mathbb{N}$, by $(m, n)R(p, q) :\Leftrightarrow m + q = p + n$. Intuitively, this means that each pair represents subtraction, i.e., $(m, n)R(p, q)$ means $m+q = p+n$, which by *manipulation* we get $m - n = p - q$. So (m, n) essentially represents $m - n$, where $-$ is a symbol. This relation R is an equivalence relation (exercise). So, $(2, 1)R(3, 2)$, likewise $(3, 2)R(4, 3)$. So, we have the equivalence classes $[(2, 1)] = [(3, 2)] = \cdots = [(1, 0)]$.

Definition 4.10. The *integer set* $\mathbb{Z}$ is defined as the quotient set $(\mathbb{N} \times \mathbb{N})/R$. We call $a \in \mathbb{Z}$ an *integer*. All integers are of the form $n = [(n, 0)]$, or its additive inverse $-n = [(0, n)]$ for $n \in \mathbb{N}$.

So, for example, $[(0, 0)] = [(m, m)]$ for all $m \in \mathbb{N}$. So $0 \in \mathbb{Z}$. We also have $[(0, 1)] = [(20, 21)] = [(m, m + 1)] = -1$. And $[(0, 2)] = [(43, 45)] = [(m, m + 2)] = -2$, and so on.

At the start of the section when we gave a non-set theoretical definition of the integers we claimed $\mathbb{N} \subseteq \mathbb{Z}$, since, intuitively, for all numbers $m \in \mathbb{N}$, we have that it is also in $\mathbb{Z}$. However, by definition, they contain different elements, for example, $1_{\mathbb{N}} = \{\emptyset\} \in \mathbb{N}$ is not the same as $1_{\mathbb{Z}} = [(\{\emptyset\}, \emptyset)] = [(\{\{\emptyset\}\}, \{\emptyset\})] = \cdots \in \mathbb{Z}$. So, $\mathbb{N}$ is not a subset of $\mathbb{Z}$, however it is *embedded* into $\mathbb{Z}$ with the *inclusion map* ι, given by $\iota : \mathbb{N} \hookrightarrow \mathbb{Z}$, by $n \mapsto [(n, \emptyset)]$. So that the embedded $\mathbb{N}$ in $\mathbb{Z}$ is then a subset of $\mathbb{Z}$.

We define *addition* on $\mathbb{Z}$ with a map $+_{\mathbb{Z}} : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ given by $[(m, n)] +_{\mathbb{Z}} [(p, q)] := [(m + p, n + q)]$, for any $m, n, p, q \in \mathbb{N}$ and where $+$ is addition on $\mathbb{N}$. This is well-defined, since for any integers $[(a, b)], [(c, d)], [(e, f)], [(g, h)]$ such that $[(a, b)] = [(e, f)]$ and

$[(c,d)] = [(g,h)]$, then $[(a,b)] +_{\mathbb{Z}} [(c,d)] = [(e,f)] +_{\mathbb{Z}} [(g,h)]$. This is true because we have $a+f = b+e$ and $c+h = d+g$, so then by associativity and commutativity of $+$, we get $(a+c)+(f+h) = a+f+c+h = b+e+c+h = b+e+d+g = (b+d)+(e+g)$, so then $(a+c, b+d)R(e+g, f+h)$ and the result follows.

Let us add m with its additive inverse, $-m$. So, $m +_{\mathbb{Z}} (-m) = [(m,0)] +_{\mathbb{Z}} [(0,m)] = [(m,m)] = [(0,0)] = 0$. Clearly $0 = -0$. So now we can define *subtraction* on $\mathbb{Z}$ as $m -_{\mathbb{Z}} n := m +_{\mathbb{Z}} (-n)$ for any $m, n \in \mathbb{Z}$.

We define *multiplication* on $\mathbb{Z}$ by the map $\cdot_{\mathbb{Z}} : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$, given by $[(m,n)] \cdot_{\mathbb{Z}} [(p,q)] := [(m \cdot p + n \cdot q, m \cdot q + n \cdot p)]$, for any $m, n, p, q \in \mathbb{N}$ and where $\cdot$ is multiplication defined on $\mathbb{N}$. For example, $5 \cdot_{\mathbb{Z}} (-3) = [(5,0)] \cdot_{\mathbb{Z}} [(0,3)] = [(5 \cdot 0 + 0 \cdot 3, 5 \cdot 3 + 0 \cdot 0)] = [(0,15)] = -15$. This is well-defined (exercise).

Clearly, for any integer m we have, $1 \cdot_{\mathbb{Z}} [(m,0)] = [(1,0)] \cdot_{\mathbb{Z}} [(m,0)] = [(1 \cdot m + 0 \cdot 0, 1 \cdot 0 + 0 \cdot m)] = [(m,0)] = m$ and similarly we see $1 \cdot_{\mathbb{Z}} [(0,m)] = -m$. Also, $0 \cdot_{\mathbb{Z}} m = 0$. Now given any natural numbers m, n. We have $-m \cdot_{\mathbb{Z}} -n = [(0,m)] \cdot_{\mathbb{Z}} [(0,n)] = [(m \cdot n, 0)] = m \cdot_{\mathbb{Z}} n$.

Also note that multiplication and addition are both commutative and associative on $\mathbb{Z}$ (easy exercise).

We also define the relation $<$ on $\mathbb{Z}$ (which we denote $<_{\mathbb{Z}}$ for now), where given any integers $[(a,b)], [(c,d)]$, we say $[(a,b)] <_{\mathbb{Z}} [(c,d)]$ if $a+d < b+c$, where $+, <$ are defined on $\mathbb{N}$. This is well-defined and the proof is analogous to how $<_{\mathbb{Q}}$ is well-defined in the next section. We call an integer a, *positive* if $0 <_{\mathbb{Z}} a$, and *negative* if $a <_{\mathbb{Z}} 0$.

Note that it follows from section §4.I.I that the integers is a countable set. And so $|\mathbb{Z}| = |\mathbb{N}| = \aleph_0$.

An optional exercise to conclude this section would be to check that this indeed forms an ordered ring. And from this, we can see that the structure of the embedded $\mathbb{N}$ into $\mathbb{Z}$ is well preserved. So, just as in usual mathematics, we can regard $\mathbb{N}$ as the embedded $\mathbb{N}$ into $\mathbb{Z}$, so $\mathbb{N} \subseteq \mathbb{Z}$. And so when referring to the natural numbers and integers, we can use the same symbols (for instance, instead of $<_{\mathbb{Z}}$ and $<$, we can say $<$).

4.III Construction of $\mathbb{Q}$.

We will now construct the rational numbers from the integers. The idea here is to define the notion of *division*.

Definition 4.11. We first define a relation R on $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$, given by $(p,q)R(r,s) :\Leftrightarrow p \cdot s = q \cdot r$. Now we can define the *rational numbers* $\mathbb{Q}$ by the quotient set $(\mathbb{Z} \times \mathbb{Z} \setminus \{0\})/R$. We call $a \in \mathbb{Q}$, a *rational number*. We let the rational number $[(a,b)]$ be represented by $\frac{a}{b}$. For example, $\frac{5}{7} = [(5,7)] = [(10,14)]$.

Similarly to $\mathbb{N}$ being embedded into $\mathbb{Z}$, we also have $\mathbb{Z}$ embedded into $\mathbb{Q}$ with the inclusion map $\iota : \mathbb{Z} \hookrightarrow \mathbb{Q}$ given by $p \mapsto [(p,1)]$.

We now define the relation $<$ on $\mathbb{Q}$, and denote it only for now as $<_{\mathbb{Q}}$. For any rational numbers $[(a,b)]$ and $[(c,d)]$, we say $[(a,b)] <_{\mathbb{Q}} [(c,d)]$ if $a \cdot d < b \cdot c$. We see that this is well defined since if we have rational numbers $[(a,b)], [(c,d)], [(e,f)]$ and $[(g,h)]$ such that $[(a,b)] = [(c,d)]$ and $[(e,f)] = [(g,h)]$, then we have $[(a,b)] <_{\mathbb{Q}} [(e,f)]$ if and only if $[(c,d)] <_{\mathbb{Q}} [(g,h)]$. We will now show the only if part. The if part is left as an exercise.

Proof. This is clear since if $[(a,b)] <_{\mathbb{Q}} [(e,f)]$, then we have $a \cdot f < b \cdot e$. And we already

know $a \cdot d = b \cdot c$ and $e \cdot h = f \cdot g$. So then $(a \cdot f) \cdot (d \cdot g) < (b \cdot e) \cdot (b \cdot g)$. And also $(a \cdot f) \cdot (d \cdot g) = (b \cdot c) \cdot (f \cdot g) = (b \cdot c) \cdot (e \cdot h)$. Finally, we get $(b \cdot e) \cdot (c \cdot h) < (b \cdot e) \cdot (d \cdot g)$, and so since $<$ is well defined on $\mathbb{Z}$ and the embedded $\mathbb{N}$ in $\mathbb{Z}$, we have $c \cdot h < d \cdot g$. □

We can then also define addition and multiplication on $\mathbb{Q}$, by the maps $+_{\mathbb{Q}} : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ given by $[(p, q)] +_{\mathbb{Q}} [(r, s)] := [(p \cdot s + r \cdot q, q \cdot s)]$, and $\cdot_{\mathbb{Q}} : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ given by $[(p, q)] \cdot_{\mathbb{Q}} [(r, s)] := [(p \cdot r, q \cdot s)]$. Both $\cdot_{\mathbb{Q}}$ and $+_{\mathbb{Q}}$ are well-defined. We will now show $+_{\mathbb{Q}}$ is well defined, and we leave $\cdot_{\mathbb{Q}}$ to the reader as an easy exercise.

Proof. We need to show that for rational numbers $[(a, b)], [(c, d)], [(g, h)], [(e, f)]$ such that $[(a, b)] = [(c, d)]$ and $[(g, h)] = [(e, f)]$, then $[(a, b)] +_{\mathbb{Q}} [(e, f)] = [(b, c)] +_{\mathbb{Q}} [(g, h)]$. So we need to show $[((a \cdot f) + (b \cdot e), b \cdot f)] = [((c \cdot h) + (d \cdot g), d \cdot h)]$, i.e.,

$$(a \cdot f) + (b \cdot e) \cdot (d \cdot h) = (b \cdot f) \cdot ((c \cdot h) + (d \cdot g)).$$

Now $[(a, b)] = [(c, d)]$ means $a \cdot d = b \cdot c$, and $[(g, h)] = [(e, f)]$ means $g \cdot f = h \cdot e$. So, by noting that $+$ and $\cdot$ are commutative, associative and have distributive properties on $\mathbb{Z}$ and the embedded $\mathbb{N}$, we get

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

And so the result follows. □

Note that, similarly as before, we say that a rational number q is positive if $0 <_{\mathbb{Q}} q$, and negative if $q <_{\mathbb{Q}} 0$. And so if we have a rational number $p = [(a, b)]$, then we can define its negative $-p$ as $[(-a, b)]$ or $[(a, -b)]$, and, as usual, we write it as $-\frac{a}{b}$ instead of $\frac{-a}{b}$ or $\frac{a}{-b}$. So, we can now define *subtraction* on $\mathbb{Q}$. Given any $p, q \in \mathbb{Q}$, we have $p -_{\mathbb{Q}} q = p + (-q)$.

The last operation we need to define is *division* on $\mathbb{Q}$. The map $/_{\mathbb{Q}} : \mathbb{Q} \times \mathbb{Q} \setminus \{0\} \to \mathbb{Q}$ is given by $p/_{\mathbb{Q}} q := p \cdot_{\mathbb{Q}} q^{-1}$ for any rational numbers p and $q = [(a, b)]$, where q^{-1} is given by $[(b, a)]$.

Proposition 4.12. *Multiplicative inverses of non-zero elements in* $\mathbb{Q}$ *exist.*

Proof. Let $p = [(a, b)] \neq 0$. Then $p \cdot_{\mathbb{Q}} p^{-1} = [(a, b)] \cdot_{\mathbb{Q}} [(b, a)] = [(a \cdot b, b \cdot a)]$. And we can see $(a \cdot b, b \cdot a) R (1, 1)$, and so the result follows. □

Note that it follows from section 4.I.I and that integers is a countable set, that the rationals is also a countable set. So, $|\mathbb{Q}| = |\mathbb{Z}| = |\mathbb{N}| = \aleph_0$.

Similarly to the previous section, a nice exercise to conclude this section is to check that this forms an ordered field. And from there, we see that the structure of $\mathbb{Z}$ is well preserved as an ordered ring, and hence from now on, as usual, we regard $\mathbb{Z}$ as the embedded $\mathbb{Z}$ into $\mathbb{Q}$, so that $\mathbb{Z} \subseteq \mathbb{Q}$. And so we can likewise use the same symbols for $\mathbb{N}, \mathbb{Z}$, and $\mathbb{Q}$.

4.IV Construction of $\mathbb{R}$.

Now since $\mathbb{Q}$ is not complete (see Example 5.40), in the sense that it does not contain any "irrational number." As an exercise, show $x \notin \mathbb{Q}$ such that $x \cdot_{\mathbb{Q}} x = 2$, i.e., $\sqrt{2} \notin \mathbb{Q}$. We need to "complete" the set. There are two well-known approaches to construct the real numbers: Cantor's construction via rational Cauchy sequences and Dedekind cuts on $\mathbb{Q}$. We will construct our real numbers via rational Cauchy sequences as this will help us see things more clearly in the later chapters. We claim that the set $\mathbb{R}$ is the *completion* of $\mathbb{Q}$. Here completion essentially is the same as we will see in §5.IV, where a normed vector space is said to be complete if every Cauchy sequence converges, and a complete normed vector space $\hat{V}$ is said to be a completion of a normed vector space V, if V is a dense subset of $\hat{V}$.

So, we begin with defining the *absolute value* on $\mathbb{Q}$ by $|x|$ as the map $\mathbb{Q} \to \mathbb{Q}_{\geq 0}$ given by

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

$\mathbb{Q}_{\geq 0}$ here just means for all $x \in \mathbb{Q}$, $x \geq 0$. Note that we have the triangle inequality $|a + b| \leq |a| + |b|$.

A *sequence* in $\mathbb{Q}$ is a map $\mathbb{N} \to \mathbb{Q}$ given by $n \mapsto a_n$. We write (a_n) for the sequence $\{a_1, a_2, \ldots\}$. We say $(a_n) \to q$ (*converges*) if for all $\mathbb{Q} \ni \epsilon > 0$, there exists $N \in \mathbb{N}$, such that for all $n > N$, we have $|a_n - q| \leq \epsilon$.

Definition 4.13. A *Cauchy sequence* on $\mathbb{Q}$ is a sequence (x_n) such that for all $\mathbb{Q} \ni \epsilon > 0$, there exists $N \in \mathbb{N}$, such that $|x_m - x_i| < \epsilon$ for all $m, i \geq N$.

It is easy to see that if (a_n) is a convergent sequence in $\mathbb{Q}$, then it is Cauchy. It is left as an exercise (see Proposition 5.35). The converse, however, is not true: if (a_n) is a Cauchy sequence in $\mathbb{Q}$, then it is not necessarily convergent. An example of this could be the sequence

$$\left(\frac{3}{1}, \frac{31}{10}, \frac{314}{100}, \ldots, \frac{314159}{100000}, \ldots\right),$$

(which we know converges to π) which does not converge to any number in $\mathbb{Q}$. Another result that is useful, is that any Cauchy sequence is *bounded*, i.e., there exists $M \in \mathbb{Q}$ such that $|a_n| \leq M$ for all n (see Proposition 5.36).

And so now we define a relation R on the set of all Cauchy sequences in $\mathbb{Q}$, which is denoted as $\mathfrak{C}_{\mathbb{Q}}$, given by $(a_n)R(b_n) :\Leftrightarrow (a_n - b_n) \to 0$.

Let us show that this is in fact an equivalence relation.

Proof. For $(a_n) \in \mathfrak{C}_{\mathbb{Q}}$, we have $(a_n)R(a_n)$ since $|a_n - a_n| = 0$, and so it is reflexive. It is also symmetric, since for $(a_n), (b_n) \in \mathfrak{C}_{\mathbb{Q}}$, we have if $(a_n)R(b_n)$, then $|a_n - b_n| \to 0$ and so $|-1(b_n - a_n)| = |b_n - a_n| \to 0$, which means $(b_n)R(a_n)$. Finally, if we have $(a_n)R(b_n)$ and $(b_n)R(c_n)$ for $(a_n), (b_n), (c_n) \in \mathfrak{C}$, then $|a_n - b_n| \to 0$ and $|b_n - c_n| \to 0$. And so, by the triangle inequality, $|a_n - c_n| = |(a_n - b_n) + (b_n - c_n)| \leq |a_n - b_n| + |b_n - c_n| \to 0$, since for $\mathbb{Q} \ni \epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n > N$, $|a_n - b_n| < \epsilon/2$, and there exists $M \in \mathbb{N}$ such that for all $n >_{\mathbb{N}} M$, $|b_n - c_n| < \epsilon/2$; and by letting $n > \max(N, M)$, then $|a_n - c_n| \leq |a_n - b_n| + |b_n - c_n| < \epsilon/2 + \epsilon/2 = \epsilon$. Hence, $(a_n)R(c_n)$, which shows transitivity. And we are done. □

Definition 4.14. So, now we can define the set of real numbers $\mathbb{R}$ as $[(a_n)]$ for all $(a_n) \in \mathfrak{C}_{\mathbb{Q}}$. The equivalence class $[(a_n)]$ is called a *real number*.

We define multiplication and addition on $\mathbb{R}$, with maps $+_{\mathbb{R}} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $[(a_n)] +_{\mathbb{R}} [(b_n)] := [(a_n + b_n)]$, and $\cdot_{\mathbb{R}} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $[(a_n)] \cdot_{\mathbb{R}} [(b_n)] := [(a_n \cdot b_n)]$. It is clear that for any $r \in \mathbb{R}$, we have $r \cdot_{\mathbb{R}} 1 = r$, and $r \cdot_{\mathbb{R}} 0 = 0$. There also exists inverses for any $r \in \mathbb{R} \setminus \{0\}$, i.e., there exists some $r^{-1} \in \mathbb{R}$ such that $r \cdot_{\mathbb{R}} r^{-1} = 1$.

These operations are well defined, we will only show that multiplication on $\mathbb{R}$ is well defined, addition is left as an exercise. It is easy to see that these two operations are associative, commutative and distributive (left as an exercise to show).

Proof. To show that it is well defined, we need to show that if $[(a_n)] = [(c_n)]$ and $[(b_n)] = [(d_n)]$, then $[(a_n \cdot b_n)] = [(c_n \cdot d_n)]$. So, let $[(a_n)] = [(c_n)]$ and $[(b_n)] = [(d_n)]$, for $(a_n), (b_n), (c_n), (d_n) \in \mathfrak{C}_{\mathbb{Q}}$. Since they are Cauchy sequences, we know that they are bounded, so let $A, B \in \mathbb{Q}$ such that $|b_n| \leq B$ and $|c_n| \leq C$ for all $n \in \mathbb{N}$. Now, let $A = B + C$. Then we have $(a_n - c_n) \to 0$ and $(b_n - d_n) \to 0$, i.e., for any $\mathbb{Q} \ni \epsilon > 0$ there exists $N, M \in \mathbb{N}$ such that for all $n > N$, we have $|a_n - c_n| < \epsilon/C$ and for all $n > M$, we have $|b_n - d_n| < \epsilon/C$. Now we have $|a_n \cdot b_n - c_n \cdot d_n| = |b_n \cdot (a_n - c_n) + c_n \cdot (b_n - d_n)| \leq |b_n| \cdot |a_n - c_n| + |c_n| \cdot |b_n - d_n| < \frac{B\epsilon}{A} + \frac{C\epsilon}{A} = \epsilon$. And so $(a_n) \cdot b_n - c_n \cdot d_n \to 0$, which means $[(a_n \cdot b_n)] = [(c_n \cdot d_n)]$. And we are done. □

We also define the relation $<$ on $\mathbb{R}$ (denoted as $<_{\mathbb{R}}$ for now). Given real numbers $[(a_n)], [(b_n)]$, we say $[(a_n)] <_{\mathbb{R}} [(b_n)]$ to mean that there exists $\mathbb{Q} \ni \epsilon > 0$ and $N \in \mathbb{N}$ such that $a_n \leq b_n - \epsilon$ for all $n \geq N$. One can show that this relation is not reflexive (i.e., for all $[(a_n)] \in \mathbb{R}$, $[(a_n)] \not<_{\mathbb{R}} [(a_n)]$), transitive (i.e., for all $[(a_n)], [(b_n)], [(c_n)] \in \mathbb{R}$, if we have $[(a_n)] <_{\mathbb{R}} [(b_n)]$ and $[(b_n)] <_{\mathbb{R}} [(c_n)]$, then $[(a_n)] <_{\mathbb{R}} [(c_n)]$) and linear (i.e., for all $[(a_n)], [(b_n)] \in \mathbb{R}$, either $[(a_n)] <_{\mathbb{R}} [(b_n)]$ or $[(a_n)] = [(b_n)]$ or $[(b_n)] <_{\mathbb{R}} [(a_n)]$).

$\mathbb{Q}$ is embedded in $\mathbb{R}$ with the inclusion map $\iota : \mathbb{Q} \hookrightarrow \mathbb{R}$ given by $x \mapsto [(x, x, x, \ldots)]$.

The next result tells us $\mathbb{Q}$ is *dense* in $\mathbb{R}$.

Lemma 4.15. *For all $x \in \mathbb{R}$ and $\mathbb{Q} \ni \epsilon > 0$, there exists $a \in \mathbb{Q}$ such that $|x - a| < \epsilon$.*

Proof. Let $\mathbb{R} \ni x = [(x_i)]$ for $(x_i) \in \mathfrak{C}_{\mathbb{Q}}$ for all i. So then, for any $\mathbb{Q} \ni \epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n > N$, we have $|x_n - x_m| < \epsilon$. Now there exists $a \in \mathbb{Q}$, and $k > N$, so then we can define $a := [(x_k, x_k, \ldots)]$. And so $x - a = [(x_i)] - [(x_k)] = [(x_i - x_k)] < \epsilon$. Also, $a - x = [(x_k)] - [(x_i)] = [(x_k - x_i)] < \epsilon$. So, then $|x - a| < \epsilon$. □

A nice remark is that, $|[(x_n)]| := [(|x_n|)]$.

The set of real numbers $\mathbb{R}$ is uncountable ([12]). What is the cardinality of $\mathbb{R}$? *Cantor's theorem* states $\mathcal{P}(\mathbb{N})$ is an uncountable set. And from this we see $|\mathcal{P}(\mathbb{N})| = |2^{\mathbb{N}}|$, where $2^{\mathbb{N}}$ is the set of all functions on $\mathbb{N}$ into 2. So a general corollary to this is $|\mathcal{P}(X)| = |2^X|$ for any set X. And, since the real numbers is a subset of the Cartesian product of the power set of the rational numbers, then $|\mathcal{P}(\mathbb{Q}) \times \mathcal{P}(\mathbb{Q})| = 2^{\aleph_0}$. And thus, we can conclude $|\mathbb{R}| = 2^{\aleph_0}$.

So, we can now conclude this section by saying $\mathbb{R}$ forms an ordered field (exercise). And so considering the embedded $\mathbb{Q}$ into $\mathbb{R}$, we have $\mathbb{Q} \subseteq \mathbb{R}$. So we can essentially replace any symbol from $\mathbb{Q}$ to $\mathbb{R}$. And we can show $\mathbb{R}$ is complete, by showing that any Cauchy sequence of real numbers converges to a real number (see Proposition 5.41).

4.V Construction of $\mathbb{C}$.

We can define $\mathbb{C}$ as the ordered pair (a, b) of $\mathbb{R} \times \mathbb{R}$, where $(a, b) := a + bi$. Addition and multiplication are the maps $+_{\mathbb{C}}, \cdot_{\mathbb{C}} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ given by $(a, b) +_{\mathbb{C}} (c, d) := (a + c, b + d)$ and $(a, b) \cdot_{\mathbb{C}} (c, d) := (a \cdot c - b \cdot d, b \cdot c + a \cdot d)$. We can then define the *conjugate* as $\bar{z} := (a, -b) = a - bi$, where $z := (a, b) = a + bi$. We see $i^2 = i \cdot i = (0, 1) \cdot (0, 1) = (-1, 0) = -1$.

This completes the first part (*Logic & the Foundations of Mathematics*) of this essay.

5 Normed Spaces

In this chapter we will learn the basics of normed vector spaces, metric spaces and toplogical spaces. Readers who are not comfortable with convergence and continuity in metric and topological spaces should find this easy to follow.

This chapter is based on the presentations of [1], [4], [22], [24], [27], [28], [33], [38], [40], and [45].

We let $\mathbb{F}$ denote $\mathbb{R}$ or $\mathbb{C}$ throughout this essay.

5.I Norms, Metrics and Topology

To do analysis on a vector space we need to introduce the concept of a normed vector space. In this section, we will see how any normed vector space is a metric space, which is a topological space.

Definition 5.1. Let V be a $\mathbb{F}$-vector space. A *norm* on V is a function $\|\cdot\| : V \to \mathbb{R}$ such that

(i) $\|v\| \geq 0$ for all $v \in V$ with equality if and only if $v = 0$

(ii) $\|\lambda v\| = |\lambda|\|v\|$ for all $\lambda \in \mathbb{F}$ and $v \in V$

(iii) $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$

V equipped with a norm, $(V, \|\cdot\|)$, is called a *normed vector space*.

Remark. If V is a normed vector space and W is a vector subspace of V, then W is also a normed vector space.

Example 5.2. For all $x \in \mathbb{R}$, we have the norm $\|x\| = |x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x \leq 0 \end{cases}$.

And for $z \in \mathbb{C}$, we have the norm $\|z\| = |z| = |x + iy| = \sqrt{x^2 + y^2}$.

A norm on a vector space is essentially thought of as the size of an element in the space.

Now let X be a set. A *metric* on X is the function $d : X \times X \to \mathbb{R}$ such that for all $x, y, z \in X$, we have

(i) $d(x, y) \geq 0$ with equality if and only if $x = y$

(ii) $d(x, y) = d(y, x)$

(iii) $d(x, z) \leq d(x, y) + d(y, z)$

A set equipped with a metric is called a *metric space*.

Turns out that all normed spaces are metric spaces.

Proposition 5.3. *A norm induces a metric on a $\mathbb{F}$-vector space V, i.e., $d(v,w) = \|v-w\|$ for any $v, w \in V$.*

Proof. By definition we have $d(v,w) = \|v-w\| \geq 0$ with equality if and only if $v = w$ for any $v, w \in V$. We also have $d(v,w) = \|(-1)(w-v)\| = |-1|\|w-v\| = d(w,v)$ for any $v, w \in V$. Finally, for any $u, v, w \in V$, we have $d(u,v) + d(v,w) = \|u-v\| + \|v-w\| \geq \|u-w\| = d(u,w)$. □

Given a metric d on a $\mathbb{F}$-vector space X. Let $\|x\| = d(x,0)$ for $x \in X$. This defines a norm on the metric space X if and only if we have that the metric is *translation-invariant* (i.e., $d(x,y) = d(x+z, y+z)$ for $x, y, z \in X$) and *dilation-invariant* (i.e., $d(\lambda x, \lambda y) = |\lambda| d(x,y)$ for all $x, y \in X$ and $\lambda \in \mathbb{F}$). This is left as an exercise.

A *topology* τ on a set X is a collection of subsets on X such that

(i) $\emptyset, X \in \tau$

(ii) if $U_\gamma \in \tau$ for $\gamma \in \Gamma$, then $\bigcup_{\gamma \in \Gamma} U_\gamma \in \tau$

(iii) if $U_1, \ldots, U_n \in \tau$, then $\bigcap_{i=1}^n U_i \in \tau$.

The elements of τ are called *open sets* in the topology. A set equipped with a topology is a *topological space*.

Every metric space is a topological space. We can define the topology on the metric space X by first defining an open ball, then defining the open sets with Definition 5.5.

Since every normed space is a metric space, we will use normed spaces for our definition of an open ball by replacing the induced metric with the norm (i.e., $d(x,y) = \|x-y\|$).

Definition 5.4. Let $v \in V$ and $r > 0$. The *open ball* $B(v,r)$ is $\{w \in V : \|v-w\| < r\}$.

Definition 5.5. A subset $U \subseteq V$ is *open* if for all $u \in U$, there exists $r > 0$ such that $B(u,r) \subseteq U$.

Proposition 5.6. *Let V be a normed vector space. Then*

(i) V and $\emptyset$ are open in V

(ii) if $U_i \subseteq V$ is open for all i, then $\bigcup_i U_i$ is open

(iii) if $U_1, \ldots, U_n$ are open in V, then $\bigcap_{i=1}^n U_i$ is open.

Proof.

(i) It is clear V is open in V, by definition.
Suppose $\emptyset$ is not open. So, there exists $v \in \emptyset$ such that $B(v,r) \not\subseteq \emptyset$. This is a contradiction since no such v can exist. Thus $\emptyset$ is open in V.

(ii) Let U_i be open subsets in V for all i. And let $v \in \bigcup U_i$. Then we have $v \in U_i$ for some i. So there exists $r > 0$ such that $B(v,r) \subseteq U_i \subseteq \bigcup_i U_i$. So, $\bigcup_i U_i$ is open.

(iii) Let $U_i = U_1 \cap U_2 \cap \cdots \cap U_n$ be a collection of open sets in V. And let $v \in U_i$ for every i. So, for each i, there exists $\epsilon_i > 0$ such that $B(v, \epsilon_i) \subseteq U_i$. Now let ϵ be the minimum of ϵ_i. Then, for each i, we have $0 < \epsilon \leq \epsilon_i$, so then $B(v,\epsilon) \subseteq B(\epsilon_i) \subseteq U_i$. Thus U_i is open.

□

So, we can conclude that every normed space is also a topological space. And so any results in metric and topological spaces are also applied to normed spaces.

Proposition 5.7. *Let V be a normed vector space. Let $v, w \in V$, then $|\|v\| - \|w\|| \leq \|v - w\|$.*

Proof. We have $\|v\| = \|v - w + w\| = \|(v - w) + w\| \leq \|v - w\| + \|w\|$ by definition. So, $\|v\| - \|w\| \leq \|v - w\|$. We also have, $\|w\| = \|(w - v) + v\| \leq \|w - v\| + \|v\|$. And so $\|w\| - \|v\| \leq \|(-1)(v - w)\| = \|v - w\|$. Therefore, we have $|\|v\| - \|w\|| \leq \|v - w\|$. □

We can show that the *open* ball is, indeed, an *open* set.

Proposition 5.8. *The open ball in space V is an open subset.*

Proof. Let $u \in B(v, t)$. Then $\|u - v\| < t$. Now let $r = t - \|u - v\|$ so $r > 0$. Also let $w \in B(v, r)$. Then we have $\|u - w\| < r$. And so $\|v - w\| \leq \|v - u\| + \|u - w\| < \|v - u\| + r = t$. So then $w \in B(v, t)$. And so then $B(u, r) \subseteq B(v, t)$, which means $B(v, t)$ is an open subset. □

The following result is known as the *Minkowski inequality*. Given $(a_k) \in \mathbb{R}^n$, define $\|a_k\|_p = (\sum_{k=1}^n |a_k|^p)^{1/p}$.

Proposition 5.9. *If $p > 1$ and $a_k, b_k > 0$ then,*

$$\left(\sum_{k=1}^{n} |a_k + b_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{n} |a_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |b_k|^p\right)^{\frac{1}{p}},$$

which can be rewritten as $\|a_k + b_k\|_p \leq \|a_k\|_p + \|b_k\|_p$.

Proof. We need two results in order to prove this, namely, Young's inequality, and Holder's inequality.

We begin with the following result, called, *young's inequality*.

Lemma 5.10. *If $p, q > 1$ such that $1/p + 1/q = 1$ and $x, y \leq 0$, then*

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Proof (idea). Assume $x, y > 0$ otherwise if x or y is 0 then we have that the inequality is trivial. So we first let the real-valued function $f(a) = \dfrac{a^p}{p} + 1/q - a$, so that the stationary point ($f'(a) = 0$) is, then, $a = 1$. And this is a minimum ($f''(a) \geq 0$), so the function is increasing, i.e., $f(a) \geq f(1) = 0$. So, we have $a \leq \dfrac{a^p}{p} + 1/q$. And, now let $a = xy^{1/(1-p)}$. And so,

$$xy^{1/(1-p)} \leq \frac{x^p y^{p/(1-p)}}{p} + 1/q.$$

Now, multiply both sides by $y^{-p/(1-p)}$, and the result then follows. □

We also use the following result, called *Holder's inequality*.

Lemma 5.11. *If $p, q > 1$ such that $1/p + 1/q = 1$, and $a_k, b_k \in \mathbb{R}^n$,*

$$\sum_{i=1}^{n} |a_k||b_k| \leq \left(\sum_{i=1}^{n} |a_k|^p \right)^{1/p} \left(\sum_{i=1}^{n} |b_k|^q \right)^{1/q},$$

which can be rewritten as, $\|a_k b_k\|_1 \leq \|a_k\|_p \|b_k\|_p$.

Proof. Let $a_k, b_k > 0$, otherwise it is trivial. Now let $x_k = |a_k|(\sum_{k=1}^n |a_k|^p)^{-1/p}$ and $y_k = |b_k|(\sum_{k=1}^n |b_k|^q)^{-1/q}$. And so by Young's inequality, we get $x_k y_k \leq {x_k}^p/p + {y_k}^q/q$ for each k. So, now we have $\sum_{k=1}^n {x_k}^p = 1$ and $\sum_{k=1}^n {y_k}^q = 1$. So then, $\sum_{k=1}^n x_k y_k \leq 1/p + 1/q = 1$. And, $\sum_{k=1}^n x_k y_k = \|x_k y_k\|_1 \leq 1$. Finally, $\|a_k b_k\|_1 \leq \|a_k\|_p \|b_k\|_q$. □

Now, we can prove this. We have

$$\begin{aligned} |a_k + b_k|^p &= |a_k + b_k|^{p-1}|a_k + b_k| \leq |a_k + b_k|^{p-1}(|a_k| + |b_k|) \\ &= |a_k + b_k|^{p-1}|a_k| + |a_k + b_k|^{p-1}|b_k| \end{aligned}$$

Now, let $p - 1 = p/q$, where $q > 1$. And so, we have

$$\begin{aligned} \|a_k + b_k\|_p^p &= \sum_{k=1}^{n} |a_k - b_k|^p \\ &\leq \sum_{k=1}^{n} |a_k||a_k + b_k|^{p/q} + \sum_{k=1}^{n} |b_k||a_k + b_k|^{p/q} \end{aligned}$$

And so by Holder's inequality, we have $\|a_k + b_k\|_p^p \leq \||a_k + b_k|^{p/q}\|_q (\|a_k\|_p + \|b_k\|_p)$. And so, then

$$\begin{aligned} \||a_k + b_k|^{p/q}\|_q &= \left(\sum_{k=1}^{n} (|a_k + b_k|^{p/q})^q \right)^{1/q} \\ &= \left(\sum_{k=1}^{n} |a_k + b_k|^p \right)^{1/q} = \|a_k + b_k\|_p^{p/q} \end{aligned}$$

Finally, we have $\|a_k + b_k\| \leq \|a_k + b_k\|_p^{p/q}(\|a_k\|_p + \|b_k\|_p)$, and since $p - p/q = 1$, we have $\|a_k + b_k\|_p^{p-p/q} = \|a_k + b_k\|_p \leq \|a_k\|_p + \|b_k\|_p$. □

Example 5.12. Let $p \geq 1$, then $\mathbb{R}^n$ has a norm $\|(x_1, \ldots, x_n)\|_p = (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$.

We check the axioms,

(i) if $v \in \mathbb{R}^n$ and $\|v\|_p = 0$, then $(|v_1|^p + \cdots + |v_n|^p)^{\frac{1}{p}} = 0$ and so $|v_1|^p + \cdots + |v_n|^p = 0$, since $|v_i| \geq 0$ for all i, $v_1 = \cdots = v_n = 0$ so $v = 0$

(ii) if $v \in \mathbb{R}^n$ and $\lambda \in \mathbb{F}$, then $\|\lambda v\|_p = (|\lambda v_1|^p + \cdots + |\lambda v_n|^p)^{\frac{1}{p}} = (|\lambda|^p(|v_1|^p + \cdots + |v_n|^p)^{\frac{1}{p}} = |\lambda|\|v\|_p$

(iii) if $v, w \in \mathbb{R}^n$, then by using the Minkowski inequality,

$$\begin{aligned}\|v+w\|_p &= (|v_1+w_1|^p + \cdots + |v_n+w_n|^p)^{\frac{1}{p}} \\ &\leq (|v_1|^p + \cdots + |v_n|^p)^{\frac{1}{p}} + (|w_1|^p + \cdots + |w_n|^p)^{\frac{1}{p}} = \|v\|_p + \|w\|_p\end{aligned}$$

Example 5.13. Let $l^\infty = \{(a_n) : a_i \in \mathbb{F},\ (a_n) \text{ is bounded}\}$, then it has a norm $\|(a_n)\|_\infty = \sup\{|a_n| : n \in \mathbb{N}\}$.

We check the axioms:

(i) if $(a_n) \in l^\infty$ and $\|(a_n)\|_\infty = 0$, then $\sup\{|a_n| : n \in \mathbb{N}\} = 0$ and so $(a_n) = 0$ for all n, so $(a_n) = 0$

(ii) if $(a_n) \in l^\infty$ and $\lambda \in \mathbb{F}$, then
$\|\lambda(a_n)\|_\infty = \sup\{|\lambda(a_n)| : n \in \mathbb{N}\} = \sup\{|\lambda||(a_n)| : n \in \mathbb{N}\} = |\lambda|\|(a_n)\|_\infty$

(iii) if $(a_n), (b_n) \in l^\infty$, then

$$\begin{aligned}\|(a_n)+(b_n)\|_\infty &= \sup\{a_n + b_n : n \in \mathbb{N}\} \\ &\leq \sup\{|a_n| + |b_n| : n \in \mathbb{N}\} \\ &\leq \sup\{|a_n| : n \in \mathbb{N}\} + \sup\{|b_n| : n \in \mathbb{N}\} = \|(a_n)\|_\infty + \|(b_n)\|_\infty\end{aligned}$$

5.II Convergence

Now we will discuss convergence of sequences in normed spaces.

Definition 5.14. Let (v_n) be a sequence in the normed vector space V, and $v \in V$. We say (v_n) *converges* to v if for all $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $\|v_n - v\| < \epsilon$. We say $v_n \to v$ as $n \to \infty$ in norm.

Remark. We say (v_n) converges in norm to a limit $v \in V$ if and only if $\lim_{n\to\infty} \|v_n - v\| = 0$.

Definition 5.15. Let $S \subseteq V$. We say S is *closed* if, when (v_n) is a sequence in S which converges to a limit $v \in V$, we have $v \in S$.

Example 5.16. $[0,1]$ is closed, but $[0,1)$ is not. $(0,1)$ is open. $(-\infty, 0] \cup [1, \infty)$ is closed.

Remark. Note that subsets may be any of the following: open and not closed (such as $(0,1) \subseteq \mathbb{R}$), closed and not open ($[0,1] \subseteq \mathbb{R}$), neither open nor closed ($[0,1) \subseteq \mathbb{R}$), and both open and closed ($\mathbb{R} \subseteq \mathbb{R}$).

Definition 5.17. Let $S \subseteq V$. The *closure* of S is $\bar{S}$, and is the set of all elements $v \in V$ that are norm limits of S.

Corollary 5.18. *If $S \subseteq V$, we have S is closed if $\bar{S} = S$. If V is closed, then $\bar{S} \subseteq V$.* □

Proposition 5.19. *Let $A \subseteq V$. A is open if and only if the complement $V \setminus A$ is closed.*

Proof.

($\Rightarrow$) Let A be open in V. We want to show that for any sequence (v_n) in $V \setminus A$, such that $(v_n) \to v$, then $v \in V \setminus A$. So let $(v_n) \in V \setminus A$ converge to v. And suppose

$v \in A$. So, then, there exists $r > 0$ such that $B(v, r) \subseteq A$ by definition of openness. Now, since $v_n \to v$, then there exists $v_n \in B(v, r)$. But, then $v_n \in A$. This is a contradiction since $v \in V \setminus A$ by assumption. Hence, $v \notin A$, so $v \in V \setminus A$. Hence $V \setminus A$ is closed since it contains its limit points.

($\Leftarrow$) Now let $V \setminus A$ be closed. Suppose A is closed. So there exists $a \in A$ such that $B(a, r) \subseteq A$, which means $B(a, r) \cap (V \setminus A) \neq \emptyset$, for all $r > 0$. Hence, for each $n \in \mathbb{N}$, there exists $v_n \in B(a, 1/n)$ with $v_n \notin A$, by definition of closedness. And since $\|a - v_n\| < 1/n$, then $|-1|\|v_n - a\| < 1/n$, and so $v_n \to a$. But, then $a \in V \setminus A$ since $V \setminus A$ is closed. This is a contradiction since we assumed $a \in A$. Thus, A is open.

□

Definition 5.20. Let V be a normed vector space and $K \subseteq V$. K is *compact* if for any sequence (v_n) in K, we have a subsequence (v_{n_k}) such that (v_{n_k}) converges in norm to some $v \in K$.

Example 5.21. In $[0, 1]$, consider the sequence

$$\frac{1}{2}, 1, \frac{1}{3}, 0, \frac{1}{4}, 1, \frac{1}{5}, 0, \ldots.$$

This sequence does not converge. But, $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ does converge. $\frac{1}{2}, 0, \frac{1}{3}, 0, \frac{1}{4}, 0, \ldots$ also converges.

Proposition 5.22. *Let (v_n) be a sequence in V such that $v_n \to v \in V$. Then any subsequence (v_{n_k}) of (v_n) also converges to v.*

Proof. Let (v_{n_k}) be any subsequence of $(v_n) \in V$ and $v_n \to v \in V$. Then for any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for $n > N$, $\|v_n - v\| < \epsilon$. Let $K \in \mathbb{N}$ such that $n_K > N$. And if $k > K$, then $n_k > n_K > N$, and so $\|v_{n_k} - v\| < \epsilon$. □

Definition 5.23. $B \subseteq V$ is *bounded* if there exists c such that $\|v\| \leq c$ for all $v \in B$.

The following result is known as the *Bolzano-Weierstrass theorem*.

Theorem 5.24. *Every sequence in $\mathbb{R}$ that is bounded has a converging subsequence.*

Proof (idea). We give a sketch of the proof, which can be found in any Real Analysis textbook, such as [39], [40], and [41].

Definition. A *monotone* sequence is a sequence where all the terms are increasing or decreasing.

The next step is to show the following.
Lemma. Every real sequence has a monotone subsequence. □

The proof of this is left as an exercise for the reader.

Then we come to our main point.
Lemma. If a real sequence is monotone and bounded, then it converges. □

The idea of the proof is by first seeing what happens when the monotone sequence is increasing, and then decreasing. You can see that when it is increasing, the sequence is converging to the supremum of the sequence, while when it is decreasing, it is converging to the infimum of the sequence.

It then follows from the above statements, that indeed if a real sequence is bounded, then it has a monotone subsequence, which is also bounded, and thus is convergent. □

The following is a corollary to Bolzano-Weirstrass, which can be shown using induction (left as an exercise).

Corollary 5.25. *Every sequence in $\mathbb{R}^n$ that is bounded for some n, has a converging subsequence.* □

Lemma 5.26. *Norm limits are unique, i.e., for a sequence (v_n) in a normed vector space V such that $v_n \to v$ and $v_n \to v'$, then $v = v'$.*

Proof. Since v_n converges to both v and v', we have that for any $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that $\|v_n - v\| \leq \epsilon/2$ if $n > N$; and there exists some N' such that $\|v_n - v'\| < \epsilon/2$ if $n > N'$. So we have that if $n > \max(N, N')$ we get $0 \leq \|v - v'\| \leq \|v - v_n\| + \|v_n - v'\| \leq \epsilon$. So then $0 \leq \|v - v'\| \leq \epsilon$ for all $\epsilon > 0$, and so $\|v - v'\| = 0$ which implies $v = v'$. □

The following result is known as *Heine-Borel theorem*.

Theorem 5.27. *Let V be a finite-dimensional normed vector space and $K \subseteq V$. K is compact if and only if it is closed and bounded.*

Proof.

(⇒) Let K be compact. We first want to show that if $v_n \in K$ and $v_n \to v$, then $v \in K$, i.e., K is closed. So suppose $v_n \to v \in K$ for a sequence $(v_n) \in K$. So, then, by definition of compactness, there exists $v_{n_k} \to v' \in K$. Then we also have $v_{n_k} \to v$ since (v_{n_k}) is a subsequence of (v_n). And so, by the above lemma, we have $v = v'$. Thus, K is closed.
Now, suppose K is not bounded. Then there exists $(v_n) \in K$ such that for all c we have $\|v_n\| > c$. But, then no subseqeunce of (v_n) is bounded and so none would converge to some element in K. This is a contradiction since K is compact by assumption. Hence K is bounded.

(⇐) Now, let K be bounded and closed. Let $(v_n) \in K$. Since K is bounded, then there exists c such that $\|w_n\| \leq c$. And by Bolzano-Weirstrass, (v_n) has a subsequence (v_{n_k}) that converges to some v. And since K is closed, any sequence in K will contain its limit points, and thus $v \in K$. Hence, K is compact.

□

5.III Continuity

Now we will talk about the important notion of continuity of functions between normed spaces.

Definition 5.28. Let V, W be normed vector space, $X \subseteq V$ and $Y \subseteq W$. The function $f : X \to Y$ is *continuous* if for all $x_0 \in X$, $\epsilon > 0$ there exists $\delta > 0$ such that $\|x - x_0\| < \delta$ implies $\|f(x) - f(x_0)\| < \epsilon$.

Proposition 5.29. *Given a function between normed spaces $f : V \to W$, and $v \in V$. Then the definition of continuity above is equivalent to saying that, for sequences $(v_n) \in V$ such that $v_n \to v$, then we have $f(v_n) \to f(v) \in W$.*

Proof.

($\Rightarrow$) Suppose f is continuous at $v \in V$. Now let $(v_n) \in V$ such that $v_n \to v$. We want to show that for any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n > N$, we have $\|f(v_n) - f(v)\| < \epsilon$. Now, let $\epsilon > 0$ there exists $\delta > 0$ such that for all $w \in V$ such that $\|w - v\| < \delta$ implies $\|f(w) - f(v)\| < \epsilon$. And so there exists $N \in \mathbb{N}$ such that for all $n > N$, we have $\|v_n - v\| < \delta$. Which means $\|f(v_n) - f(v)\| < \epsilon$. Hence, $f(v_n) \to f(v)$.

($\Leftarrow$) Now suppose f is not continuous at v. Then there exists $\epsilon > 0$ such that for all $\delta > 0$ there exists $w \in V$ such that $\|w - v\| < \delta$ and $\|f(w) - f(v)\| \geq \epsilon$. Let $\delta = 1/n$, then there exists some $(v_n) \in V$ such that $\|v_n - v\| < 1/n$ and $\|f(v_n) - f(v)\| \geq \epsilon$. This is a contradiction to our assumption since $v_n \to v$ but $f(v_n) \not\to f(v)$.

□

The following result shows us that to talk about continuity, we can forget about metrics and norms and use open and closed sets.

Lemma 5.30. *Given a function between normed spaces $f : X \to Y$. We have the following equivalence.*

(i) f is continuous

(ii) for every closed $U \subseteq Y$, the preimage $f^{-1}(U) = \{x \in X : f(x) \in U\} \subseteq X$ is closed

(iii) for every open $U \subseteq Y$, the preimage $f^{-1}(U) \subseteq X$ is open

Proof.

$(i) \Rightarrow (ii)$ Let $f : X \to Y$ be continuous and let $U \subseteq Y$ be closed. We want to show that for any $(x_n) \in f^{-1}(U)$ such that $x_n \to x$, then $x \in f^{-1}(U)$. So let (x_n) be a sequence in $f^{-1}(U)$ such that $x_n \to x$, so then we have $f(x_n) \to f(x)$. Since $x_n \in f^{-1}(U)$, we have $f(x_n) \in U$. And so $f(x) \in U$ and $x \in f^{-1}(U)$. Hence $f^{-1}(U)$ is closed.

$(ii) \Rightarrow (iii)$ Now let $U \subseteq Y$ be open. Then the complement $Y \setminus U$ is closed. So $f^{-1}(Y \setminus U)$ is closed by (ii). Now let $x \in f^{-1}(Y \setminus U$, so $f(x) \in Y \setminus U$, which means $f(x) \notin U$, and so we have $x \in X \setminus f^{-1}(U)$; thus $f^{-1}(Y \setminus U) = X \setminus f^{-1}(U)$ and so $X \setminus f^{-1}(U)$ is closed, hence $f^{-1}(U)$ is open in X.

$(iii) \Rightarrow (i)$ Finally, let $x \in X$ and $r > 0$. Now we have $f(x) \in B(f(x), r) \subseteq Y$, so $x \in f^{-1}(B(f(x), r))$ is open. Therefore for some $\delta > 0$, we have $B(x, \delta) \subseteq f^{-1}(B(f(x), r))$. Thus f is continuous since $f(B(x, \delta)) \subseteq B(f(x), r)$.

□

Proposition 5.31. *Given two continuous functions $f : U \to V$ and $g : V \to W$ between normed spaces. Then we have that the function $g \circ f : U \to W$ is also continuous.*

Proof. Let $\mathcal{U} \subseteq W$ be an open subset. We want to show $(g \circ f)^{-1}(\mathcal{U})$ is open. We first want to show $(g \circ f)^{-1}(\mathcal{U}) = f^{-1}(g^{-1}(\mathcal{U}))$.
If $x \in (g \circ f)^{-1}(\mathcal{U})$, then we have $g(f(x)) = (g \circ f)(x) \in \mathcal{U}$. And so $f(x) \in g^{-1}(\mathcal{U})$, and then we have $x \in f^{-1}(g^{-1}(\mathcal{U}))$. And for the converse, if $x \in f^{-1}(g^{-1}(\mathcal{U}))$, then we have $f(x) \in g^{-1}(\mathcal{U})$, and so $g(f(x)) = (g \circ f)(x) \in \mathcal{U}$. Finally, we have $x \in (g \circ f)^{-1}(\mathcal{U})$. Thus, $(g \circ f)^{-1}(\mathcal{U}) = f^{-1}(g^{-1}(\mathcal{U}))$.
By the above, since g is continuous, we have $g^{-1}(\mathcal{U}) \subseteq V$ is open. Similarly, since f is continuous, we have $f^{-1}(g^{-1}(\mathcal{U})) \subseteq U$ is also open. Thus, $g \circ f : U \to W$ is continuous. □

Lemma 5.32. *Let $f : X \to Y$ be continuous and $K \subseteq X$ be compact. Then $f(K) = \{f(x) : x \in K\}$ is also compact.*

Proof. Let (y_n) be any sequence in $f(K)$. Then for each n, there exists $x_n \in K$ such that $f(x_n) = y_n$. Since K is compact, there exists a subsequence (x_{n_k}) such that it converges to $x \in K$. Since f is continuous, we have $y_{n_k} = f(x_{n_k})$ converges to $f(x) \in f(K)$. So $f(K)$ is compact since (y_n) has a subsequence converging to some point in $f(K)$. □

Theorem 5.33. *Let $K \subseteq V$ be compact and $f : K \to \mathbb{F}$ be continuous. Then f is bounded.*

Proof. By the above lemma, $f(K)$ is compact. And by Heine-Borel theorem, we know that any compact set is bounded. □

5.IV Cauchy, complete spaces and Banach spaces

In this section, we will learn about completeness (Banach spaces are complete normed spaces), and completions. We will conclude the chapter with the fact that any normed vector space has a completion. We will prove this properly in the next chapter with dual spaces. However, we state that we can find completions, in a constructive method, similarly to the way we constructed the real numbers in §4.

Recall the definition of a Cauchy sequence on $\mathbb{Q}$ from before (Definition 4.13). But, we rewrite it so that we can use the norm limit instead.

Definition 5.34. Let (v_n) be a sequence in a normed vector space V. Then (v_n) is *Cauchy* if for all $\mathbb{R} \ni \epsilon > 0$ there exists $N \in \mathbb{N}$ such that $m, n \geq N$ implies $\|v_m - v_n\| < \epsilon$.

Proposition 5.35. *If (v_n) is a norm convergent sequence in a normed vector space V, then (v_n) is a Cauchy sequence. (i.e., every convergent sequence is Cauchy.)*

Proof. Let v be the limit of (v_n), and $\epsilon > 0$. There exists $N \in \mathbb{N}$ such that if $n \geq N$, then $\|v_n - v\| < \epsilon/2$. So if $m, n \geq N$, then $\|v_m - v_n\| = \|v_m - v + v - v_n\| \leq \|v_m - v\| + \|v - v_n\| = \|v_m - v\| + \|v_n - v\| < \epsilon$. □

Proposition 5.36. *A Cauchy sequence is bounded.*

Proof. By definition, there is some N such that for all $n \geq N$, we have $\|x_N - x_n\| < 1$. So, $\|x_n\| < 1 + \|x_N\|$ for $n \geq N$. So, for all n, we have $\|x_n\| \leq \max\{\|x_1\|, \ldots, \|x_{N-1}\|, 1 + \|x_N\|\}$. □

Proposition 5.37. *Let (v_n) be a Cauchy sequence in V. If (v_n) has a subsequence $v_{n_k} \to v$, then (v_n) also converges to v.*

Proof. Let $\epsilon > 0$. Since (v_n) is Cauchy, there exists N such that $\|v_n - v_m\| < \epsilon/2$ for all $n, m > N$. Since $v_{n_k} \to v$, there exists K such that $\|v_{n_k}\| < \epsilon/2$ for all $k > K$. Choose $M \in \mathbb{Z}$ such that $M > N$ and $M > K$ and let $m = n_M$. Then $m \geq M > N$. Now let $n > m$, then $n, m > N$ so $\|v_n - v_m\| < \epsilon/2$. But we have $v_m = v_{n_M}$ is in the subsequence and $M > K$ so $\|v_m - v\| < \epsilon/2$. Hence $\|v_n - v\| \leq \|v_n - v_m\| + \|v_m - v\| < \epsilon/2 + \epsilon/2 = \epsilon$. □

Definition 5.38. A normed vector space V is *complete* if every Cauchy sequence converges. A complete normed vector space is a *Banach space*.

Example 5.39. Any finite dimensional normed vector space is a Banach space.
l^p is a Banach space.
Vector space $C[-1, 1]$ defined by $\|f\| = \int_{-1}^{1} |f(x)|\, dx$, where $f \in C[-1, 1]$, is not a Banach space. But, it is a Banach space with $\|f\| = \sup\{|f(x)| : x \in [-1, 1]\}$.

Example 5.40. $\mathbb{Q}$ is not complete since the sequence

$$\frac{1}{1}, \frac{14}{10}, \frac{141}{100}, \dots$$

in $\mathbb{Q}$ has its limit as $\sqrt{2} \notin \mathbb{Q}$, but it is Cauchy since $\sqrt{2} \in \mathbb{R}$.

Proposition 5.41. $\mathbb{R}$ *is a Banach space.*

Proof. We want to show that every real Cauchy sequence is convergent. Let $(v_n) \in \mathbb{R}$ be a Cauchy sequence. From above results, we know (v_n) is bounded. And from the previous section, we saw the Bolzano-Weirstrass theorem, which says any bounded real sequence has a converging subsequence. So let (v_{nk}) be a subsequence of (v_n), such that $v_{nk} \to v \in \mathbb{R}$. By the above proposition, we see $v_n \to v$. □

Corollary 5.42. $\mathbb{C}$ *is a Banach space.* □

Definition 5.43. The subset $W \subseteq V$ is *dense* if $\bar{W} = V$.

The following result shows that our previous definition of dense from Lemma 4.15 is equivalent to the one above.

Proposition 5.44. *Given a normed vector space V, and $U \subseteq V$, then the following are equivalent.*

(i) for all $v \in V$ and $\epsilon > 0$, there exists $u \in U$, such that $\|u - v\| < \epsilon$

(ii) $\bar{U} = V$

Proof.

(*i*) ⇒ (*ii*) Let $v \in V \setminus U$, so that for some $r > 0$, there exists an open ball $B(v, r)$ that contains some $v_n \in U$. So, then $\|v_n - v\| < \epsilon$, i.e., $v_n \to v$. And so $v \in \bar{U}$. So we now have $V \setminus U \subseteq \bar{U}$. So, $V \subseteq \bar{U}$. And we know $\bar{U} \subseteq V$, therefore $V = \bar{U}$.

(*ii*) ⇒ (*i*) Now let $v \in V \setminus U$. So, then $v \in \bar{U} \setminus U$, and there exists $(v_n) \in U$ such that $v_n \to v$. And so for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|v_n - v\| < \epsilon$.

□

Definition 5.45. A Banach space $\hat{V}$ is a *completion* of V, if V is a dense subset of $\hat{V}$.

Following Lemma 4.15, we saw $\mathbb{Q}$ is dense in $\mathbb{R}$, where $\mathbb{Q}$ is embedded into $\mathbb{R}$ with the inclusion map $\iota : \mathbb{Q} \hookrightarrow \mathbb{R}$ given by $x \mapsto [(x, x, x, \ldots)]$. The method of completion we did in §4 gives us an idea of how to prove the following theorem. One way is applying a similar method to constructing the real numbers from the rationals. We can generalise such by saying that the completion of any normed space is the set of equivalence classes of Cauchy sequences in the normed space.

Theorem 5.46. *Any normed vector space V has a completion $\hat{V}$.*

We will prove this theorem (via dual spaces) at the end of the next chapter.

6 Spaces of Bounded Linear Maps

In this chapter we will learn about bounded linear maps; the Hahn-Banach theorems (proof omitted); and dual spaces. We will conclude the chapter with the fact that any normed vector space has a completion. We prove this with dual spaces.

This is based on the presentations of [1], [4], [27], [28], [38], and [40].

6.I Bounded linear maps

In this section we will learn about bounded, equivalently continuous (see below), linear maps between normed vector spaces.

Definition 6.1. Let U, V be $\mathbb{F}$-vector spaces. Then $T : U \to V$ is a *linear map* if

(i) $T(u + v) = T(u) + T(v)$ for all $u, v \in U$

(ii) $T(\lambda u) = \lambda T(u)$ for all $\lambda \in \mathbb{F}$ and $u \in U$.

Remark. We can instead say that T is a linear map if $T(\lambda u + \mu v) = \lambda T(u) + \mu T(v)$.

Definition 6.2. Let V, W be normed vector spaces. The linear map $T : V \to W$ is a *bounded* if there exists $C > 0$ such that $\|Tv\| \leq C\|v\|$ for all $v \in V$.

Theorem 6.3. *Let V, W be normed vector spaces, and $T : V \to W$ a linear map. Then the following are equivalent.*

(i) T is continuous

(ii) T is continuous at 0

(iii) T is bounded

Proof.

$(i) \Rightarrow (ii)$ By definition.

$(ii) \Rightarrow (iii)$ Let $\epsilon = 1$, there exists $\delta > 0$ such that $\|v - 0\| = \|v\| < \delta$ implies $\|Tv - 0\| = \|Tv\| < 1$. If $v = 0$, then $\|Tv\| = 0 \leq C\|v\|$. Now let $v \in V \setminus \{0\}$, then

$$\left\| T\left(\frac{\delta v}{2\|v\|}\right) \right\| = \left\| \frac{\delta}{2\|v\|} \right\| \|Tv\| = \frac{\delta}{2\|v\|}\|Tv\| < 1$$

and so $\|Tv\| < \dfrac{2}{\delta}\|v\|$. So if we let $C = 2/\delta$, then T is bounded.

(iii) $\Rightarrow$ *(i)* Since T is bounded, there exists $C > 0$ such that $\|Tv\| \leq C\|v\|$ for all $v \in V$. Let $\epsilon > 0$, so for $v, w \in V$, there exists $\|v - w\| < \epsilon/C$. Then $\|Tv - Tw\| = \|T(v - w)\| \leq C\|v - w\| < \epsilon$. And so T is continuous.

□

Definition 6.4. Let V, W be normed vector spaces. $\mathrm{Hom}(V, W)$ is the set of all bounded linear maps from $T : V \to W$.

Definition 6.5. Given a bounded linear map $T : V \to W$. There exists a *norm*

$$\|T\| = \sup_{v \in V, \|v\| \leq 1} \|Tv\|.$$

We also have the inequality $\|Tv\| \leq \|T\|\|v\|$ for all $v \in V$, such that $\|T\|$ is the minimum C such that $\|Tv\| \leq C\|v\|$.

Remark. $\mathrm{Hom}(V, W)$ is a normed vector space with the above norm and operations defined by $(S + T)(v) = S(v) + T(v)$ and $(\lambda T)(v) = \lambda T(v)$ for $S, T \in \mathrm{Hom}(V, W)$, $v \in V$ and $\lambda \in \mathbb{F}$.

6.II Dual spaces

Definition 6.6. Let V be a normed vector space. The *dual space*, V^*, is $\mathrm{Hom}(V, \mathbb{F})$.

Theorem 6.7. *Let V be a normed vector space and W a Banach space. Then $\mathrm{Hom}(V, W)$ is a Banach space.*

Proof. Let $(T_n) \in \mathrm{Hom}(V, W)$ be a Cauchy sequence and $v \in V$. We want to show that there exists some $T \in \mathrm{Hom}(V, W)$ such that $T_n \to T$.

Now we have $(T_n(v))$ is Cauchy in W since for all $n, m \in \mathbb{N}$, $\|T_n(v) - T_m(v)\| \leq \|T_n - T_m\|\|v\|$. And so since W is Banach, there exists some $w \in W$ such that $T_n(v) \to w$ for $n \to \infty$. Now let $T : V \to W$, where $T(v) := \lim_{n\to\infty} T_n(v)$. Now we have $T(\lambda u + \mu v) = \lim_{n\to\infty} T_n(\lambda u + \mu v) = \lambda \lim_{n\to\infty} T_n(u) + \mu \lim_{n\to\infty} T_n(v) = \lambda(T(u)) + \mu(T(v))$, for all $\lambda, \mu \in \mathbb{F}$ and $u, v \in V$, i.e., T is a linear map.

Since $T_n(v) \to T(v)$, we have that for $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n > N$, we have $\|T_n(v) - T(v)\| < \epsilon$, and so $\|T(v)\| = \|T_n(v) - T(v) - T_n(v)\| \leq \|T_n(v) - T(v)\| + \|T_n(v)\| < \epsilon + \|T_n(v)\|$. And since $(T_n) \in \mathrm{Hom}(V, W)$, we have that T is bounded. So T is in $\mathrm{Hom}(V, W)$.

Now, since (T_n) is Cauchy, let $\epsilon > 0$, then there exists $N \in \mathbb{N}$ such that $\|T_m - T_n\| < \epsilon/2$ for $m, n \geq N$. And since $T_n(v) \to T(v)$ in W, we have that for $v \in V$ and $\|v\| \leq 1$, then $\|T_m(v) - T_n(v)\| < \epsilon/2$. Now take the limit as $m \to \infty$, for $n \geq N$, $\|T(v) - T_n(v)\| \leq \epsilon/2$. Thus,

$$\|T - T_n\| = \sup_{v \in V, \|v\| \leq 1} \|T(v) - T_n(v)\| \leq \epsilon/2 < \epsilon.$$

So, then T is the norm limit of (T_n) in $\mathrm{Hom}(V, W)$, which means $\mathrm{Hom}(V, W)$ is complete.

□

Corollary 6.8. *Let V be a normed vector space, then V^* is a Banach space.*

Proof. By the above theorem, $\mathrm{Hom}(V, \mathbb{F})$ is a Banach space since $\mathbb{F}$ is complete. So V^* is a Banach space whether V is complete or not. □

Proposition 6.9. *Let V be a Banach space, and W a closed subspace of V. Then W is a Banach space.*

Proof. Let (v_n) be a Cauchy sequence in W. Then it is also Cauchy in V, and therefore converges to some $v \in V$. And since W is closed in X, we have $v \in W$. Hence W is a Banach space. □

6.III The Hahn Banach theorem

The Hahn-Banach theorem is a fundamental result in functional analysis. The proof is omitted here. The interested reader is directed to one of the following for a proper proof: [4], [7], [16], [22], [24], [27], [33], [37], or [38].

The next result is known as the *Hahn-Banach theorem*.

Theorem 6.10. *Let V be a normed vector space, and $W \subseteq V$ be a subspace, and let $f : W \to \mathbb{F}$ be a bounded linear map. Then there exists a bounded linear map $F : V \to \mathbb{F}$ such that $F|_w = f$, where $F|_w = F : W \to \mathbb{F}$, and $\|F\| = \|f\|$.* □

What this essentially says is that for a normed vector space V and a subspace $W \subseteq V$, every norm on a bounded linear map on W can be extended to V.

6.IV Double Duals

Definition 6.11. Given a normed vector space V, we have the *double dual* $V** := (V^*)^*$. We define the map $\phi : V \to V^{**}$ by $\phi(v)(f) = f(v)$, for $v \in V$ and $f \in V^*$.

$\phi(v)$ is linear, since for all $v \in V$, $f, g \in V^*$, and $\lambda, \mu \in \mathbb{F}$, we have $\phi(v)(\lambda f + \mu g) = (\lambda f + \mu g)(v) = \lambda f(v) + \mu g(v) = \lambda\phi(v)(f) + \mu\phi(v)(g)$. And it is bounded since $\|\phi(v)(f)\| = \|f(v)\| \leq \|v\|\|f\|$ so $\|\phi(v)\| \leq \|v\|$ for all $v \in V$. So, indeed $\phi(v)$ is a bounded linear map, $\phi(v) \in (V^*)^*$.

Proposition 6.12. *$\|\phi(v)\| = \|v\|$ for all $v \in V$ (i.e., ϕ is an isometry).*

Proof. We need some $F \in V^*$ such that $\|\phi(v)(F)\| = \|v\|\|F\|$, i.e., $\|F(v)\| = \|v\|\|F\|$. If $v = 0$, then it is clear. So, let $v \neq 0$ and define $f : \mathrm{Span}(v) \to \mathbb{F}$ by $f(\lambda v) = \lambda\|v\|$. So

$$\|f\| = \sup_{0 \neq \lambda \in \mathbb{F}} \frac{\|f(\lambda v)\|}{\|\lambda v\|} = \frac{|\lambda|\|v\|}{|\lambda|\|v\|} = 1.$$

Now, by the Hahn-Banach theorem, we have $F : V \to \mathbb{F}$ such that $\|F\| = 1$ and $F(\lambda\|v\|) = \lambda\|v\|$ for all $\lambda \in \mathbb{F}$. Hence, $\|v\|\|F\| = \|v\| = \|F(v)\| = \|\phi(v)(F)\|$. So $\|\phi(v)\| = \|v\|$ for all $v \in V$. □

Now we prove Theorem 5.46 (i.e., any normed vector space has a completion).

Proof. We know from above, that any dual space is a Banach space. So $\phi(V) \subseteq (V^*)^*$ is Banach. So then, $\overline{\phi(V)}$ is also Banach. Thus, $\phi(V)$ is dense in $\overline{\phi(V)}$. So $\hat{V} = \overline{\phi(V)}$. □

Example 6.13. In Example 5.39, we said $C[-1,1]$ defined with that norm is not Banach thus not complete. But, it has a completion, $L^p[0,1]$.

7 Hilbert Spaces

In this chapter, we will learn about inner product spaces, orthogonality in vector spaces, and adjoints.

This is based on the presentations given in [27] and [4].

7.I Inner product spaces

In this section, we will see how an inner product induces a norm, so that any inner product space is a normed vector space.

Definition 7.1. Let V be a $\mathbb{F}$-vector space. An *inner product* on V is a map $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}$ such that

(i) $\langle u, \lambda v + \mu w\rangle = \lambda\langle u, v\rangle + \mu\langle u, w\rangle$ for all $u, v, w \in V$ and $\lambda, \mu \in \mathbb{F}$

(ii) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$

(iii) $\langle u, u\rangle \geq 0$ and equality holds if and only if $u = 0$.

Here, $\bar{\alpha}$ means the complex conjugate, for $\alpha \in \mathbb{F}$. So, $\overline{x + iy} = x - iy$ for $x, y \in \mathbb{R}$. An *inner product space* is $(V, \langle\cdot,\cdot\rangle)$, i.e., a vector space equipped with an inner product.

The following result is known as the *Cauchy-Shwarz inequality*.

Lemma 7.2. *Let V be an inner product space, and $v, w \in V$. Then*

$$|\langle u, v\rangle| \leq \sqrt{\langle u, u\rangle\langle v, v\rangle}.$$

Proof. Assume $v \neq 0$, otherwise it is trivial ($\langle u, v\rangle = 0$ and $\langle v, v\rangle = 0$). So $\langle v, v\rangle \neq 0$. And so, for $\lambda \in \mathbb{F}$, we have

$$\begin{aligned} 0 \leq \langle u - \lambda v, u - \lambda v\rangle &= \langle u - \lambda v, u\rangle + \langle u - \lambda v, -\lambda v\rangle \\ &= \langle u, u\rangle - \bar{\lambda}\overline{\langle u, v\rangle} - \lambda\langle u, v\rangle + \bar{\lambda}\lambda\langle v, v\rangle \end{aligned}$$

And so, let $\lambda = \frac{\overline{\langle u,v\rangle}}{\langle v,v\rangle}$, then

$$\langle u, u\rangle - \frac{\langle u, v\rangle\overline{\langle u, v\rangle}}{\langle v, v\rangle} - \frac{\overline{\langle u, v\rangle}\langle u, v\rangle}{\langle v, v\rangle} + \frac{\langle u, v\rangle\overline{\langle u, v\rangle}}{\langle v, v\rangle} \geq 0.$$

And so,

$$\langle u, u\rangle\langle v, v\rangle - 2|\langle u, v\rangle|^2 + |\langle u, v\rangle|^2 \geq 0$$

and the result follows from there. □

This next result shows that all inner product spaces are normed spaces with an induced norm.

Proposition 7.3. *If V is an inner product space and $v \in V$, then $\|v\| = \sqrt{\langle v, v\rangle}$ is a norm.*

Proof. If $\|v\| = 0$, then $\langle v, v\rangle = 0$, so $v = 0$. And if $v = 0$, then $\sqrt{\langle v, v\rangle} = 0 = \|v\|$. Now let $\lambda \in \mathbb{F}$ and $v \in V$. Then $\|\lambda v\|^2 = \langle \lambda v, \lambda v\rangle = \lambda\bar{\lambda}\langle v, v\rangle = |\lambda|^2\|v\|^2$, so $\|\lambda v\| = |\lambda|\|v\|$. Finally, let $w \in V$, then by using Cauchy-Shwarz,

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w\rangle \\
&= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\
&= \|v\|^2 + \langle v, w\rangle + \overline{\langle v, w\rangle} + \|w\|^2 \\
&= \|v\|^2 + 2\Re\langle v, w\rangle + \|w\|^2 \\
&\leq \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \\
&\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\
&= (\|v\| + \|w\|)^2
\end{aligned}$$

And so we then get $\|v + w\| \leq \|v\| + \|w\|$. □

Thus now we have show that all inner product spaces are normed spaces; all normed spaces are metric spaces; and all metric spaces are topological spaces. We have the following hierarchy of spaces: *Vector Spaces* → *Topological Spaces* → *Metric Spaces* → *Normed Spaces* → *Inner Product Spaces.*

Proposition 7.4. *Let V be a $\mathbb{F}$-vector space. A norm on the space V comes from an inner product if and only if it satisfies the parallelogram law (i.e., $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$ for all $u, v \in V$).*

Proof.

(⇒) Let V be a vector space with an inner product, and with the induced norm (i.e., $\|v\| = \sqrt{\langle v, v\rangle}$ for $v \in V$). Let $u, v \in V$, then,

$$\begin{aligned}
\|u + v\|^2 + \|u - v\|^2 &= \langle u + v, u + v\rangle + \langle u - v, u - v\rangle \\
&= \langle u + v, u\rangle + \langle u + v, v\rangle + \langle u - v, u\rangle - \langle u - v, v\rangle \\
&= \langle u, u\rangle + \overline{\langle u, v\rangle} + \langle v, v\rangle + \overline{\langle v, u\rangle} + \langle u, u\rangle \\
&\quad - \overline{\langle u, v\rangle} - \overline{\langle v, u\rangle} + \langle v, v\rangle \\
&= 2\langle u, u\rangle + 2\langle v, v\rangle = 2(\|u\|^2 + \|v\|^2)
\end{aligned}$$

(⇐) Now let the vector space V with a norm satisfy the parallelogram. We want to show that this norm is the induced norm by the inner product on V. That is, we have $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$ for all $u, v \in V$, and we want to show that there exists an inner product $\langle u, v\rangle$ for all $u, v \in V$, such that $\langle u, u\rangle = \|u\|^2$ for all $u \in V$, in particular, that this inner product is uniquely determined by the induced norm.

We claim $\|u+v\|^2 = \langle u + v, u + v\rangle$ and $\|u-v\|^2 = \langle u - v, u - v\rangle$ so $\|u+v\|^2 - \|u - v\|^2 = \langle u + v, u + v\rangle - \langle u - v, u - v\rangle = 4\Re\langle u, v\rangle$. Thus, the real part of the inner product, $\Re\langle u, v\rangle$, is $1/4(\|u+v\|^2 - \|u-v\|^2)$. And we know that the imaginary part of an equation can be found from the real part of the equation, i.e., $\Re(-i(u+iv)) = \Im(u + iv)$. So, we let $\Im\langle u, v\rangle = \Re(-i\langle u, v\rangle)$, and so we find

$$\langle u, v\rangle = \Re\langle u, v\rangle + \Im\langle u, v\rangle = \frac{\|u + v\|^2 - \|u - v\|^2 + i\|u + iv\|^2 - i\|u - iv\|^2}{4}$$

This is because $\Im(\langle u, v\rangle) = \Re(-i\langle u, v\rangle) = \Re(\langle u, iv\rangle)$.
Also, note that this is only for when $\mathbb{F} = \mathbb{C}$. When $\mathbb{F} = \mathbb{R}$, then we have

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}.$$

And we are done.

□

The following result shows us that the inner product is continuous.

Proposition 7.5. *Given an inner product space V. If (u_n) and (v_n) are sequences in V such that $u_n \to u$ and $v_n \to v$, then $\langle u_n, v_n\rangle \to \langle u, v\rangle$.*

Proof.

$$\begin{aligned}
|\langle u_n, v_n\rangle - \langle u, v\rangle| &= |\langle u_n, v_n\rangle - \langle u_n, v\rangle + \langle u_n, v\rangle - \langle u, v\rangle| \\
&= |\langle u_n, v_n - v\rangle + \langle u_n - u, v\rangle| \\
&\leq |\langle u_n, v_n - v\rangle| + |\langle u_n - u, v\rangle| \\
&\leq \|u_n\|\|v_n - v\| + \|u_n - u\|\|v\|
\end{aligned}$$

And since $u_n \to u$, we have $\|u_n - u\| \to 0$; similarly, as $v_n \to v$, we have $\|v_n - v\| \to 0$; and also $\|u_n\| \to \|u\|$. So, then, we have $|\langle u_n, v_n\rangle - \langle u, v\rangle| \to 0$, and so $\langle u_n, v_n\rangle \to \langle u, v\rangle$. □

Definition 7.6. An inner product space that is complete is a *Hilbert space*.

Given an inner product space V. We say that the Hilbert space $\hat{V}$ is a *completion* of V if V is dense in $\hat{V}$ with the same inner product. Constructing this is similar to constructing completions in Normed spaces in the previous chapter.

Example 7.7.

- $\mathbb{F}^n$ has an inner product $\langle v, w\rangle = \overline{v}^T w$.
- Consider the space l^2, the set of sequences in $\mathbb{F}$ such that $\sum_{n=1}^{\infty} |a_n|^2 < \infty$. If $(a_n), (b_n) \in l^2$ then by Cauchy-Schwarz for $\mathbb{F}^N$, we get

$$\sum_{n=1}^{N} |\overline{a_n} b_n| \leq \left(\sum_{n=1}^{N} |a_n|^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^{N} |b_n|^2\right)^{\frac{1}{2}}$$

  hence $\sum_{n=1}^{\infty} \overline{a_n} b_n$ converges absolutely and hence converges. Now, define the inner product on l^2 to be $\langle (a_n), (b_n)\rangle = \sum_{n=1}^{\infty} \overline{a_n} b_n$. This is a Hilbert space.

7.II Orthogonal complements

Definition 7.8. Let H be a Hilbert space and $S \subseteq H$. The *orthogonal complement* $S^{\perp} = \{x \in H : \langle x, y\rangle = 0 \text{ for all } y \in S\}$.

Properties of orthogonal complements:

Proposition 7.9. *Let H be a Hilbert space and $S \subseteq H$. Then $S^{\perp}$ is a closed vector subspace of H.*

Proof. We check the vector subspace criterion. We first have $0 \in S^\perp$, since $\langle 0, u\rangle = 0$ for all $u \in S$. Now let $u, v \in S^\perp$, so $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle = 0$ for all $w \in S$. Hence $\alpha u + \beta v \in S^\perp$, and $S^\perp$ is a vector subspace of H.
Now let $(s_n) \in S^\perp$ such that it converges to some $s \in \overline{S^\perp}$. Then for all $t \in S$, we have $\langle s, t\rangle = \lim_{n\to\infty}\langle s_n, t\rangle = 0$. So, $s \in S^\perp$. And so is closed, since $S^\perp = \overline{S^\perp}$. □

Proposition 7.10. *Let H be a Hilbert space and let $A, B \subseteq H$. Then $(A \cup B)^\perp = A^\perp \cap B^\perp$.*

Proof. Suppose $v \in (\bigcup S_i)^\perp$. Then we have $\langle v, x\rangle = 0$ for all $x \in S_i$ and all i. Hence $v \in S_i^\perp$ for all i, so $v \in \bigcap S_i^\perp$. Conversely, if we suppose $v \in \bigcap S_i^\perp$. Then $v \in S_i^\perp$ for all i, so $\langle v, x\rangle = 0$ for all $x \in S_i$ and all i. Thus we have $v \in (\bigcup S_i)^\perp$. This proves $(A \cup B)^\perp = A^\perp \cap B^\perp$. □

Proposition 7.11. *Let H be a Hilbert space, let $S \subseteq H$, and let $S' \subseteq S$. Then $S^\perp \subseteq (S')^\perp$.*

Proof. Let $v \in S^\perp$. Then $\langle v, x\rangle = 0$ for all $x \in S$. Since $S' \subseteq S$, we have $\langle v, x\rangle = 0$ for all $x \in S'$. So $v \in (S')^\perp$. And so $S^\perp \subseteq (S')^\perp$. □

Proposition 7.12. *Let H be a Hilbert space and V a subspace of H. Then $(\bar{V})^\perp = V^\perp$.*

Proof. We have $V \subseteq \bar{V}$, so by the above result, we know $\bar{V}^\perp \subseteq V^\perp$. Now let $w \in V^\perp$. We want to show $w \in \bar{V}^\perp$, i.e., $\langle w, v\rangle = 0$ for all $v \in \bar{V}$. There exists a sequence $(v_n) \in V$ such that $v_n \to v \in \bar{V}$. Now, by the continuity of the inner product, we have $\langle w, v_n\rangle \to \langle w, v\rangle = 0$. And so $w \in \bar{V}^\perp$. Thus $V^\perp \subseteq \bar{V}^\perp$. And we are done. □

Proposition 7.13. *Let H be a Hilbert space, and $S \subseteq S^\perp$. If $x \in S$ and $v \in S^\perp$, then $\|x + v\|^2 = \|x\|^2 + \|v\|^2$.*

Proof. $\|x + v\|^2 = \langle x, x\rangle + \langle x, v\rangle + \langle v, x\rangle + \langle v, v\rangle = \langle x, x\rangle + \langle v, v\rangle = \|x\|^2 + \|v\|^2$. □

Lemma 7.14. *Let H be a Hilbert space, and V be a closed subspace. Then $V \cap V^\perp = \{0\}$.*

Proof. Let $v \in V \cap V^\perp$. We want to show $v = 0$. So since $v \in V^\perp$, we have $\langle v, x\rangle = 0$ for all $x \in V$. Now since v is in the intersection, we can let $x := v \in V$, then $\langle v, x\rangle = \langle v, v\rangle = \|v\|^2 = 0$, and so $v = 0$. And so $V \cap V^\perp = \{0\}$. □

The following result is known as the *Hilbert space decomposition theorem*. The proof of this can be found in [27] and [4].

Theorem 7.15. *Let H be a Hilbert space. Let $V \subseteq H$ be a closed vector subspace. Then $H = V \oplus V^\perp$, i.e., $V \cap V^\perp = \{0\}$, and for all $v \in H$, we can write $v = v_1 + v_2$ for $v_1 \in V$ and $v_2 \in V^\perp$.*

So given a Hilbert space H and a closed vector space $V \subseteq H$, we have $x \in H$ can be uniquely written as $x = v + w$ where $v \in V$ and $w \in V^\perp$. And so we define the *orthogonal projection* map P onto V as $P : H \to H$ given by $P(x) = v$.

Corollary 7.16. *Let $S \subseteq H$, then* $\mathrm{Span}(S)$ *is dense in H if and only if $S^\perp = \{0\}$.*

Proof.

($\Rightarrow$) Let $\mathrm{Span}(S)$ be dense in H, then $\overline{\mathrm{Span}(S)} = H$. So $\overline{\mathrm{Span}(S)}^{\perp} = \{0\}$, i.e., $S^{\perp} = \{0\}$.

($\Leftarrow$) Now, let $S^{\perp} = \{0\}$. We have $\overline{\mathrm{Span}(S)}$ a closed subspace of H. By the Hilbert space decomposition theorem, $H = \overline{\mathrm{Span}(S)} \oplus \overline{\mathrm{Span}(S)}^{\perp} = \overline{\mathrm{Span}(S)} \oplus S^{\perp} = \overline{\mathrm{Span}(S)}$.

□

7.III Dual spaces

Let H be a Hilbert space, $v \in H$. Then, we define the linear map $J_v : H \to \mathbb{F}$ by $J_v(w) = \langle v, w\rangle$. It is clear that, by the Cauchy-Shwarz, for $w \in H$, we have $|J_v(w)| = |\langle v, w\rangle \leq \|v\|\|w\|$. So, J_v is a bounded linear map, with $\|J_v\| \leq \|v\|$.
Now, $|J_v(v)| = |\langle v, v\rangle = \|v\|^2$, so $\frac{J_v(v)}{\|v\|} = \|v\|$. Hence, $\|J_v\| \geq \|v\|$, so $\|J_v\| = \|v\|$.
The map J_v is an element of the dual space H^*.

The following result is known as *Riesz representation theorem*, which tells us that the above map is bijective.

Theorem 7.17. *Let $f : H \to \mathbb{F}$ be a bounded linear map. Then, we have $v \in H$ such that $f = J_v$.*

Proof. Consider $\ker f = \{v \in H : f(v) = 0\}$. And $\ker f$ is a vector subspace. So f is a bounded linear map, so it is continuous. So $f^{-1}(\{0\}) = \ker f$ is closed. Hence, $H = (\ker f) \oplus (\ker f)^{\perp}$. If $\ker f = H$, then $f = 0$ and $0 = J_0$, so we are done. Suppose $\ker f \neq H$, let $0 \neq x \in (\ker f)^{\perp}$. Let $v \in H$, we write $v = y + z$, for $y \in \ker H$ and $z \in (\ker H)^{\perp}$. Then, $f(v) = f(y) + f(z) = f(z)$, so $\langle x, v\rangle = \langle x, y\rangle + \langle x, z\rangle = \langle x, z\rangle$. Now, let $w \frac{x}{f(x)}$, so $f(w) = 1$, $\langle w, v\rangle = \langle w, z\rangle$. Now consider $f(v - f(v)w) = f(v) - f(v)f(w) = 0$. Hence, $v - f(v)w \in \ker f$, and $w \in (\ker f)^{\perp}$. So, $\langle w, v - f(v)w\rangle = 0$. Hence, $\langle w, v\rangle = f(v)\langle w, w\rangle$. So, define $u = \frac{w}{\|w\|^2}$, then $\langle u, v\rangle = f(v)$. □

Theorem 7.18. *Let H, H' be Hilbert spaces, and $T : H \to H'$ be a bounded linear map. Then we have a unique bounded linear map $T^* : H \to H$ called the adjoint of T, such that $\langle T^*x, y\rangle = \langle x, Ty\rangle$ for all $x \in H'$ and $y \in H$.*

Proof. Let $x \in H$. Then we have a bounded linear map $y \mapsto \langle x, Ty\rangle$. Now by Riesz's representation, we have a unique element $T^*x \in H$ such that $\langle T^*x, y\rangle = \langle x, Ty\rangle$ for all $y \in H$. □

Proposition 7.19. *Let $S, T : H \to H$ be bounded linear maps and let $\alpha, \beta \in \mathbb{F}$, then*

(i) $(\alpha S + \beta T)^* = \bar{\alpha}S^* + \bar{\beta}T^*$

(ii) $(S^*)^* = S$

(iii) $(ST)^* = T^*S^*$

Proof.

(i) for all $x, y \in H$, we have

$$\begin{aligned}\langle x, (\alpha(S) + \beta(T))y\rangle &= \alpha\langle x, Sy\rangle + \beta\langle x, Ty\rangle \\ &= \alpha\langle S^*x, y\rangle + \beta\langle T^*x, y\rangle\end{aligned}$$

$$= \langle \bar{\alpha} S^* x, y \rangle + \langle \bar{\beta} T^* x, y \rangle$$
$$= \langle \bar{\alpha} S^* x + \bar{\beta} T^* x, y \rangle = \langle (\bar{\alpha} S^* + \bar{\beta} T^*) x, y \rangle$$

(ii) for all $x, y \in H$, we have, $\langle x, S^* y \rangle = \overline{\langle S^* y, x \rangle} = \overline{\langle y, Sx \rangle} = \langle Sx, y \rangle$. Therefore we have $(S^*)^* = S$.

(iii) for all $x, y \in H$, we have

$$\langle x, (ST) y \rangle = \langle S^* x, Ty \rangle$$
$$= \langle T^* S^* x, y \rangle$$

Therefore, we have $(ST)^* = T^* S^*$.

□

Example 7.20. Consider $R : l^2 \to l^2$,
$R(a_1, a_2, a_3, \ldots) = (0, a_1, a_2, a_3, \ldots)$
$\langle (a_1, a_2, \ldots,), T(b_1, b_2, \ldots) \rangle = \langle (a_1, a_2, \ldots), (0, b_1, b_2, \ldots) \rangle = \overline{a^2} b_1 + \overline{a^3} b_2 + \overline{a_4} b_3 + \cdots = \langle (a_2, a_3, \ldots), (b_1, b_2, \ldots) \rangle = \langle L(a_1, a_2, \ldots), (b_1, b_2, \ldots) \rangle$, where $L(a_1, a_2, \ldots) = (a_2, a_3, \ldots)$.
So, $R^* = L$ and $L^* = R$.

Remark. If V is a finite dimensional vector space $T : V \to V$, which is injective, then $\ker T = \{0\}$; and by the Rank-Nullity theorem, $\dim \ker T + \dim \operatorname{im} T = \dim V$, we have $\dim \operatorname{im} T = \dim V$, which means $\operatorname{im} T = V$, so T is surjective.
In an infinite dimensional space $R : l^2 \to l^2$, we have R is injective but not surjective.

We call the elements $x, y \in H$, where H is a Hilbert space and $x \neq y$, *orthogonal* if $\langle x, y \rangle = 0$. And, indeed, if x, y are orthogonal, we see $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2$.

8 Banach Algebras

Now, we come to the important bit of Functional Analysis: Normed and Banach Algebras. We will learn the basics of Banach Algebras, and Spectral theory; concluding with the fact that the spectrum is non-empty.

This is mainly based on the presentations given in [4], [27], and [38].

Definition 8.1. A $\mathbb{F}$-vector space V is an *algebra* if there is a multiplication operation $V \times V \to V$, $(u, v) \mapsto u, v$ such that for all $u, v, w \in V$ and $\lambda, \mu \in \mathbb{F}$,

(i) $u(v + w) = uv + uw$ and $(u + v)w = uw + vw$

(ii) $(\lambda u)(\mu v) = (\lambda \mu)(uv)$

Definition 8.2. An algebra V over $\mathbb{C}$ is a *Banach algebra* if V is a Banach space and for all $u, v \in V$, we have $\|uv\| \leq \|u\| \|v\|$.

Remark. We call the algebra V over $\mathbb{C}$ a *normed algebra* if V is a normed space and it satisfies $\|xy\| \leq \|x\| \|y\|$ for all $x, y \in V$. If V is a complete normed vector algebra then it is a Banach algebra.
If V is a normed algebra with multiplicative identity, then V is a unital algebra.
If V is an algebra that is commutative as a ring, then it is a commutative algebra.

It can easily be seen that the operations of addition and multiplication in a normed algebra are continuous. Also, the operations of addition and multiplication extend continuously to the completion of a normed algebra; a completion of a normed algebra is a Banach algebra.

Example 8.3. Let V be a Banach space. We define $\mathcal{L}(H) = \{T : V \to V : \text{where } T \text{ is a bounded linear operator}\}$. We notice that, for $S, T \in \mathcal{L}(v), v \in V$, (multiplication in composition), $\|(ST)(v)\| = \|S(Tv)\| \leq \|S\| \cdot \|Tv\| \leq \|S\| \cdot \|T\| \cdot \|v\|$. So, for $v \neq 0$, we have $\frac{\|(ST)(v)\|}{\|v\|} \leq \|S\| \cdot \|T\|$ so then $\|ST\| \leq \|S\| \cdot \|T\|$.

Example 8.4. Let K be a compact topological space, then $C(K) = \{f : K \to \mathbb{C} \; f \text{ is continuous}\}$. For $f, g \in C(K)$ and $x \in K$, addition: $(f+g)(x) = f(x) + g(x)$; scalar multiplication: $(\lambda f)(x) = \lambda f(x)$; and multiplication: $(fg)(x) = f(x)g(x)$ (multiplication is commutative). The norm: $\|f\| = \sup\{|f(x) : x \in K\}$. So, for $f, g \in C(K), x \in K$, we see $|f(x)| \leq \sup_{x \in K} f(x) = \|f\|$, and same for g, so then $|f(x)g(x)| = |f(x)| \cdot |g(x)| \leq \|f\| \cdot |g(x)| \leq \|f\| \cdot \|g\|$, so after taking the supremum, we get $\|fg\| \leq \|f\| \cdot \|g\|$.

Let V be a unital Banach algebra over $\mathbb{C}$, i.e., we have a unit $1 \in V$ such that $1x = x1 = x$ for all $x \in A$. x is *invertible* if we have x^{-1} such that $x^{-1}x = xx^{-1} = 1$. It can easily be shown that, indeed, for each element x, there exists a unique inverse. Suppose not, and let $x \in A$, so there exists $y_1, y_2 \in A$ such that $y_1 x = 1 = xy_1$ and $y_2 x = 1 = xy_2$; and so $y_1 = 1y_1 = (y_2 x)y_1 = y_2 1 = y_2$.

Remark. $\|x^2\| \leq \|x\|\|x\| = \|x\|^2$, $\|x^3\| \leq \|x^2\|\|x\| \leq \|x\|^3$, etc....

The following result is known as the *Murray-von Nuemann criterion*.

Theorem 8.5. *Let A be a unital Banach algebra and $x \in A$. If $\|x\| < 1$, then $1 - x$ is invertible.*

Proof. We first notice that in $\mathbb{C}$ if $|\lambda| < 1$, then $1/(1-\lambda) = 1 + \lambda + \lambda^2 + \lambda^3 + \cdots$. So, we claim $(1-x)^{-1} = 1 + x + x^2 + \cdots$, if $\|x\| < 1$. Let $S_n = 1 + x + x^2 + \cdots + x^n$. For $n > m$, we have $S_n - S_m = x^{m+1} + x^{m+2} + \cdots + x^n = x^{m+1}(1 + x^2 + \cdots + x^{n-m-1})$. So, then

$$\begin{aligned}
\|S_n - S_m\| &\leq \|x^{m+1}\|\|1 + x + \cdots + x^{n-m+1}\| \\
&\leq \|x\|^{m+1}(1 + \|x\| + \|x\|^2 + \cdots + \|x\|^{n-m-1}) \\
&\leq \|x\|^{m+1}(1 + \|x\| + \|x\|^2 + \cdots) \qquad (\|x\| < 1) \\
&= \frac{\|x\|^{m+1}}{1 - \|x\|} \to 0 \text{ as } m \to \infty
\end{aligned}$$

So, (S_n) is Cauchy, and so it converges as A is a Banach algebra. Now, let $S_n \to S$ as $n \to \infty$, then $S_n(1-x) = (1+x+\cdots+x^n)(1-x) = 1+x+\cdots+x^m - x - x^2 - \cdots - x^{n+1} = 1 - x^{n+1}$. So, $\|S_n(1-x) - 1\| \leq \| - x^{n+1}\| \leq \|x\|^{n+1} \to 0$ as $n \to \infty$ as $\|x\| < 1$. So $S(1-x) - 1 = 0$, i.e., $S(1-x) = 1$, similarly, $(1-x)S = 1$. □

Corollary 8.6. *Given a unitral Banach algebra A, let $x \in A$ and $\lambda \in \mathbb{F} \setminus \{0\}$. If $\|x\| < |\lambda|$, then $\lambda - x$ is invertible.*

Proof. Since $\|x\| < \lambda$, then $\|x/\lambda\| < 1$, and by the above result, we know $1 - x/\lambda$ is invertible in A. We claim the inverse to be $\frac{(1-x/\lambda)^{-1}}{\lambda}$. And, indeed,

$$(\lambda - x)\frac{1 - x/\lambda}{\lambda} = (1 - x/\lambda)(1 - x/\lambda)^{-1} = 1$$

So, by the above result again, we find the inverse to be

$$(\lambda - x)^{-1} = \frac{1}{\lambda}(1 + \frac{x}{\lambda} + \frac{x^2}{\lambda^2} + \cdots).$$

□

Definition 8.7. Let A be a unital complex Banach algebra, $x \in A$. Then we define the *spectrum*, $\sigma(x) = \{\lambda \in \mathbb{C} : x - \lambda I \text{ is not invertible}\}$.

Example 8.8. Let V be a Banach space, and $T \in \mathcal{L}(v)$. Let λ be an eigenvalue of T, then $Tv = \lambda v$ for some $v \neq 0$. i.e. $(T - \lambda I)_v = 0$ where $v \neq 0$. This implies $\ker(T - \lambda I) \neq \{0\}$, and so $T - \lambda I$ is not injective and not invertible.
If V is finite-dimensional, then if $T - \lambda I \in \mathcal{L}(v)$ is not invertible, then $\det(T - \lambda I) = 0$. So, λ is an eigenvalue.

Theorem 8.9. *Let A be a unital complex Banach algebra, $x \in A$. Then, $\sigma(x)$ is a non-empty compact subset of $\mathbb{C}$.*

We first prove part of the statement (i.e., it is compact). Then we will soon prove the rest of the statement.

Proof (I). $\sigma(x)$ is closed by the Murray-von Nuemann criterion. And $\sigma(x)$ is bounded, since if we let $\lambda > \|x\|$, then $\|x/\lambda\| < 1$, so $1 - x/\lambda$ is invertible. And so, $-\lambda(1 - x/\lambda) = x - \lambda$ is invertible and $\lambda \notin \sigma(x)$. So, if $\lambda \in \sigma(x)$, then $|\lambda| \leq \|x\|$. And so $\sigma(x)$ is compact by Heine-Borel. □

Definition 8.10. Let A be a unital Banach algebra, $x \in A$. The *resolvent set* of x is the complement of the spectrum, $\rho(x) = \mathbb{C} \setminus \sigma(x)$, i.e., $\rho(x) = \{\lambda \in \mathbb{C} : x - \lambda \text{ is invertible}\}$. We also define the *resolvent* map $R_{\rho(x)} : \rho(x) \to A$, given by $\lambda \mapsto (\lambda - x)^{-1}$. So if $\lambda \in \rho(x)$, then the resolvent is $R_\lambda(x) = (\lambda - x)^{-1}$.

The next result shows the resolvent is bounded.

Proposition 8.11. *Let $x \in A$, then $\|R_\lambda(x)\| \to 0$ as $\lambda \to \infty$.*

Proof. $\|(\lambda - x)^{-1}\| = \|\lambda^{-1}(1 - x/\lambda)^{-1}\|$, by taking limits, we see $\|(\lambda - x)^{-1}\| \to 0$, as $\lambda \to \infty$. □

The following result shows the resolvent identity.

Lemma 8.12. *Let $\lambda, \mu \in \rho(x)$. Then $R_\lambda - R_\mu = (\mu - \lambda)R_\lambda R_\mu$.*

Proof. We first see $(\mu - x) - (\lambda - x) = \mu - \lambda$. And so $R_\lambda - R_\mu = R_\lambda((\mu - x) - (\lambda - x))R_\mu = (\mu - \lambda)R_\lambda R_\mu$. □

Definition 8.13. Let $D \subseteq \mathbb{C}$ be a connected open set. Let V be a complex Banach space. A function $f : D \to V$ is *holomorphic* if there is a function $f' : D \to V$, the derivative of f, such that

$$\lim_{h\to 0} \left\| \frac{f(z+h) - f(z)}{h} - f'(z) \right\| = 0.$$

Lemma 8.14. *Let A be a complex unital Banach algebra, $x \in A$, and $f(\lambda) = R_\lambda(x)$. Then there exists a holomorphic function $f : \rho(x) \to A$.*

Proof. Let $g(\lambda) = -R_\lambda(x)^2 = -(\lambda - x)^{-2}$. We want to show that as $h \to 0$,

$$\left\| \frac{f(\lambda+h) - f(\lambda)}{h} - g(\lambda) \right\| \to 0.$$

So, then

$$\left\| \frac{f(\lambda+h) - f(\lambda)}{h} - g(\lambda) \right\| = \left\| \frac{1}{h}((\lambda - x + h)^{-1} - (\lambda - x^{-1})) + (\lambda - x)^{-2} \right\|.$$

Now, let for $h > 0$, by the above results, $\lambda - x + h$ is invertible, and so we have $(\lambda - x + h)^{-1} = (\lambda - x)^{-1}(1 - h(\lambda - x)^{-1} + h^2(\lambda - x)^{-2} - \cdots)$. And so

$$\left\| \frac{f(\lambda+h) - f(\lambda)}{h} - g(x) \right\| = \|h(\lambda - x)^{-3} - h^2(\lambda - x)^{-4} + \cdots\|.$$

And so the result follows from there. So $f(\lambda)$ is holomorphic with $g(\lambda)$ as its derivative. □

The next result is known as *Liouville's theorem*.

Theorem 8.15. *Let A be a complex Banach space. Let $f : \mathbb{C} \to A$ be a bounded holomorphic function. Then f is constant.*

Sketch (idea). Let $g \in A^*$, so g is a map $A \to \mathbb{C}$. We want to show $g \circ f : \mathbb{C} \to \mathbb{C}$ is a bounded holomorphic function, so that we can use Liouville's theorem on $\mathbb{C}$ to show that it is constant. So, $|g \circ f(\lambda)| \leq \|g\|\|f\|$, and since f is bounded and holomorphic, and by Hahn-Banach, g is also bounded and holomorphic, we have that by Liouville's theorem on $\mathbb{C}$, the function is constant and holomorphic. So, $g(f(\lambda) - f(\mu)) = 0$, i.e. $g(f(\lambda)) = g(f(\mu))$. And by Hahn Banach again, we see $f(\lambda) = f(\mu)$, thus f is constant. □

Now, we can prove that the spectrum is non-empty (Theorem 8.9).

Proof (II). Suppose $\sigma(x) = \emptyset$, then $\rho(x) = \mathbb{C}$. And by the above lemma, there exists a holomorphic function $f : \mathbb{C} \to A$ where $f(\lambda) = (\lambda - x)^{-1}$. And by above results, we know $\|f(\lambda)\| \to 0$ as $\lambda \to \infty$. So, it follows by Louiville's theorem $f(\lambda) = 0$ for all $\lambda \in \mathbb{C}$, i.e., $(x - \lambda)^{-1}$ for all λ. This is a contradiction since $R_\lambda \neq R_\mu$ for $\lambda \neq \mu$. □

The next result is called the *Spectral mapping theorem for polynomials*. The proof for this can be found in [27] and [4].

Theorem 8.16. *Let A be a unital Banach algebra, $a \in A$, and let p be a polynomial. Then $\sigma(p(x)) = p\sigma(x)$.* □

8.I C*-algebras Brief

In this section we will briefly study C^*-algebras, which is a Banach algebra together with an involution.

This is based on the presentations given in [8] and [26].

Definition 8.17. An *involution* on an algebra A over $\mathbb{C}$ is a map $A \to A$, given by $x \mapsto x^*$, such that

(i) $(\lambda x + \mu y)^* = \bar{\lambda} x^* + \bar{\mu} y^*$ for all $x, y \in A$ and $\lambda, \mu \in \mathbb{F}$

(ii) $(x^*)^* = x$ for all $x \in A$

(iii) $(xy)^* = y^* x^*$ for all $x, y \in A$

A *C^*-algebra* A is a Banach algebra over $\mathbb{C}$ together with an involution, such that $\|x\|^2 = \|x^* x\|$ for all $x \in A$.

We say that a norm on a Banach algebra A together with an involution such that $\|x\|^2 \leq \|x^* x + y^* y\|$ for all $x, y \in A$ is a *C^*-norm*. The completion of a Banach algebra with an involution, is a C^*-algebra.

Example 8.18. The set of bounded operators on a Hilbert space H, $\mathcal{L}(H)$, is a C^*-algebra, by defining the involution to be the adjoints in H.
The complex numbers $\mathbb{C}$ forms a C^*-algebra, by defining the involution $\lambda^* = \bar{\lambda}$ and where the norm is $\|\lambda\| = |\lambda|$.

Given a C^*-algebra A, if A has a unit $1 \in A$, i.e., an identity element. Then it can easily be shown $1^* = 1$. This can be seen by $1 = (1^*)^* = (1^* 1)^* = (1^* (1^*)^*) = 1^* 1 = 1^*$.

Proposition 8.19. *Given a C^*-algebra A, the involution is continuous.*

Proof. Let $x_n \in A$ such that $x_n \to x \in A$. So, $\|x_n^* - x^*\| = \|(x_n - x)^*\|$. Now we want to show $\|(x_n - x)^*\| = \|x_n - x\|$. So $\|x_n - x\|^2 = \|(x_n - x)^*(x_n - x)\| \leq \|(x_n - x)^*\| \|x_n - x\|$, so if $x_n - x \neq 0$ then $\|x_n - x\| \leq \|(x_n - x)^*\|$ (otherwise if $x_n - x = 0$, then we are done). So, $\|(x_n - x)^*\| \leq \|((x_n - x)^*)^*\| = \|x_n - x\|$. So, then, $\|(x_n - x)^*\| = \|x_n - x\|$.
So, we now have $\|x_n^* - x^*\| = \|x_n - x\| \to 0$ since $x_n \to x$. Therefore, the involution is continuous. □

Proposition 8.20. *If A is a C^*-algebra, then $\sigma(x^*) = \sigma(x)$ for all $x \in A$.* □

Given a C^*-algebra A. We call $x \in A$ *normal*, if $xx^* = x^* x$. We call $x \in A$ *Hermitian* if $x = x^*$.

The following result is left as an exercise to the reader.

Proposition 8.21. *If x is an element of a C^*-algebra A. Then $x^* x$ is Hermitian.* □

Definition 8.22. Given C^*-algebras A and B, a **-homomorphism* is an algebra homomorphism $\alpha : A \to B$ such that $\alpha(x^*) = \alpha(x)^*$ for all $x \in A$. A *-homomorphism, α, between unital C^*-algebras is itself *unital* if $\alpha(1) = 1$.

Theorem 8.23. *Any *-homomorphism is continuous, with closed image, and is bounded with norm at most one. Any injective *-homomorphism is an isometry.* □

So, the norm on a C^*-algebra is determined by the rest of the algebraic structure.

A *-homomorphism is an *isomorphism* of C^*-algebras if it is bijective.

Proposition 8.24. *The inverse of a bijective *-homomorphism is also a *-homomorphism.* □

A List of Inference Rules

Conjunction:

$$\frac{A \quad B}{A \wedge B}\,\wedge\text{-int} \qquad \frac{A \wedge B}{A}\,\wedge\text{-elim} \qquad \frac{A \wedge B}{B}\,\wedge\text{-elim}$$

Disjunction:

$$\frac{A}{A \vee B}\,\vee\text{-int} \qquad \frac{B}{A \vee B}\,\vee\text{-int}$$

$$\frac{A_i \vee \cdots \vee A_j \qquad \begin{array}{c}[A_i]\\ \vdots\\ C\end{array} \qquad \cdots \qquad \begin{array}{c}[A_j]\\ \vdots\\ C\end{array}}{C}\,\vee\text{-elim}$$

Implication & *modus ponendo ponens*:

$$\frac{A \quad A \Rightarrow B}{B}\,\Rightarrow\text{-elim} \qquad \frac{\begin{array}{c}[A]\\ \vdots\\ C\end{array}}{A \Rightarrow C}\,\Rightarrow\text{-int}$$

falsum & contradictum:

$$\frac{\bot}{A}\,\bot\text{-elim} \qquad \frac{A \quad \neg A}{\bot}\,\textit{contradictum}$$

reductio ad absurdum:

$$\frac{\begin{array}{c}[A \to \bot]\\ \vdots\\ \bot\end{array}}{A}\,\neg\text{-elim}$$

Equality:

$$\frac{}{t = t}\,=\text{-int} \qquad \frac{s = t \quad \varphi[s/x]}{\varphi[t/x]}\,=\text{-elim}$$

For All Quantifier:

$$\frac{\forall x : \varphi}{\varphi[t/x]}\,\forall\text{-elim} \qquad \frac{\left[\begin{array}{c}t\\ \vdots\\ \varphi[t/x]\end{array}\right]}{\forall x : \varphi}\,\forall\text{-int}$$

Exists Quantifier:

$$\frac{\varphi[t/x]}{\exists x : \varphi}\,\exists\text{-int} \qquad \frac{\exists x : \varphi \qquad \left[\begin{array}{c}t\\ \varphi[t/x]\\ \vdots\\ p\end{array}\right]}{p}\,\exists\text{-elim}$$

B List of ZFC Axioms

Axiom 1 - Empty set. $\exists x : \forall y : (y \notin x)$
There exists a set that contains no elements.

Axiom 2 - Extensionality. $\forall x : \forall y : (x = y) \Leftrightarrow ((x \subseteq y) \wedge (y \subseteq x))$
Two sets are equal if and only if they have the same elements.

Axiom 3 - Pair set. $\forall x : \forall y : \exists m : \forall u : (u \in m) \Leftrightarrow ((u = x) \vee (u = y))$
Given sets x and y. There exists a set whose elements are exactly x and y. We will define this set, m, as $\{x, y\}$.

Axiom 4 - Union set. $\forall x : \exists u : \forall y : ((y \in u) \Leftrightarrow (\exists s : (y \in s) \wedge (s \in x)))$
Given a set x. There exists a set u, whose elements are the elements of the elements of x. We denote this set u as, $\bigcup x$.

Axiom 5 - Replacement schema.

$$\forall x : [\forall y, z, u : ((\phi(y, z) \wedge \phi(y, u)) \Rightarrow (z = u)) \Rightarrow (\exists m : \forall t : (t \in m) \Leftrightarrow (\exists s \in x : \phi(s, t)))]$$

Given a set x and a formula $\phi(y, t)$, where for any y there exists at most one t such that $\phi(y, t)$. Then there exists a set m of all those t, written as $\{t : \exists y \in x : \phi(y, t)\}$.

Axiom 6 - Power set. $\forall x : \exists y : \forall a : (a \in y) \Leftrightarrow (a \subseteq x)$
Given a set x, there exists a set y, denoted as $\mathcal{P}(x)$, whose elements are the subsets of x.

Axiom 7 - Infinity. $\exists x : (\emptyset \in x) \wedge (\forall y : (y \in x) \Rightarrow (y \cup \{y\} \in x))$
Given a set y. There exists a set that contains the empty set, y, and $y \cup \{y\}$.

Axiom 8 - Choice. $\forall x : P(x) \Rightarrow \exists y : \forall a \in x : \exists! b \in a : a \in y$, where $P(x) \Leftrightarrow (\exists a : (a \in x) \wedge (\forall a : \forall b : (a \in x) \wedge (b \in x)) \Rightarrow \bigcap\{a, b\} = \emptyset)$
Give a set x, whose elements are non-empty and mutually disjoint. There exists a set y which contains exactly one element of each element of x.

Axiom 9 - Foundation. $\forall x : \exists a : (a \in x) \Rightarrow (\exists y \in x : \bigcap\{x, y\} = \emptyset)$
Every non-empty set has a minimal element. More precisely, every non-empty set x contains an element y that has none of its elements in common with x.

C Latin Phrases (Glossary)

ad infinitum: "to infinity" or "and so on"

contradictum: "contradiction"

de facto: "in fact"

ex falso sequitur quodlibet: "from falsehood anything follows"

falsum: "false"

id est: "that is" or more commonly abbreviated as 'i.e.'

ipso facto: "follows from the fact"

modus ponendo ponens: "mode by affirming afirms" which means given the antecedent and the implication (antecedent $\Rightarrow$ conclusion), then we can deduce the conclusion

modus tollendo tollens: "mode by denying denies" which is the *opposite* of the above. Given the negation of the conclusion and the implication, then we can deduce the negation of the antecedent

reductio ad absurdum: "reduction to absurdum" also known as proof by contradiction

tertium non datur: "the third is not given" or known by its inference law 'law of excluded middle (LEM)' which just means that either p is true or $\neg p$ is true

verum: "true"

References

[1] Banach, S., *Théorie des Opérations Linéaires*, Monografje Matematyczne, vol. 1, Warsaw, 1932.

[2] Barnes, D. W. & Mack J. M., *An Algebraic Introduction to Mathematical Logic*, Springer, 2006.

[3] Bowers, A. & Kalton, N. J., *An Introductory Course in Functional Analysis*, Springer, 2014.

[4] Cambridge Part II, *Linear Analysis*, 2019.

[5] Cambrdige Part II, *Logic and Set Theory*, 2019.

[6] Cambridge Part III, *Formal Logic*, 2019.

[7] Conway, J. B., *A Course in Functional Analysis*, Springer GTM, 1990.

[8] Davidson, K. R., *C*-Algebras by Example*, AMS, 1996.

[9] Dedekind, R., *Essays on the Theory of Numbers*, Chicago, The Open Court Pub. Co., 1901.

[10] Devlin, K., *The Joy of Sets*, Springer-Verlag, 1979.

[11] Enderton, H. B., *Elements of Set Theory*, Elsevier, 1977.

[12] Goldrei, D., *Classic Set Theory*, Chapman & Hall, 1998.

[13] Grattan-Guinness, I., *The Search For Mathematical Roots*, Princeton University Press, 2000.

[14] Haggarty, R., *Fundamentals of Mathematical Analysis*, Addison-Wesley, 1992.

[15] Hamilton, A. G., *Numbers, Sets, and Axioms*, Cambridge University Press, 1982.

[16] Hewitt, E. & Stromberg K., *Real and Abstract Analysis*, Springer, 1975.

[17] Kneebone, G. T., *Mathematical Logic and the Foundations of Mathematics*, D. Van Nostrand Co. Ltd., 1963.

[18] Kunen, K., *Set Theory*, Elsevier, 1980.

[19] Landau, E., *Grundlagen Der Analysis*, Chelsea Pub. Co., 1951.

[20] Levy, A., *Basic Set Theory*, Springer, 1979.

[21] Levy, A., Fraenkel, A. A., & Bar-Hillel, Y., *Foundations of Set Theory*, Elsevier, 1973.

[22] MacCluer, B. D., *Elementary Functional Analysis*, Springer, 2009.

[23] Mendelson, E., *Introduction to Mathematical Logic*, Taylor & Francis Group, 2010.

[24] Magic Course, *Functional Analysis*, 2019.

[25] Magic Course, *Set Theory*, 2020.

[26] Mitchener, P., *Analytic K-theory*, 2011.

[27] Mitchener, P., *MAS439 Functional Analysis*, The University of Sheffield, Lecture notes, 2019.

[28] Mitchener, P., *Real and Complex Analysis*, 2010.

[29] Moller, F. & Struth, G., *Modelling Computing Systems*, Springer, 2013.

[30] Moschovakis, Y., *Notes on Set Theory*, Springer, 2000.

[31] Oxford Part B, *B1.1 Logic*, 2020.

[32] Oxford Part B, *B1.2 Set Theory*, 2020.

[33] Oxford Part B, *B4.1 Functional Analysis I*, 2019.

[34] Oxford Part B, *B4.2 Functional Analysis II*, 2020.

[35] Oxford Part C, *C1.1 Model Theory*, 2020.

[36] Oxford Part C, *C1.4 Axiomatic Set Theory*, 2020.

[37] Ringrose, J. R. & Kadison, R. V., *Fundamentals of the Theory of Operator Algebras*, AMS, 1997.

[38] Rudin, W., *Functional Analysis*, McGraw-Hill, 1991.

[39] Rudin, W., *Principles of Mathematical Analysis*, McGraw-Hill, 1976.

[40] Rudin, W., *Real and Complex Analysis*, McGraw-Hill, 1987.

[41] Schechter, E., *Handbook of Analysis and Its Foundations*, Academic Press, 1997.

[42] Shapiro, S., *Philosophy of Mathematics and Logic*, Oxford University Press, 2005.

[43] Sternberg, S. & Loomis, L. H., *Advanced Calculus*, Addison-Wesley Pub. Co., 1968.

[44] Suppes, P., *Axiomatic Set Theory*, D. Van Nostrand Co. Inc., 1960.

[45] The University of Sheffield Lectures, *Metric Spaces*, 2019.

Index