Mathematical Foundations

Harsh Prajapati

9.01.26

These notes were prepared between October 2025 and January 2026 **(Last update: January 4, 2026)**.
If you find any mistakes or typos, please report them to **caccacpenguin@gmail.com**. I would really appreciate it.
I often use informal language to make the ideas easier to grasp. My goal is to make the material feel approachable, while still respecting the rigor that makes mathematics what it is.
I hope you will find these notes helpful :D!

References

These lecture notes closely follow these two lecture notes' preliminary sections (section 1 in both notes). However, they are written in German.

- Analysis I — WiSe 2016/17, by Franz Merkl, Facultät für Mathematik, Informatik und Statistik, LMU München.
- Lineare Algebra I — WiSe 2023/24, by Roland Herzog, Interdisciplinary Center for Scientific Computing, Heidelberg University.

There are many books that can serve as excellent resources, some of them as listed below.

- Michael Junk and Jan-Hendrik Treude, Beweisen lernen Schritt für Schritt, Springer-Verlag, 2020.
- Felix Göbler, Alex Küronya, Einstieg in die beweisorientierte Mathematik, Springer-Verlag, 2023.

Contents

1 Introduction

You must have been told that to study courses like Analysis and Algebra, you first need to learn proof writing and logic. But why is that? After all, we studied mathematics in school too, without any issues, so why the sudden need for formality?

The usual answer is, "You don't just learn methods now, you prove them." But what does it really mean to prove something? And why must we do it in this particular, formal way? We used to do "proofs" in school too, so what's changed? Isn't that enough? The common reply is, "Those were just simple informal deductions." But then, what exactly is the difference between an informal and a formal proof?

Some of you might be thinking, "*Isn't this just how maths is supposed to be? You need to show that something is true, so why even ask such questions?*"

Sure, teachers told us that mathematics is about proving things axiomatically. But remember how, in school, we were given a "visual proof" of the Pythagorean theorem? You draw a right triangle, construct squares on each side, and if you measure their areas, the two smaller ones together equal the larger one on the hypotenuse. Isn't that a proof too? After all, teachers presented it confidently, and we accepted it without question.

But if you try the same argument in a university-level maths course, you'll for sure get cooked.

Visual proofs like these can convince us that a statement *must* be true, but they're not considered formal proofs, not until they can be derived *purely* from axioms and logical reasoning.

So what makes a formal proof so necessary?

These questions aren't silly at all. In fact, they're the very same ones that puzzled mathematicians like Hilbert, Russell, Cantor, and Gödel during the late 19th and early 20th centuries, a period often called the "*crisis of rigour*" in mathematics. And these questions should puzzle you too.

Part of what made this crisis so fascinating is that it wasn't just about maths, it was about the limits of human reasoning itself. Our brains aren't naturally built to handle pure abstraction. You're far more likely to accept that the volume of a cylinder is three times that of a cone with the same height and base area if someone demonstrates it to you using flasks and water, rather than writes a string of abstract symbols on a piece of paper.

We're wired to trust what we can see more than what we can only think. That's why theories like Einstein's relativity or quantum mechanics feel so counterintuitive, they defy the patterns our senses are tuned to believe. And for a long time, mathematicians felt the same way: a convincing demonstration or geometric diagram was considered proof enough.

So what changed in the early 20th century? Why did mathematicians begin to trust formal proofs, long chains of logical deductions, more than vivid demonstrations or intuitive arguments?

And yet, after all their efforts, mathematics still isn't completely consistent or perfectly justified from within itself. The cracks they uncovered are subtle but profound, mathematics contains both incompleteness and undecidability, and understanding these ideas is part of understanding what mathematics really is.

The **prerequisites** for this course is only some background in elementary mathematics taught in schools, I don't even assume the readers to be 'mathematically matured', rather I will try to develop that thinking from ground up.

This course is essentially the "first installment" in a series of lecture notes i'll be writing on mathematics and mathematical physics :3. Starting from foundational topics

like analysis and algebra and going towards advanced topics in algebraic topology, PDEs, Geometry, and Stochastic Processes. In Mathematical Physics, i treat the subjects very rigorouly (sometimes more than required and typically taught in undergraduate courses) to ensure mathematical sophistication for advanced research. The aim of these notes is to train a future researcher in mathematics or/and mathhematical physics.

HuiiiHuiii :3

2 Logic

2.1 Where do we begin?

Before we even start throwing around symbols, let's take a step back and ask some of the most fundamental questions about the foundations of mathematics. *Where do we actually begin? What's the first object, the first axiom or the most primitive structures?*

Most will say: you should start from sets and logic, treat sets as the primitive and accept Extensionality, the separation schema, and the inference rules of first-order logic as your "first axioms". ZFC together with first-order logic can, in principle, encode virtually all of ordinary mathematics[1]— but you still need some underlying meta-logic to even state ZFC properly, and our formal systems of logic itself has some undecidability.

Others prefer types or categories, where sets themselves are reconstructed as types with rules for forming terms, and categorical foundations in the ETCS approach take objects and morphisms as primary and stress structure-preserving maps. But even those approaches rely on informal notions once you step outside the formal system.

These frameworks aren't really isolated, they overlap quite a bit and can often sort of "inter-translate" into one another. In practice you can often model type-theoretic constructions as set-theoretic objects, and you can formalise categories (and even categories of models) within set theory, although that gets tricky due to size issues of small vs large categories.

You could also start from talking about formal systems themselves — studying what can be proved within them in proof theory, reverse mathematics, and other areas of metamathematics.

Or, we could go beyond that, discussing the philosophical stances like logicism, formalism, intuitionism and structuralism which differ in how they view the relationship between mathematics, logic, and meaning.

There are some modern frameworks like Homotopy Type Theory and Topos Theory which are attempts to solve he problems of formalisation, whereas Feferman's system of explicit mathematics and Voevodsky's Univalent Foundations are some active research areas which blur the boundary between object and meta-level.

You might be thinking, *"But, I don't understand any of this jargon! What do you even mean?"*

Don't worry, nobody really understands all of that jargon, people just hide behind it.

What I mean is: you can't define everything from some absolute bottom layer without already standing on some other framework. And if you seriously start thinking about the foundations, it can really keep you up at night!!

Of course I don't give such an absurd amount of coverage, the above discussion was just to give you a "zoom out" overview of the how the landscape of the foundations looks like.

The takeaway is that foundations aren't about discovering a single metaphysical "first object" — they're about choosing what your primitive notions and rules will be. Even though sets and logic are where almost every university mathematics program begins, they aren't the only possible starting point. That tradition mostly comes from a mix of pedagogy, culture, and the historical development of mathematics education.

[1]the mathematics where you "do maths" over objects and proof theorem (such as in analysis or algebra) rather than discussing about their philosophy which would be "foundational mathematics"

2.1.1 Epistemology and Regress Argument*

In most textbooks, **logic** is defined as, "the study of reasoning".

Okay, but what is *reasoning*?

Well, reasoning is, "the act of logical deduction".

And... what is logical deduction...?

Logical deduction is, "the process of deducing a conclusion from premises using symbolic logic."

But-huh...? W-WHAT?!

See, if you try to define everything precisely you can easily run into circulalarity or getting spooked by Euthypohro's ghost [2].

Each definition presuppose some other informal notion we haven't defined yet, that's the **regress problem** in epistemology [3]. Any definition, idea of belief which can be infinity quesitioned, results in endless regress.

So what do we do in such situations? Well, one solution is to accept that some chain of beliefs start with a **basic belief** which does not need to be justified by some other belief, and all beliefs that follow it are justified by the basic belief. It's sort of like building architecture, you start with a base and everything on top of it stands on the base. that's the **foundationalism** approach, trying to escape the regress argument.

"But what about the circle?"

Yes, in our case, the idea of logic and reasoning doesn't seem to have a clear base, rather it seems to be circular. That's where **coherentism** comes in, which says that beliefs are justified a coherent system of mutually supporting beliefs. So we have a heurtistic web instead of a linear chain.

But, since it relies on the idea that circular reasoning is acceptable, in this view, one belief ultimately supports itself. Coherentists reply that it is not just that the belief is supporting itself, but that belief along with the totality of the other statements in the whole system of beliefs support each other.

Now, one observation here is that one belief can cohere and justify two different beliefs without any of the three being true. But Coherentists argue that it is unlikely for a whole system of beliefs to be false if they cohere well together, if some parts of the system were untrue it would certainly be inconsistent with some other part of the system.

Okay, but if coherentism claims that every belief is justified by its coherence with other beliefs, then what about those beliefs which seem to arise from our experiences and not from other beliefs? Such as the white canopy bed example: you look into a totaly dark room, the lights turn on momentarily and you see a white canopy bed in the room. You may say, *"I saw a white canopy bed, therefore I believe that there is a white canopy bed inside this room."* This belief seem to based entirely on your experience, you don't need other beliefs to justify it, because you *"saw it"*. So, it seems like beliefs can be justified by concepts other beliefs, such as experiences and perceptions which coherentism doesn't takee into account. But others have argued that the experiences of seeing the bed is indeed dependent on other beliefs, about what a bed, a canopy and so on, actually looks like.

Now, **Infinitism** argues that the chain can go on forever. But that seems like a restatement of the problem itself and not a solution[4].

[2] see *Plato, "Euthyphro"* for an account of the dialogue between Socrates and Euthyphro

[3] see *Nicolas Rescher, "Epistemology: An Introduction to the Theory of Knowledge"* for a detailed account.

[4] this is very simplified explanation but the point here is that infinitism doesn't help us as much as the other two.

You may feel a bit skeptic about the above approaches and argue that the beliefs cannot be justified without doubts. And if you closely examine the so far discussed approaches it does seem like they point to the same conclusion that it's really hard to justify some beliefs, axioms or laws which seem fundamental. And that is the takeaway of all of this jargon.

"So, what does all of this have to do with logic, sets and proofs?"

In mathematical logic, we don't really deal with 'beliefs' but rather axioms and propositions[5], but the idea is the same. And this is the focus point of the section, where we connect all of this jargon to actual mathematics and discuss that all axioms cannot be proven and we must accept them as unjustified facts, and our formal systems cannot be justified internally without certain assumptions that are outside the system.

The epistemological problem of justification and the logical problem of incompleteness are of the same shape.

2.1.2 What is Mathematical Logic?

We stumbled upon the problem of defining concepts in mathematics and by a lot of discussion reached a conclusion that it is not possible to define everything precisely, so how do we do mathematics then?

Mathematics is just made up of lots of concepts and we try to reduce these concepts to certain axioms which are do not need to be justified. Axioms are analogous to the basic beliefs we discussed in the section 2.1.1.

"But, why only *these* axioms?"

This is quite hard to reason without relying on our informal idea of what we feel should be the primitive laws from which we choose to derive other laws. This is very similar to the foundationalist approach, but mathematics not just foundationalist, it is a mix of this philosophy aliong with coherentism.

2.1.3 Axiom System

Just like basic beliefs, an **axiom** is a primitive law which we accept without proof. All other laws called **theorems** can be derived from axioms. With mathhematical concepts we have cetain **basic concepts** which are not to be quesitioned, and **derived concepts** which can be derived from the basic concepts.

In any statement we can replace the derived concepts with basic concepts such as with axioms. We assume that all concepts which appear in axioms are basic concepts.

"Okay, but what do we do with it?"

So, to develope any mathematical theory we present certain basic concepts and axioms and we explain those axioms until it's clear that the axioms are true. Then we use those axioms to prove theorems and derive other derived concepts.

The entire edifices of basic concepts, derived concepts, axioms and theorems is called an **axiom system**. It could be an axioms system for all of mathematics or just a specific part, such as non-Euclidean geometry.

This definition of axiom system assumes that the axioms cannot be derived or proven from other fundamental concepts, hence they is sometimes also called **classical** axiom systems. However, it is sometimes possible to realise the axioms from other concepts, in such case mathematicians frame axiom systems in which axioms are large number of

[5]these terms will be defined precisely in the next two sections

concepts, such as the axioms for groups. Such an axiom system is called **modern** axiom system. The difference between the two is not huge, it merly depends on the intentions of the framer of the system.

This is exactly how every mathematical theory is studied, you will notice that this is actually the structure a formal textbook of mathematics follows. And we shall start the study of mathematical logic with the study of axiom systems.

2.1.4 Formal System

We introduced axiom in the previous section but what does axiom really represent? It turns out that axioms can be interpreted in two sense, as a statement: objects which the axiom mentions when we write it on paper, or as the *meaning* of the sentence: the fact being claimed by the axiom.

"What is the difference??" The difference is that first is the *syntactic* view, we are interested in the "objects"(vocabulary/symbols) the axiom is talking about and the second is *semantic* view, we are interested the *fact*(or relationship) about(between) the objects being claimed.

Let's see an example to make this clear.

Example 2.1. Consider the axiom: *Thomas is taller than Max.*

- **Syntactic view**: "Thomas" and "Max" are the objects of the axiom.
- **Semantic view**: The relationship being asserted is: "One object *is taller than* the other object"

Of course, this axiom assumes prior understanding of the concept of "taller than" and what it means for two objects to have such relationship, without knowing that, the two objects alone do not convey the fact being asserted. This hightlights that the semantic view relies on interpretation of the symbols in the axiom.

This differentiation of the structure of a sentence is very useful. Separating the syntactic and semantic part of our study of axiom systems and by using suitable language, the structure of the sentence would reflect the meaning of the axiom to some extent.

2.2 Propositional Logic

To be able to work with mathematical statements and be able to prove them we need precise language. English or any human language are not well suited for such purposes since there are many ambiguity when using these languages, for example, if you say "I will study mathematics or physics in university" we typically assume you mean you will 'only' study 'one of these' subjects and 'not both' however we use similar phrases or wording for when we mean to include both options.

Mathematical Logic is sort of similar to high school algebra, you assign letter to well-defined sentences just like variable and you can sort of 'add', 'subtract' and do other operations on them. Although this description is very simplifed, it helps to give an idea of what you need to know to be able to learn the skills of dealing with mathematical logic.

2.2.1 Proposition

In **Propositional logic** (also called Sentential logic, or Propositional Calculus), we discuss about **propositions** and how they connect using logical connectives.

Definition 2.2. A **statement** is a sentence which can either be true or false. There is no other possiblity and it cannot be both.

A statement can be an ordinary sentence or a mathematical sentence.

Example 2.3. Examples of what a proposition *is*,

1. It is raining.
2. If 5 is not divisible by 2, then 5 is not an even number
3. $2 + 7 = 3$

You might have noticed that in the last example, $2 + 7 = 3$ is a false statement, that means it is not necessary for a sentence to be true as to be considered a propotision.

Now, let's see some statements which are not a proposition,

Example 2.4. Examples of what a proposition *is not*

1. Will it rain today?
2. Munich is 781 km away from Hamburg.
3. 10 is not divisible by x^2
4. $2 + xy = 3y$

What is the difference? The difference is that we can confidently assign a **truth value** to the sentences in Example 2.3, T for *true*, and F for *false*[6].

In simple words, you can confidently say whether $3 > 2$ is true or not (in this case it is, so it's truth value is T), but we can't say whether $3+x > 5$ is *always* true or false, you need to know what x is in order to decide it's truth value. This distinction is very important to keep in mind. Statements like those in Example 1.1.2. are called **expressions**.

In the second sentence, it is not clear what is meant by "Munich" and "Hamburg" and with what tolerance should the distance be measured?

To deal with complicated statements in mathematics, it is convenient to use the notation of symbols. We will use capital letter to denote sentences such as, $A :=$ "3 is greater than 2", the symbol ":=" stands for "is defined by".

[6]in computer science and information science, 1 and 0 are used for true and false respectively

2.2.2 Logical Connectives

From the primitive propositions, we can make compound propositions by using **logical connectives** on one or more propositions. The truth value of the new proposition formed is determined by the truth value of the primitive ones.

We can define the negation of A as:

Definition 2.5 (Negation)**.** $\neg A$ is the **negation** of A, $\neg A$ is true when A is false and $\neg A$ is false when A is true. This can be represented in a **truth table**:

A	$\neg A$
T	F
F	T

Two statements, A and B, can be combined using **conjunction** $\wedge$: $A \wedge B$ is true if both A and B are true and false in all other cases. And **disjunction** $\vee$: $A \vee B$ is false if both A and B are false and true in all other cases.

The disjunction is sometimes said to mean the *inclusive or*, it is true when either A or B or both A and B are true and not 'exclusive or'[7] which would be false if both are true, and true when *only* one of them is true.

The next type of connective is the **implication**: $A \Rightarrow B$ is false if A is true and B is false (because A *implies* B) and true in all other cases. It is defined as

$$A \Rightarrow B := (\neg A) \vee B$$

It is common for the implication to be written in English as: "A *implies* B", "If A then B", "B is necessary for A", or "A is sufficient for B". Implication is the most common type of statement in mathematics.

Its **converse** is defined as: $A \Leftarrow B \coloneqq B \Rightarrow A$.

The **equivalence** is the conjunction of the implication and its converse:

$$A \Leftrightarrow B \coloneqq (A \Rightarrow B) \wedge (B \Rightarrow A)$$

A common way to say it is "A if and only if B" or in short "A iff B". Some other ways include: A is a necessary and sufficient condition for B (or vice versa).

All the logical connectives can be represented on the truth table which makes their definitions clear:

A	B	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftarrow B$	$A \Leftrightarrow B$
T	T	T	T	T	T	T
T	F	F	T	F	T	F
F	T	F	T	T	F	F
F	F	F	F	T	T	T

Now, let's take some examples to understand how symbols can help us to analyse their logical forms:

Example 2.6. Analyse the logical form of the following statements:

[7] Exclusive or: $A \oplus B := (A \vee B) \wedge \neg(A \wedge B)$. You can check using truth table that it is true only when one of the two is true.

1. The election result is not decided even though the votes have been counted.

2. Either Thomas and Anja are both telling the truth, or neither of them is.

3. You will not pass the course if you either miss the exam or do not do every exercise sheet.

4. 4 is less than 5 if and only if, the sum of 3 and 7 is not 10 and 6 is a multiple of 2.

Solution. Try the problems yourself before looking at the solutions. Following the logic of the statements, we'll translate them into symbolical logic step-by-step.

1. Let $E \coloneqq$ "The election result is decided". Let $V \coloneqq$ "The votes have been counted". Then, the negation of E, "The election result is not decided" is $\neg E$. Now, "even though" is just another way of saying "but" which is the conjunction, so the final statement is $\neg E \wedge V$.

2. Let $T \coloneqq$ "Thomas is telling the truth". Let $A \coloneqq$ "Anja is telling the truth". "Thomas and Anja are both telling the truth" can be written as: $T \wedge A$. Similarly, "neither of them (Thomas and Anja) is (telling the truth)" can be written as: $(\neg T) \wedge (\neg A)$. So, we have: Either $T \wedge A$, or $(\neg T) \wedge (\neg A)$. This is a disjunction which can be written as: $(T \wedge A) \vee [(\neg T) \wedge (\neg A)]$.

3. Let $P \coloneqq$ "You will pass the course". Let $M \coloneqq$ "You miss the exam". Let $E \coloneqq$ "You do every exercise sheet". This is an implication where "You miss the exam or (you) do not do every exercise sheet" is the premise $M \vee \neg E$ and the conclusion "You will not pass the course" is $\neg P$. So, the final statement is $[M \vee (\neg E)] \Rightarrow \neg P$.

4. Let $L \coloneqq 4 < 5$. Let $S \coloneqq (3 + 7 = 10)$. Let $M \coloneqq$ "6 is a multiple of 2". By now, it you might have figured how to translate. This is an equivalence statement, so $L \Leftrightarrow [(\neg S) \wedge M]$. This statement is always false but here we are not talking about the truth value of the statements, we just want to practice writing them in formal logic.

□

Converting ordinary statements into formal logical symbols is not always straightforward, some assumptions are implicit. For example, in example 4, we could have assumed that the statement says "4 is less than 5 if and only if, the sum of 3 and 7 is not 10 *and* 6 is a multiple of 2". Here, we would have written $[L \Leftrightarrow (\neg S)] \wedge M$ but the comma after "if and only if" tells us implicitely that both $\neg S$ *and* M are to be taken inside the conclusion of the implication. You might have also notice that nuances of the language is lost in favour for formal precision.

Statements that are always true are called **tautology** (represented by $\top$) and the ones that are always false are called **contradictions** (represented by $\bot$).

2.2.3 Precedence of Logical Operators

We have been using brackets and parentheses to specify the order in which the connectives apply. However, for highly complicated statements this can get very tedious. So, to reduce that we specify the order of precedence of connectives:

Connective	Precedence
$\neg$	1
$\wedge$	2
$\vee$	3
$\Rightarrow$	4
$\Leftrightarrow$	5

Remark 2.7. Precedence is "transitive" meaning that if $\neg$ precedes $\wedge$, and $\wedge$ precedes $\vee$, then $\neg$ precedes $\vee$ as well. This might seem obvious but transitivity is not always guranteed for every property[8].

Just like the rule of precedence in arithmetic, our rule works sort of the same way. In the arithmetic expression, $1 + 4 \times 2$, you multiply before adding, in the same way for the formula $A \wedge \neg B \Rightarrow C$ you first consider the negation $\neg B$ then the conjunction $A \wedge \neg B$ and then the implication.

By using these rules we can rewrite the statements of Example 2.6 in a much cleaner way: In the second statement, removing parentheses over the negation, $(T \wedge A) \vee (\neg T \wedge \neg A)$. We can remove the rest of the parentheses as well since $\wedge$ precedes $\vee$: $T \wedge A \vee \neg T \wedge \neg A$. Similarly, we can write $M \vee \neg E \Rightarrow \neg P$ for the third statement, and $L \Leftrightarrow \neg S \wedge M$ for the fourth.

2.2.4 Propositional Equivalence

You may have already noticed that some statements are equivalent to some other statement and it's often convenient to write statements in some other way to do Propositional calculus. Below I'll present key implication and equivalence laws, they can be verified using truth tables.

Theorem 2.8. *For statements A, B and C. We have the following implication and equivalence laws:*

- **Double Negation Law:** $\neg\neg A \Leftrightarrow A$.
- **verum ex quodlibet:** $A \Rightarrow \top$.
- **ex falso quodlibet:** $\bot \Rightarrow A$.
- **Indempotent Laws:** $A \wedge A \Leftrightarrow A$ *and* $A \vee A \Leftrightarrow A$.
- **Neutrality Laws:** $A \wedge \top \Leftrightarrow A$ *and* $A \vee \top \Leftrightarrow A$.
- **Absorption Laws:** $A \wedge \bot \Leftrightarrow \bot$ *and* $A \vee \top \Leftrightarrow \top$.
- **Complementarity Laws:** $A \wedge \neg A \Leftrightarrow \bot$ *and* $A \vee \neg A \Leftrightarrow \top$.
- **Commutativity Law:** $A \wedge B \Leftrightarrow B \wedge A$.
- **Commutativity Law:** $A \vee B \Leftrightarrow B \vee A$.
- **Associativity Law:** $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$.

[8]Some mathematicians treat $\vee$ and $\wedge$ as having the same precedence and some treat the precedence of $\Rightarrow$ and $\Leftrightarrow$ as equivalent, but we'll follow a strict order to avoid any confusion.

- **Associativity Law:** $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$.
- **De Morgan's Law:** $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$.
- **De Morgan's Law:** $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$.
- **Distributivity Law:** $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$.
- **Distributivity Law:** $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$.

From the truth table, you can verify that

$$A \Rightarrow B \Leftrightarrow \neg B \Rightarrow \neg A.$$

This is called the **contrapositive** of the implication. This statement is very useful for proving implications.

To say that a statement A is true whenever B is true, we write:

$$A :\Leftrightarrow B.$$

This stands for "A is true, by definition, if B is true."

2.2.5 Logical Puzzles!!

Formal logic is not just some dry symbolic formalisation for mathematics, we can use it to solve some interesting logical puzzles as well :D!

All puzzles can't really be solved by logical reasoning, there are some types of puzzles that you only solve. But they still very fun and interesting to solve. The way to solve these puzzles is just by coverting the statements into formal logic and then applying logical deduction on them. Let's solve an interesting puzzle posed by Raymond Smullyan:

Example 2.9. You are on an island that has two kinds of inhabitants, *knights*, who always tell the truth, and their opposites, *knaves*, who always lie. You encounter two people, A and B. Determine, what A and B are if they address you in the ways described:

1. A says "The two of us are both knights" and B says "A is a knave."
2. A says "I am a knave or B is a knight" and B says nothing.

Solution. Let $P :=$ "A is a knight" and $Q :=$ "B is a knight". Then $\neg P$ and $\neg Q$ would be the statements, "A is a knave" and "B is a knave" respectively.

1. Let's first consider that A is a knight meaning P is true. In this case A must be telling the truth, so $P \wedge Q$ must be true. But if Q is true then B is also a knight and his statements $\neg P$ should also be true, which is not the case. Hence, we conclude that A is not a knight. If A is a knave, i.e. $\neg P$ is true, then $P \wedge Q$ must be false, which is equivalent to $\neg P \vee \neg Q$. Since P is already false, Q may or may not be true. If we consider that Q is true then B's statement $\neg P$ must be true, which it is. Hence, we conclude that $\neg P \wedge Q$ is true meaning A is a knave and B is a knight.
2. Let's first consider that A is a knight, so P is true. Then $\neg P \vee Q$ must be true. Since P is true then Q must be true, meaning B must be a knight. If A was a knave, $\neg P$ wouls be true and $\neg P \vee Q$ would be false meaning $P \wedge \neg Q$ must be true, but P is false, so this cannot be the case. We conclude that $P \wedge Q$ is true, A and B are both knights.

□

These puzzles were not very challenging, so now we'll add a third type of inhabitants *normals*, which may either lie or tell the truth. This increases the difficulty and makes the puzzle more fun!

Example 2.10. You are on an island that has three kinds of inhabitants, *knights*, who always tell the truth, *knaves*, who always lie, and *normals*, who may lie or tell the truth. You encounter three people, A, B and C. Determine, what these three are if they address you in the ways described:

1. A says "C is a knave,", B says "A is a knight," and C says "I am the normal."

2. A says "I am a knave or B is a knight" and B says nothing.

Solution. Let $P \coloneqq$ "A is a knight" and $Q \coloneqq$ "B is a knight". Then $\neg P$ and $\neg Q$ would be the statements, "A is a knave" and "B is a knave" respectively.

1. Let's first consider that A is a knight meaning P is true. In this case A must be telling the truth, so $P \wedge Q$ must be true. But if Q is true then B is also a knight and his statements $\neg P$ should also be true, which is not the case. Hence, we conclude that A is not a knight. If A is a knave, i.e. $\neg P$ is true, then $P \wedge Q$ must be false, which is equivalent to $\neg P \vee \neg Q$. Since P is already false, Q may or may not be true. If we consider that Q is true then B's statement $\neg P$ must be true, which it is. Hence, we conclude that $\neg P \wedge Q$ is true meaning A is a knave and B is a knight.

2. Let's first consider that A is a knight, so P is true. Then $\neg P \vee Q$ must be true. Since P is true then Q must be true, meaning B must be a knight. If A was a knave, $\neg P$ wouls be true and $\neg P \vee Q$ would be false meaning $P \wedge \neg Q$ must be true, but P is false, so this cannot be the case. We conclude that $P \wedge Q$ is true, A and B are both knights.

□

2.3 Predicate Logic

2.3.1 Variables and Quantifiers

Propositional logic is powerful, but it is not sufficient to do represent all kinds of mathematical statements. To represent statements like, "If n is an even integer, then n^2 is also an even interger", caould be written in propositional logic but it wouldn't really bring out the true essense of the statements which is why we need statements that can take in variables.

For this, we need predicate logic. In predicate logic, we can separate the statement from the object it describes. It is an extension of propositional logic, we just take variable into account as well.

A typical statement in predicate logic is represented as $E(x)$ which says "x has property E," replacing x which an object would change this a simple proposition.

Example 2.11. Some examples of predicate statements:

1. $A(x) \coloneqq x$ is a mean cat.
2. $B(x) \coloneqq x$ is an odd number.
3. $C(x, y) \coloneqq x$ thinks y is dumb.

The truth value of these statements depends on the value of the **free variables** (x in the 1 and 2, x and y in 3), variable which are placeholders and can be replaced by any other variable without changing the meaning of the expressions, for example $A(y)$ and the original statement mean the same thing. The number os variables in a statement is called its **arity**, 1 in for $A(x)$ and $B(x)$ and 2 for $C(x, y)$.

Of course we have not defined what type of object x can be and to which class it belongs to. But this is a general statement which applies to everything. For example, you could replcae x with "Juli" and form a statement: "Juli is a mean cat", even though Juli might be mean but she may not be a cat if she does not belong to that class of objects.

We call the class which the objects belong to as the **universe of discourse** U, for now, to avoid complications we'll assumes that U is not empty.

On of the most important features of predicate logic is that we can now quantify statements, since we can talk about some objects in a class, we can write statements about whether the statement is about all objects in that class or some.

To quantify the statement for *all* objects in a class we use the **universal quantifier** $\forall$, pronounced as "for all".

To quantify the statement for *at least one* object in a class, we use the **existential quantifier** $\exists$. The **uniqueness quantifier**, $\exists!$ is used to specify that *only one* object has the property of the statement.

Example 2.12. Some examples of quantified statements:

1. $\forall n \in \mathbb{N}(n \geq 0)$: "All natural numbers are greater than or equal to 0." (true)
2. $\forall n, m \in \mathbb{N}(n \geq m)$: "Every natural numbers has another natural number greater than or equal to itself." (true)
3. $\exists n \in \mathbb{N}(n < 0)$: "There is at least one natural numbers less than zero." (false)
4. $\exists x \in \mathbb{R}(x \leq 0)$: "THere is at least one real numbers less than zro." (true)

Let's see some examples of using quantifiers over predicates:

Example 2.13. In the universe of discourse $U \coloneqq \mathbb{N}$, write the following statements in symbolic form:

1. There exists at least one even prime natural number.
2. Every natural numbers is either even or not a prime numbers.
3. Not every prime natural number is odd.

Solution. Let $E(x) \coloneqq$ "x is an even number", and $P(x) \coloneqq$ "x is a prime number". Then we can the above statements as:

1. $\exists x \in \mathbb{N}(E(x) \wedge P(x))$.

2. $\forall x \in \mathbb{N}(E(x) \vee \neg P(x))$. Another way to write this is by using the equivalence laws $\forall x \in \mathbb{N}(P(x) \Rightarrow E(x))$.

3. This can be written as $\neg$(Every prime natural number is odd) which is an implication, $\neg$(If a natural number is prime then it is $\neg$ even), which can be finally written as $\neg\forall x \in \mathbb{N}(P(x) \Rightarrow \neg E(x))$.

Statement 2 is obviously false because we can find certain values of x for which the statement does not hold true, such as 3, 5 and 7. If a universal quantifier does not hold true for even a single value of x, then it does not hold for all. □

You may notice that statement 3 can also be written as "There exists at least one natural number which is prime and even." In symbols, $\exists x \in \mathbb{N}(P(x) \wedge E(x))$. Which is same as statement 1. This is not a coincidence, there is a general rule for negating quantified statements:

Theorem 2.14. *Let A and B be predicates of arity one, and C be a predicate of arity two, then we have the following implications and equivalences for quantified statements:*

1. $\neg(\forall x A(x)) \Leftrightarrow \exists x(\neg A(x))$ *(Negation of universal quantifier).*
2. $\neg(\exists x A(x)) \Leftrightarrow \forall(\neg A(x))$ *(Negation of existential quantifier).*
3. $\forall x \forall y C(x, y) \Leftrightarrow \forall y \forall x C(x, y)$.
4. $\exists x \exists y C(x, y) \Leftrightarrow \exists y \exists x C(x, y)$.
5. $\forall x(A(x) \wedge B(x)) \Leftrightarrow \forall x A(x) \wedge \forall x B(x)$.
6. $\exists x(A(x) \vee B(x)) \Leftrightarrow \exists x A(x) \vee \exists x B(x)$.
7. $(\forall x A(x)) \vee (\forall x B(x)) \Rightarrow \forall x(A(x) \vee B(x))$.
8. $\exists x(A(x) \wedge B(x)) \Rightarrow \exists x A(x) \wedge \exists x B(x)$.
9. $\forall x(A(x) \Rightarrow B(x)) \Rightarrow ((\forall x A(x)) \Rightarrow (\forall x B(x)))$.
10. $\exists x(A(x) \Rightarrow B(x)) \Leftrightarrow ((\exists x A(x)) \Rightarrow (\exists x B(x)))$.

3 Proofs

Now that we have discussed the foundations of logic and sets, we're in a good position to do proofs. Here, you'll learn basic proof stratergies and techniques such as direct proofs, contraposition, contradiction and existence and uniqueness proofs. The final section will give an introduction to limits of formal systems and proofs.

3.1 Formalisation and Formal Proofs

3.2 Inference Rules

Below we'll see some more important tautologies which are very useful for proofs.

Theorem 3.1 (modus ponendo ponens)**.** *Let A and B be two statements, then the following statement holds:*

$$A \wedge (A \Rightarrow B) \Rightarrow B$$

.

Proof. Proof can be done via the truth table. □

If we are given, or we can prove A and $A \Rightarrow B$ then by *modus ponendo ponens* we can conclude B.

Theorem 3.2 (modus tollendo tollens)**.** *Let A and B be two statements, then the following statement holds:*

$$\neg B \wedge (A \Rightarrow B) \Rightarrow \neg A$$

.

Proof. Proof can be done via the truth table. □

Since, $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$, from $\neg B$ and $A \Rightarrow B$, by *modus tollendo tollens* we can conclude $\neg A$.

Theorem 3.3 (modus ponendo tollens)**.** *Let A and B be two statements, then the following statement holds:*

$$A \wedge (A \Rightarrow \neg B) \Rightarrow \neg B$$

.

Proof. Proof can be done via the truth table. □

Theorem 3.4 (modus tollendo ponens)**.** *Let A and B be two statements, then the following statement holds:*

$$\neg A \wedge (\neg A \Rightarrow B) \Rightarrow B$$

.

Proof. Proof can be done via the truth table. □

An important transitive law:

Theorem 3.5 (Chain Inference)**.** *Let A, B and C be three statements, then the following statement holds:*

$$(A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C)$$

.

Proof. Proof can be done via the truth table. □

Now, let's see some uses of these inference rules.

Example 3.6. Suppose $P \Rightarrow Q$ and $R \Rightarrow \neg Q$ are both true. Prove that $P \Rightarrow \neg R$ is true.

Solution. **Given:** $P \Rightarrow Q$, $R \Rightarrow \neg Q$. **To prove:** $P \Rightarrow \neg R$.
Assume P and make $\neg R$ the goal.
Given: $P \Rightarrow Q$, $R \Rightarrow \neg Q$, P. **To prove:** $\neg R$.
We get $P \wedge (P \Rightarrow Q) \Rightarrow Q$ by *modus ponendo ponens*.
Given: $P \Rightarrow Q$, $R \Rightarrow \neg Q$, P, Q. **To prove:** $\neg R$.
We get $Q \wedge (R \Rightarrow \neg Q) \Rightarrow \neg R$ by *modus ponendo tollens*. □

Lemma 3.7. *Suppose $P \Rightarrow Q$ and $R \Rightarrow \neg Q$. Then $P \Rightarrow \neg R$.*

Proof. Suppose P. Then $P \wedge (P \Rightarrow Q) \Rightarrow Q$ by *modus ponendo ponens*. Then $Q \wedge (R \Rightarrow \neg Q) \Rightarrow \neg R$ by *modus ponendo tollens*. □

Sometimes you can use the inference rules to work backwards. If one of your given has the form $A \Rightarrow B$ and the goal is B, if you could prove A then you could conclude B, so you can change your goal to A and see if it makes the problem easier.

Let us now see two predicate logic inference rules for handling quantifiers.

1. Removing the $\forall$-quantifier from the assertion:

1. To prove: $\forall x$ of Type T : $B(x)$ given A.
2. Let x of Type T be given.
3. Now, to prove: $B(x)$ given A, x of Type T.

If x is already used in the assumption A, then you must *rename* it to a new unused variable such as, $\tilde{x}$ and replace the goal $B(x)$ with $B(\tilde{x})$. This avoids **variable capture**.

Essentially, we are just choosing an *arbitrary* x of a particular type and then proving the goal. If $B(x)$ is true for an arbitrary x, it must be true for all x, since we did not assume anything else about x.

2. Removing the $\exists$-quantifier from the assertion:

1. To prove: $\exists x$ of Type T : $B(x)$ given A.
2. Choose $x =$ Term. Here, the Term must be of Type T.
3. Now, to prove: $B(x)$ given A, $x =$ Term.

Again, you must rename x, if it occurs freely in A!

Sine, we assinged $x =$ Term, the goal can be formulated alternatively as:

To prove: B(Term) given A.

We usually use our intuition or side calculation to carefully choose the "Term" before constructing the proof. Choosing a Term is just like choosing a concrete instance.

Now we will discuss two more predicate logic inference rules which will allow us to use a given universal and existential statements.

1. Removing the $\forall$-quantifier from a given statement:

1. To prove: Φ. Given: A, $\forall x$ of Type T : $B(x)$.

2. We apply the given universal statement to $x =$ Term. Here, the Term must be of Type T.

3. Still, to prove: Φ. Given: A, B(Term), $\forall x$ of Type $T : B(x)$.

In this inference, B(Term) is added as a newly given statement. The statement to be proven, Φ, remains unchanged. The universal statement $\forall x$ of Type $T : B(x)$ also remains given; it can be applied to other terms when needed.

2. Removing the $\exists$-quantifier from a given statement:

1. To prove: Φ. Given: A, $\exists x$ of Type $T : B(x)$.

2. We take such an x of Type T and call it y. Here, y must be a new variable not occuring freely.

3. Still, to prove: Φ. Given: A, y of Type T, $B(y)$.

Again, the statement to be proven, Φ, remains unchanged. The existential statement $\exists x$ of Type $T : B(x)$ also remains given, but it no longer useful to list it in the list of givens. It contains no new information that $B(y)$ already does not have with the new free variable y of Type T.

3.3 Proof Techniques

In this section we will mainly focus on leaning proof techniques which we use to prove mathematical statements while also talking about why they work.

There isn't any strict rule for how proofs must be written and how much formality there should be. When you start writing your own proofs, you'll realise they are very difficult, and unfortunately, there aren't any step-by-step procedure to do proofs either. But as with every difficult problem, we can break proof problems into simpler problems by applying the tautologies we have studied.

The most common type of statements in mathematics is an implication and then equivalences, and they are usually open statements. To prove an implication, we have to show that $A \Rightarrow B$ or $A(x) \Rightarrow B(x)$ is a tautology for all x in the *universe of discourse*. Usually the premise itself is more complex and consists of many statements joined by connectives.

We will discuss different techniques we can use to prove an implication, some of these techniques can be used for statements of other form as well with some modification. First we will study some proofs that aren't common in mathematics but are important reminder about the truth table of implications.

3.3.1 Trivial and Vacuous Proofs

We know that the statement $A(x) \Rightarrow B(x)$ is true if $B(x)$ is true for all x regardless of whether $A(x)$ is true or not. So, if we can prove $B(x)$ for all x then the implication is true and thus the proof is complete. Such a proof is called a **trivial proof**.

However, it might not be clear whether we can do a trivial proof and most of the time we would have to use our intuition.

Example 3.8. Let $x \in \mathbb{Z}$. Prove that if $x + x^2 + |x - 5| \leq 5$, then $|x^2 - 1| \geq 0$.

Solution. **Given:** $x \in \mathbb{Z}$. **To Prove:** $x + x^2 + |x - 5| \leq 0 \Rightarrow |x^2 + 1| \geq 0$.

We intuitively know that showing $|x^2 - 1| \geq 0$ directly is easy without even using the premise. So, we just need to show that the conclusion is true for an arbitrary $x \in \mathbb{Z}$.

Given: $x \in \mathbb{Z}$. **To Prove:** $|x^2 + 1| \geq 0$.

Since, $x^2 \geq 0$ for all $x \in \mathbb{Z}$. Adding -1 from both sides gives, $x^2 - 1 \geq -1$. Then taking the absolute value, $|x^2 - 1| \geq 0$.

Since the conclusion is true for all x, the implication is true regardless of the premise. □

Lemma 3.9. *For all $x \in \mathbb{Z}$, if $|x + x^2| + |x - 5| \leq 0$, then $|x^2 - 1| \geq 0$.*

Proof. Since, $\forall x \in \mathbb{Z}(x^2 \geq 0)$. Then, $x^2 - 1 \geq -1$. By taking the absolute value on both sides, $|x^2 - 1| \geq 0$. Hence, if $|x + x^2| + |x - 5| \leq 0$, then $|x^2 - 1| \geq 0$. □

In the above example, we didn't care about the premise, $x + x^2 + |x - 5| \leq 5$, we just focused on the conclusion. We can often say that the proof of the above lemma follows **trivially**.

The implication $A(x) \Rightarrow B(x)$ is also true if $A(x)$ is false for all x regardless of whether $B(x)$ is true or not. This allows us to prove the implication by showing the premise is false, such a proof is called a **vacuous proof**.

Example 3.10. Let $n \in \mathbb{N}$. Prove that if $|n - 1| + |n + 1| \leq 1$, then $|n^2 - 1| \geq 4$.

Solution. **Given:** $n \in \mathbb{N}$. **To Prove:** $|n - 1| + |n + 1| \leq 1 \Rightarrow |n^2 - 1| \geq 4$.

We intuitively know that $|n^2 - 1| \geq 4$ is false. So, we just need to show that the premise is false, i.e. prove that $|n - 1| + |n + 1| > 1$ for all $n \in \mathbb{N}$.

Given: $n \in \mathbb{N}$. **To Prove:** $|n - 1| + |n + 1| > 1$.

Since, $n \geq 0$ for all $n \in \mathbb{N}$. Adding 1 on both sides we get, $n + 1 \geq 1$, then taking the absolute value on both sides gives, $|n + 1| \geq 1$. Similarily, adding -1 and taking absolute value on both sides of $n \geq 0$ gives, $|n - 1| \geq 0$. Now, adding the positive number (or zero) $|n - 1|$ to the inequality $|n + 1| \geq 1$ gives, $|n + 1| + |n - 1| > 1$. We can remove the equality since, $|n + 1| + |n - 1| = 1$ is not possible for 0, 1 or any $n \in \mathbb{N}$.

Since the premise is false for all x, the implication is true regardless of the conclusion. □

Lemma 3.11. *For all $n \in \mathbb{N}$, if $|n - 1| + |n + 1| \leq 1$, then $|n^2 - 1| \geq 4$.*

Proof. Since, $\forall n \in \mathbb{N}(n \geq 0)$. Adding 1 and taking the absolute value on both sides gives, $|n + 1| \geq 1$. Similarily, adding -1 and taking absolute value on both sides gives, $|n - 1| \geq 0$. Adding the two inequality gives, $|n + 1| + |n - 1| > 1$. Removing the equality since, $|n + 1| + |n - 1| = 1$ is not possible for any $n \in \mathbb{N}$. Hence, if $|n - 1| + |n + 1| \leq 1$, then $|n^2 - 1| \geq 4$. □

3.3.2 Direct Proof

Now, we will introduce the first proof technique.

To prove a goal of the form $A \Rightarrow B$: *Assume A is true, then show that B must logically follow from A and hence the implication is true.*

This method is called **direct proof**, in this method we *directly* derive the conclusion from axioms and previously proven theorems using the rules of inference.

This might not seem to be much of a help but an example will make it concrete. We will first do an informal scratch work before writing the formal proof.

Example 3.12. Suppose a and b are two positive real numbers. Prove that if $a > b$ then $\frac{1}{a} < \frac{1}{b}$.

Solution. Given: $a, b \in \mathbb{R}^+$. To prove: $(a > b) \Rightarrow (\frac{1}{a} < \frac{1}{b})$.

The goal is of the form $A \Rightarrow B$ so, we apply our strategy, assume the premise is true and add it to the list of given.

Given: $a, b \in \mathbb{R}^+$ and $a > b$. To prove: $\frac{1}{a} < \frac{1}{b}$.

We don't have any more proof strategies to use to simplify our problem futher, so we have work with this from here. We start from the statement $a > b$, multiply this by $\frac{1}{ab} > 0$ on both sides and since multiplying by positive real numbers preserves inequality, we get the required conclusion: $\frac{1}{b} > \frac{1}{a} \Leftrightarrow \frac{1}{a} < \frac{1}{b}$. □

Below I'll provide a formal way to write a theorem and it's proof.

Theorem 3.13. *Suppose a and b are two positive real numbers. If $a > b$ then $\frac{1}{a} < \frac{1}{b}$.*

Proof. Suppose $a > b$. Multiplying the inequality by the positive real number $\frac{1}{ab}$ on both sides, we get $\frac{1}{a} < \frac{1}{b}$. Hence, $(a > b) \Rightarrow (\frac{1}{a} < \frac{1}{b})$. □

This proof is quite concise and leaves out some assumptions implicit, such as th fact that multiplying by a positive number preserves the inequality. But those details are assumed to be filled by the reader.

The formality of a proof often depends on the field of mathematics and the intended audience. In a course of analysis, you would typically write concise proofs like this, but in a course of mathematical logic, you would be invoking every god damn axiom and inference rule in existence and write multiple pages of a long proof for something trivial.

Every mathematician would write proofs in somewhat different styles, some may not use logical symbols (especially in Anglo-American tradition), some may use more words. All of that is not a big deal as long as the logic is clear and the proof is valid.

Usually after completing a proof, we make a box □ at the end to indicate that the proof is complete, or like in the older days, write "Q.E.D." (quod erat demonstrandum) which means "which was to be demonstrated" in English.

Note that the proof is about the implication $A \Rightarrow B$ and not the conclusion B itself. Our proof does not shows that B is true, rather it shows that B follows *if* A is true.

3.3.3 Proof by Contraposition

Another way to prove an implication is to prove its contrapositive $\neg B \Rightarrow \neg A$. Such a proof is called "Proof by Contraposition" or **indirect proof**, instead of proving the implication we do a *direct proof* of its contrapositive which *indirectly* shows that the implication is true.

Example 3.14. Suppose $A \backslash B \subset C \cap D$ and $x \in A$. Prove that if $x \notin D$ then $x \in B$.

Solution. Given: $A \backslash B \subset C \cap D$ and $x \in A$. To prove: $(x \notin D) \Rightarrow (x \in B)$.

We'll prove the contrapositive, $(x \notin B) \Rightarrow (x \in D)$. Assume $x \notin B$.

Given: $A \backslash B \subset C \cap D$, $x \in A$ and $x \notin B$. To prove: $x \in D$.

If $x \in A$ and $x \notin B$ is true, then $A \backslash B$ is true, and from $A \backslash B \subset C \cap D$, it follows $x \in C \cap D$ which implies that $x \in D$. □

Theorem 3.15. *Suppose $A \backslash B \subset C \cap D$ and $x \in A$. If $x \notin D$ then $x \in B$.*

Proof. We will prove the contrapositive. Suppose $x \notin B$. From $x \in A$ and $x \notin B$, it follows $A \backslash B$, thus $x \in C \cap D$ which implies $x \in D$. Hence, $(x \notin D) \Rightarrow (x \in B)$. □

Let's write some more proofs to get better practice.

Example 3.16. Suppose x and y are real numbers. Prove that if $x^2 + y = -3$ and $2x - y = 2$ then $x = -1$.

Solution. Suppose $x^2 + y = -3$ and $2x - y = 2$. Adding the two equation gives $x^2 + 2x + 1 = 0$ which implies $(x + 1)^2 = 0$, and thus $x = -1$.

Hence, $x^2 + y = -3 \wedge 2x - y = 2 \Rightarrow x = -1$. □

3.3.4 Proof by Contradiction

Sometimes we can use the equivalence $(A \Rightarrow B) \Leftrightarrow (A \wedge \neg B) \Rightarrow \bot$. This means that to prove $A \Rightarrow B$, we can assume A is true and B is false, and then derive a contradiction. This method is called **proof by contradiction**.

Proof by contradiction is a very useful strategy, it can be used for goals of any logical form, it is especially useful when the conclusion B is a negative statement of the form $\neg C$. In such cases, we assume A is true and C is true, and then derive a contradiction. It is very similar to proof by contraposition (assume $\neg B$ and deriving $\neg A$), but here we show that assuming B is false and A is true leads to a contradiction.

In Section 3.3.6 we will see how to this strategy it to prove negations. But for now let's take an example of an implication.

Example 3.17. Suppose x is a real number. Prove that $\sin(x) + \cos(x) \neq \frac{5}{3}$

Solution. To prove: $x \in \mathbb{R} \Rightarrow \sin(x) + \cos(x) \neq \frac{5}{3}$.

1. We prove by contradiction, To prove: $x \in \mathbb{R} \wedge \sin(x) + \cos(x) = \frac{5}{3} \Rightarrow \bot$.

2. This is an implication so we can use our previous technique and assume $x \in \mathbb{R}$ and $\sin(x) + \cos(x) = \frac{3}{2}$. Given: $x \in \mathbb{R}$ and $\sin(x) + \cos(x) = \frac{5}{3}$. To prove: $\bot$.

3. To show a contradiction, we will assume that a certain $x_0 \in \mathbb{R}$ exists with the property $\sin(x_0) + \cos(x_0) = \frac{5}{3}$. Since our claim should be true for any real numbers, just one counterexample is enough to show a contradiction.

4. Squaring both sides of this equation gives: $\sin^2(x_0) + \cos^2(x_0) + 2\sin(x_0)\cos(x_0) = \frac{25}{9}$.

5. Using trigonometric identities, $\sin^2(x) + \cos^2(x) = 1$ and $2\sin(x)\cos(x) = \sin(2x)$ for all $x \in \mathbb{R}$ (should be true for x_0 as well), we get $1 + \sin(2x_0) = \frac{25}{9} \Rightarrow \sin(2x_0) = \frac{16}{9} > 1$.

6. No value of x_0 can satify this equation since sine is defined only on -1 to 1. Hence our assumption that $x \in \mathbb{R}$ and $\sin(x) + \cos(x) = \frac{5}{3}$ is false.

Therefore, $x \in \mathbb{R} \Rightarrow \sin(x) + \cos(x) \neq \frac{5}{3}$. □

Now, let's take an example where writing out the definition of one of the givens is useful.

Example 3.18. Suppose $A \backslash B$ is disjoint from C and $x \in A$. Prove that if $x \in C$ then $x \in B$.

Solution. **Given:** $A \backslash B \cap C = \emptyset$, $x \in A$. **To Prove:** $x \in C \Rightarrow x \in B$.

We prove by contradiction and add $x \in C$ and $x \notin B$ to the list of givens and make $\perp$ as the goal.

So, we examine the givens. First, writing out the definition of the first given.

$$\forall x(x \in A \wedge x \notin B \Rightarrow x \notin C).$$

So, if $x \in A$ and $x \notin B$ are true, then $x \in C$ cannot be true. Therefore our assumption $x \notin B$ was wrong as it cases a contradiction and the proof is complete.

□

Theorem 3.19. *Suppose that $A \backslash B$ is disjoint from C and $x \in A$. If $x \in C$ then $x \in B$.*

Proof. Suppose $x \in C$. Suppose $x \notin B$. Since, $A \backslash B$ and C are disjoint, then $x \in A$ and $x \notin B$ implies $x \notin C$. This contradicts the fact that $x \in C$. Therefore, $x \in B$. Hence, $x \in C \Rightarrow x \in B$. □

3.3.5 Proof by Distinction of Cases

Sometimes we can use an additional statement C to prove an implication by distinct cases, using the equivalence $(A \Rightarrow B) \Leftrightarrow (A \wedge C \Rightarrow B) \wedge (A \wedge \neg C \Rightarrow B)$.

Using the rule for a conjuction, we first show $A \wedge C \Rightarrow B$, assuming that A and C are true then showing B is also, then proving $A \wedge \neg C \Rightarrow B$, by assume now that A is tru but C is false then B is true again. This proves that B is true in both cases when A is true, hence $A \Rightarrow B$ must be true.

It might not always be clear what C we should assume,

Example 3.20. Suppose x is an integer. Prove that $x^2 + x$ is an even integer.

Solution. Let $E(x) \coloneqq$ "x is an even integer." **To Prove:** $x \in \mathbb{Z} \Rightarrow E(x^2 + x)$.

1. Any integer is either even or odd, so we can use the equivalence,

   $[x \in \mathbb{Z} \Rightarrow E(x^2+x)] \Leftrightarrow (x \in \mathbb{Z} \wedge E(x) \Rightarrow E(x^2+x)) \wedge (x \in \mathbb{Z} \wedge \neg E(x) \Rightarrow E(x^2+x))$.

2. **Case 1:** $E(x) : x \in \mathbb{Z} \wedge E(x) \Rightarrow E(x^2+x)$. This is an implication, so assume $x \in \mathbb{Z}$ and $E(x)$.

   Given: $x \in \mathbb{Z} \wedge E(x)$. **To Prove:** $E(x^2+x)$.

   If x is even, we can write it in the form of $x = 2k$, for $k \in \mathbb{Z}$.

   Squaring both sides gives, $x^2 = 4k^2$. Adding x to the equation on both sides gives, $x^2 + x = 4k^2 + 2k = 2(2k^2 + k)$

   By closure under addition and multiplication of integers, it follows that $2k^2 + k \in \mathbb{Z}$ therefore $E(2(2k^2 + k)) \Leftrightarrow E(x^2 + x)$.

3. **Case 2:** $\neg E(x) : x \in \mathbb{Z} \wedge E(x) \Rightarrow E(x^2+x)$. Again, for an implication, so assume $x \in \mathbb{Z}$ and $\neg E(x)$.

   Given: $x \in \mathbb{Z} \wedge \neg E(x)$. **To Prove:** $E(x^2+x)$.

   Now, if x is odd, we can write it in the form of $x = 2k + 1$, for $k \in \mathbb{Z}$.

   Squaring both sides gives, $x^2 = 4k^2 + 4k + 1$. Adding x on both sides, $x^2 + x = 4k^2 + 2(3k + 1) = 2(2k^2 + 3k + 1)$

   Again, by closure under addition and multiplication of integers, $(2k^2 + 2k + 1) \in \mathbb{Z}$ therefore $E(2(2k^2 + 2k + 1)) \Leftrightarrow E(x^2 + x)$.

Since, we get the same conclusion for both cases, $x \in \mathbb{Z} \Rightarrow E(x^2 + x)$. □

Now, let's write this formally.

Theorem 3.21. *For $x \in \mathbb{Z}$, $x^2 + x$ is an even integer.*

Proof. We will prove by distinction of cases. Suppose $x \in \mathbb{Z}$ and x is even. Then, $x = 2k$, for $k \in \mathbb{Z}$. Squaring both sides and adding x we get, $x^2 + x = 2(2k^2 + k)$. By closure under addition and multiplication of integers, $(2k^2 + k) \in \mathbb{Z}$, therefore $2(2k^2 + k)$ is even, hence $(x^2 + x)$ is also an even integer. Now, Suppose x is odd. Then, $x = 2k + 1$, for $k \in \mathbb{Z}$. Again, $x^2 + x = 2(2k^2 + 3k + 1)$. Since, $(2k^2 + 3k + 1) \in \mathbb{Z}$ therefore $2(2k^2 + 3k + 1)$ is even, and hence $(x^2 + x)$ is also even. Therefore, if $x \in \mathbb{Z}$, $(x^2 + x)$ is even. □

We can also generalise this idea for several statements. We use the tautology:

$$(A_1 \vee A_2 \vee \cdots \vee A_n) \wedge (A_1 \Rightarrow B) \wedge (A_2 \Rightarrow B) \wedge \cdots \wedge (A_n \Rightarrow B) \Rightarrow B$$

To prove a statement B we first show that for the statements $A_1, A_2, \ldots, A_n$ the disjunction $A_1 \vee A_2 \vee \cdots \vee A_n$ holds. Then show that for every case B holds:

1. Assume A_1 and then show B.
2. Assume A_2 and then show B.
3. Assume A_n and then show B.

3.3.6 Proof of Negation

Proving a negative statement $\neg A$ is much harder than proving a positive statement, so our usual strategy would be to use some equivalence to reexpress the statement from something that *shouldn't* be true to something that *should* be true.

It is also helpful to write out the definition of some mathematical symbols, this example illustrates this.

Example 3.22. Suppose $A \cap B \subset C$ and $x \in A$. Prove that $x \notin C \backslash A$.

Solution. Given: $A \cap B \subset C$ and $x \in A$. To prove: $x \notin C \backslash A$.

We start with writing the goal as $\neg(x \in C \backslash A)$, then by definition, $\neg(x \in C \wedge x \notin A)$. Using De Morgan's law and the double negation law, we have $(x \notin C) \vee (x \in A)$, which is true since $x \in A$. Hence, $x \notin C \backslash A$. □

Theorem 3.23. *Suppose $A \cap B \subset C$ and $x \in A$. Prove that $x \notin C \backslash A$.*

Proof. Assume $x \notin C \backslash A$, then by definition, $\neg(x \in C \wedge x \notin A)$. Using De Morgan's law and the double negation law, we have $(x \notin C) \vee (x \in A)$, which is true since $x \in A$. Hence, $x \notin C \backslash A$. □

Sometimes this it may not be easy to reexpress a negative statement as a positive statement. In such case, we do a proof by contradiction, $\neg A \Leftrightarrow (A \Rightarrow \bot)$, assume that A is true and show that from this assumption a false statement follows. This shows that our assumption A must be false to avoid the contradiction.

Example 3.24. Suppose $x^2 + y = 5$ and $x \neq 2$. Prove that $y \neq 1$.

Solution. Given: $x^2 + y = 5$ and $x \neq 2$. To prove: $y \neq 1$.

We'll prove by contradiction. Assume $y = 1$. Then, substituting $y = 1$ in $x^2 + y = 5$, we get $x^2 + 1 = 5 \Rightarrow x^2 = 4 \Rightarrow x = 2$ or $x = -2$. This contradicts our given $x \neq 2$. Hence, our assumption is false and thus $y \neq 1$. □

Note that there were three givens but we concluded that the contradiction was because of the one we assumed and hence the negation of what we assumed must be true.

Proof by contradiction is useful as it allows us to assume that the conclusion is false, providing another given to use but it leaves us with a rather vague goal: produce a contradiction $\bot$, by proving something is false from the givens. This is not an easy task since all our proof strategies discussed so far focus on analysis of the logical form of the *goal*, we don't seem to have a strategy with specifically focuses on producing $\bot$.

In the previous examples we were forced to analyse the givens to produce a contradiction so it seems if one the given has the form $\neg A$, we can produce a contradiction by proving A. This is our first strategy based on logical form the *givens*:

To use a given of the form $\neg A$: For a proof by contradiction, make A you goal and try to prove A. This contradicts the given $\neg A$ and the proof is complete.

This strategy only works for a prove by contradiction. If you are doing some other proof, you can reexpress the given $\neg A$ as a positive statement just like the strategy for a given of this form.

To use a given of the form $\neg A$: If possible, reexpress $\neg A$ as some positive statement.

3.3.7 Proof Involving Quantifiers

In the case where there are multiple quantifiers in the formula, we handle them in the order in which they appear.

3.3.8 Proof of Conjuction and Disjunction

3.3.9 Existence and Uniquness Proofs

3.4 Mathematical Induction

4 Set Theory

Sets are the foundation of modern mathematics. Most mathematical entitics can be studied as sets or classes of objects. We could have already discussed some concpts of sets in chapter 1, they were mainly discussed in the context of 'how to use set theory'. In this chapter, I'll reintroduce them in a more axiomatic and philosophical way, instead of just giving the naive definition of sets.

We start with a brief introduction to Cantor's *naive set theory* and see why it fails when considered sets that are "too large". Then we develope the Zermelo-Fraenkel system with quite a bit of depth, starting with the axiom of extentionality and axiom schema of separation which fixes the issues of Cantor's formulation. Then we'll define operations of sets such as intersection, union, and difference of two sets. After that, we discuss ordinals and cardinals, and axiom of choice. However, a full treatment of *axiomatic set theory* will not be done here.

4.1 From Cantor to Zermelo

Georg Cantor in late 19th century developed the set theory during his studies on infinite series and related topics in analysis. His greatest accomplishment is considered to be the development of the general theory of transfinite numbers, which are now called cardinal numbers and ordinal numbers.

From the standpoint of foundations of mathematics, the philosophically revolutionary aspect of his work was his insistence on the existence of infinite sets as mathematical objects not merely a intuitive idea, but just like natural numbers or finite sets.

Historically, the concept of infinity has been very important for the foundations of mathematics. Since Aristotle, every serious philosopher has exercised with this difficult concept.

We start from somewhat informal discussion of naive set theory. A set is defined as following.

Definition 4.1. A set is an unordered *collection* of mathematical objects called the **elements** of the set. We use the symbol $\in$ to denote membership or elementhood of an object to a set. If x is an object which is an element of set M then we write $x \in M$. We write $x \notin M$ for the opposite.

Remark 4.2. As mentioned, the order of appearance of elements does not matter. For e.g., $M = \{1, 2, 4, 8\}$ is the set of all natural numbers that are factors of 8, here set M is completely defined as we have listed out all of its elements. For e.g., $2 \in M$ but $3 \notin M$. Sets themselves are considered a type of object and can be members of some other set. Such sets are more abstract, such as: $\{\alpha, \beta, \{\gamma\}\}$. Here α and β are elements of this set but γ is not, $\{\gamma\}$ is.

Remark 4.3. In "pure set theory" every object is a set, even natural numbers such as 3, 5, or 67 are also defined as sets in von Neumann ordinals. In the "impure" approach some objects such as the natural number are not considered as a set but for such approach we would need a sort of "dual system" of sets and urelements, in which natural numbers are *cardinalities of sets* rather than pure sets themselves and statements like $6 \in 7$ is meaningless. However, we will follow the modern framework in which every object can be regarded as sets.

4.1.1 Axiom of Extentionality

An important point to make before moving forward, even though we are stating some axioms straight from ZFC, originally, Cantor did not explicitely work with axioms. From the theorems he proved, it can be concluded that all of set theory can be derived from some set of axioms but none of these are directly associated to Cantor's original work.

Axiom 4.4 (Axiom of extentionality)**.** *Two sets M and N are equal, if and only if they have the same elements. In formula:*

$$\forall M \forall N((\forall x(x \in M \Leftrightarrow x \in N)) \Leftrightarrow M = N)$$

Example 4.5. $M = \{2, 3, 7, 9\}$ and $N = \{3, 7, 7, 2, 9\}$ are equal as each element of M is in N and vice-versa. The repetion of 7 does not matter as it does not change the membership 7 in N. You could even say that both are the same set.

Remark 4.6. This axiom specifies when sets are equal. It can be verified that this notion of equality is reflexive, symmetric and transitive. If $x \in M$ and $M = N$ the $x \in N$. Thus, $\in$ relation obeys the axiom of substitution. So, any property defined on $\in$ will obey the axiom of substitution in any first-order theory.

The equality of two sets can also be defined using subsets.

Definition 4.7 (Subsets)**.** A set M is a **subset** of a set N, $M \subseteq N$ iff every element of M is also an element of N. In formula:

$$\forall M \forall N((\forall x(x \in M \Rightarrow x \in N)) \Leftrightarrow M \subseteq N)$$

The equality of sets is then defined as $M = N :\Leftrightarrow M \subseteq N \wedge N \subseteq M$. The axiom of extentionality can be then formulated as:

$$\forall M \forall N(M \subseteq N \wedge N \subseteq M \Leftrightarrow M = N)$$

This principle is very important for proving equality of two sets.

Remark 4.8. Sometimes we write $N \supseteq M$ to say that "N contains M", and N is said to be the **superset** of M. M is a **proper subset** of N, i.e. $M \subset N$ or $M \subsetneq N$, if $M \subseteq N \wedge M \neq N$. The notion of subset obeys the axiom of substitution as it only involves equality and $\in$ relation. Thus for instance, $M \subseteq N \wedge M = M' \Rightarrow M' \subseteq N$. We can also see that $M \subseteq M$ (reflexive), $M \subseteq N \wedge N \subseteq O \Rightarrow M \subseteq O$ (transitive), but $M \subseteq N$ does not necessarily imply $N \subseteq M$, hence it is not symmetric.

Example 4.9. Let $M = \{0, 2, 4\}$, $N = \{0, 1, 2, 3, 4\}$ and $O = \{1, 3\}$. Then (check it for yourself!) we have, $M \subseteq N$ or $M \subset N$, $O \subseteq N$ but $M \nsubseteq O$, $O \nsubseteq M$, $N \nsubseteq M$ and $N \nsubseteq O$. And obviously we have $M \subseteq M$ as well.

4.1.2 Russell's Paradox?!

There is still some issues here, how are we to identify which objects are sets and which are not? This can be answered by the following axiom.

Axiom 4.10 (Axiom of abstraction)**.** *Given any property ϕ there exists a set whose members are just those entities having that property ϕ. In formula:*

$$\exists y \forall x (x \in y \Leftrightarrow \phi(x)).$$

In the formula $\phi(x)$, y is not free.

Thus, $\{x \mid \phi(x)\}$ is the set consisting of all objects x such that $\phi(x)$ is true. For e.g., $\{x \mid x, k \in \mathbb{N} \wedge x = 2k\}$ is the set of all even natural numbers.

This axiom, called the **axiom of abstraction** or the **unrestricted comprehension principle** was formluated by Frege in 1893 as the Axiom V. It asserts that every property correspondes to a set for which that property is true.

The third axiom is the *axiom of choice* which will not be formluated right now.

However, there are terrible consequences of accepting this axiom in our axiomatic foundations of set theory.

In 1901, Bertrand Russell found that the unrestricted comprehension principle leads to a contradiction by considering a set of all objects which are not the members of themselves. This is famously known as **Russell's Paradox**, which you have probably heard about. Here, we will formally derive this paradox from Axiom 4.10.

Russell formulated a property:

$$\phi(x) = x \notin x.$$

which says "x is a set that is not a member of itself". For e.g., $\{1, 3, 5, 7\} \notin \{1, 3, 5, 7\}$ is true because $\{1, 3, 5, 7\}$ is not one of the elements of the set. The only elements are 1, 3, 5 and 7. So, $\phi(\{1, 3, 5, 7\})$ is true.

So, we have an instant of the axiom of abstraction:

$$\exists y \forall x (x \in y \Leftrightarrow x \notin x).$$

which asserts the existence of a set $R = \{x | x \notin x\}$, a set of all sets that do not contain themselves.

Now, we consider whether R is a member of itself or not, i.e., is $R \in R$ true? If R did contain itself, then by definition, $\phi(R)$ is true, which means, "R is a set and $R \notin R$". But if R did not contain itself, i.e. $R \notin R$, then again, by definition, R would be an element of R, i.e. $R \in R$. So, in both cases, we get both $R \in R$ and $R \notin R$. What the f—!

The problem is that by admitting Axiom 4.10, we are allowing sets that are "too large". For e.g., we can create a set U which is a set of *all sets*, which means $U \in U$ as well. This is not what we want to happen.

One way to resolve this issue is to introduce **classes** which are more general than sets and restrict class formation $\{x | \phi(x)\}$ to sets x. "Too large classes" are then not considered as sets. Ot, setting up a hierarchy of objects. At the bottom of the hierarchy are the *primitive objects* — objects that are not sets. For e.g., the natural number 5. Over them, the *primitive sets* — collections of *primitive objects*, such as, $\{5, 10, 15\}$. Then there are *higher sets* which consists of both primitive objects and sets, such as, $\{5, 7, \{9, 0, 4\}\}$, and then we can build sets from these higher sets and at some level we would have *classes* and so on. At every level the elements of those objects can only be the objects below them in the hierarchy. Formulating this this hierarchy is quite complicated so we resort to differet approach.

By not allowing the unrestricted comprehension and only permitting restricted set formation, the contradiction can be resolved. So, we first notice that the axiom of

abstraction is not a *definite* assertion or just one single axiom, but a *schema* to create infinite axioms. By substituting $\phi(x)$ in the formula of Axiom 4.10 with a *definite* formula (in which y is not a free variable) we can create new axioms which are definite assertions.

An axiom which allows us to create many axioms by such substitutions is called an **axiom schema**. In the next section, we will use such an axiom schema to resolve the issue of Russell's Paradox.

4.1.3 Separation Principle

Ernst Zermelo in 1908 postulated such an axiom schema called *axiom schema of separation* (Aussonderung Axiom), which permits us to "separate off" elements of a given set which satisfy some property and use only these elements to form sets. It will formulated more rigorouly later on, for now, it goes something like this:

Axiom 4.11 (Axiom schema of separation)**.** *For every statement $\phi(x)$ about object x and every set M, there exists a set whose elements are exactly the elements of M with the property $\phi(x)$. In formula:*

$$\exists y \forall x (x \in y \Leftrightarrow x \in M \wedge \phi(x))$$

We write $y = \{x \in M | \ \phi(x)\}$ for such sets. And $y \subseteq M$.

Because we are *restricting* what can be an element of a set, this is also known as the **restricted comprehension principle**. This change asserts that existence of sets is conditional, we have to be first given a set M to be able to assert the existence of a subset y of M.

Now, using the same formula $\phi(x) = x \notin x$, we have:

$$\exists y \forall x (x \in y \Leftrightarrow x \in M \wedge x \notin x)$$

if we take $x = y = R$ and let $M = \{\{1\}, \{2\}\}$, for example, we can infer:

$$\exists R (R \in R \Leftrightarrow R \in M \wedge R \notin R)$$

which is not a contradiction, because $M \in M \wedge M \notin M$ is itself a contradiction hence the left-hand side is false, and the right-hand side is also false since $M \notin M$. Therefore the equivalence is a tautology, and the issue is resolved.

We have been using *formula* to substitute for $\phi(x)$ in both the axiom schema of abstraction and axiom schema of separation, what we mean exactly by *formula* will only be precisely defined in the next section.

Zermelo originally formulated the axiom schema of separation something like this: If a statement $\phi(x)$ is *definite* for all elements of a set M, then there is always a subset $M_\phi \subseteq M$ which contains exactly those elements x of M for which ,$\phi(x)$ is true.

Here, a statement is **definite** if it can be decided in a non-arbitrary manner whether or not any object satisfies the statement. Although the decision does not have to be by some effective or finite procedure. For e.g., $\phi(x) = (x \in \mathbb{N}) \wedge (x + 3 = 4)$ is a definite statement as it can be determined without a doubt that $x = 1$ satisfies the statement.

This notion of "definiteness" was clarified by Skolem in 1922, who characterizes definite statements as just those which satisfy his exact definition of formula. A less detailed but essentially correct explaination was given by Fraenkel as well in the same year. Zermelo

also tried to make the notion of definiteness more precise in his 1929 paper but it wasn't satisfactory.

The axioms of modern axiomatic set theory correspond very closely to Zermelo's original axioms from 1908, but in the theory of transfinite induction and ordinal arithmetic, we need to add a stronger axiom schema than that of separation, usually called the *axiom schema of replacement*, given by Fraenkel in 1922. Although we won't be discussing transfinite induction and ordinal arithmetic in these notes.

For these reasons the system of axiomatic set theory is usually called *Zermelo-Fraenkel* set theory, although it would be historically more appropriate to call it *Zermelo-Fraenkel-Skolem* set theory.

4.1.4 More Paradoxes

Some of you might be thinking, "Why did Zermelo wanted to restrict the axiom schema of separation to *definite* statements?" This is because there were other paradoxes discovered before Russell's. There are around fifteen known paradoxes in naive set theory, many of them are slight variations of each other.

F. P. Ramsey [1926] is considered to be the first person who explicitly and clearly divided the paradoxes into two classes: the logical or mathematical paradoxes, and the linguistic or semantical. Roughly speaking, the former arise from purely mathematical constructions such as the Russell's and Burali-Forti paradoxes. The secound class of paradox arise from considerations involving natural language used to discuss mathematics and logic. We will briefly and informally describe only the prominent ones. For a detailed treatment, see BETH1950.

The general idea of the **Burali-Forti Paradox** goes like this: Consider an *ordered set*, which is just a set in which the elements are just lined up in ascending or descending order, basically, it has a notion of "less than" or "greater than", we call this *total order*. More precisely, for any two element, a and b, you can unambiguisly define $a > b$, $a < b$, or $a = b$. This works for any two pairs in the set. For e.g., the set of *natural numbers*, $\{0, 1, 2, 3, \dots\}$ or *real numbers* have the order $\leq$. You can determine that $1 < 2$, $2 < 5$ and so on in the set of natural numbers.

Now, a set is called a *well-ordered set* if every non-empty subset has a *least element*. For e.g., the set of natural numbers is well-ordered with respect to $\leq$. Why? Pick any subset, say $\{3, 9, 67\}$ or $\{5, 10, 15, \dots\}$, there's always a smallest element, 3 in the first and 5 in the secound. However, the set of integers is not well-ordered with respect to $\leq$ because the subset of negative integers $\{-1, -2, -3, \dots\}$ does not have a least element (-1 is the largest negative integer).

For every well-ordered set we can assign an *ordinal number*, which is simply the position of an element in the set. The natural numbers are the simplest ordinals, each number n represents the *order type* of a set with n elements. 0 for an empty set, 1 for set with one element, 2 for set with two elements in order, 3 for set with three elements in order, and so on. These are called *small ordinals*. Essentially, it tells the numbers of elements in a well-ordered set (this is different from *cardinal numbers* as they tell the number of elements in any ordered or unordered set).

The natural numbers have the ordinal ω, which is the smallest infinite ordinal. And you can keep adding more "positions" after the natural numbers to get *large ordinal* $\omega + 1$, $\omega + 2$ and so on.

Moreover, every ordinal is itself a well-ordered set of all smaller ordinals. For e.g., $5 = \{0, 1, 2, 3, 4\}$. And, the set of all ordinal numbers can be assigned an ordinal number

$\mathcal{O}$.

4.2 Axioms of Set Theory

In this chapter, we'll develope the modern framework of *Zermelo-Fraenkel* set theory with the *Axiom of Choice*, this is sometimes called the ZFC-system. We will develop it in the framework of first-order predicate calculus covered in first chapter. The language of set theory consists of two fundamental predicates, the equality "=" and the *membership relation* "$\in$". The *formulas* of set theory are build from the atomic formulas:

$$x = y,\ x \in y$$

and the logical connectives and quantifiers. A formula with all its free variables can be written as: $\varphi(x_1, \ldots, x_n)$. A formula without free variables is calles a sentence.

We begin with von Neumann's fivefold classification of the symbols of the object language into constants, variables, sentential connectives, quantifiers or operators, and punctuation or grouping symbols.

4.2.1 Extentionality and Separation

In previous section, I said that the Russell's Paradox can be solved by introducing classes of sets. But we are working in the Zermelo-Fraenkel System so we will only introduce an informal notion of class, since they are easier to manipulate.

Definition 4.12 (Class)**.** If $\varphi(x, p_1, \ldots, p_n)$ is a formula, then a **class** is:

$$C = \{x |\ \varphi(x, p_1, \ldots, p_n)\}$$

Using Russell's formula, we get the empty subset of a set, which can be defined as:

Definition 4.13 (Empty subset)**.** There exists a set $\emptyset_M$, known as the **empty subset** of M which contains no element. For every object $x \in M$, $x \notin \emptyset_M$. In formula:

$$\emptyset_M \coloneqq \{x \in M |\ x \neq x\}$$

We get two important observation from this definition which are formulated below.

Theorem 4.14. *The empty set possesses every property. In formula:*

$$\forall x \in M(x \in \emptyset_M \Rightarrow \phi(x))$$

Proof. Let x of Type M be given. $(x \in \emptyset_M \Rightarrow \phi(x)) \Leftrightarrow (x \notin \emptyset_M \vee \phi(x))$. $x \notin \emptyset_M$ is true for x of Type M. □

Theorem 4.15. *If M and N are two given sets, then $\emptyset_M = \emptyset_N$, i.e. there exists exactly one **empty set**, $\emptyset$ which is a subset of all sets.*

Proof. From Theorem 4.14, $x \in \emptyset_M \Rightarrow x \in \emptyset_N$, hence, $\emptyset_M \subseteq \emptyset_N$. By symmetry, $\emptyset_N \subseteq \emptyset_M$. Therefore, $\emptyset_M = \emptyset_N$. □

Remark 4.16. The symbol {} is also used to denote the empty set. If a set is not equal to the empty set then it is called **non-empty set**. The next lemma is a consequence of the above theorems.

Lemma 4.17 (Single choice)**.** *Let M be a non-empty set. Then there exists an object x such that $x \in M$.*

Proof. We prove by contradiction. Suppose $x =$ Term of Type M. Suppose $x \notin M$. From Definition 4.13 $x \notin \emptyset$. Thus, $x \in M \Leftrightarrow x \notin \emptyset$. Therfore, $M = \emptyset$ by Axiom 4.4. Which contradicts the given, $M \neq \emptyset$. Therefore $\exists x(x \in M)$. □

Remark 4.18. Lemma 4.17 asserts that given a non-empty set M, we are allowed to choose an element x of M such that $x \in M$. From a given finite number of non-empty sets, $M_1, \ldots, M_n$ we can choose one element $x_1, \ldots, x_n$ from each set; this is known as "finite choice". To choose elements from an infinite number of sets, we need the *axiom of choice*, which we will discuss in Section 4.4.

4.2.2 Operations on Sets

The theorems asserting existence of the intersection and difference of two sets may be proved by use of the axiom schema of separation but not the union of two sets, and so we first introduce the union axiom, which we will show later to be redundant in terms of the full set of axioms.

Axiom 4.19 (Union Axiom)**.** *Let M and N be two subset of X. Then there exists a set V_X, which contains elements that are in M or N. In formula:*

$$\exists V_X \forall M \forall N[\forall x \in X((x \in V_X) \Leftrightarrow (x \in M \vee x \in N))]$$

Using this axiom we will now prove a theorem which will justify the definition of union of two sets.

Theorem 4.20. *There exists a unique set V_X, for every subset M and N of X such that every element $x \in X$ is in V_X iff $x \in M$ or $x \in N$. In formula:*

$$\exists! V_X \forall M \forall N[\forall x \in X((x \in V_X) \Leftrightarrow (x \in M \vee x \in N))]$$

Definition 4.21 (Union)**.** Let M and N be two subsets of X. Then the **union** of the two sets, $V \coloneqq M \cup N$ is the set containing elements from M or N. In formula:

$$M \cup N = V \Leftrightarrow [\forall x \in X((x \in V) \Leftrightarrow (x \in M \vee x \in N)) \wedge V]$$

You might be wondering why did we write V_X in Theorem 4.20 but V in Definition 4.21. It is indeed tempting to just use one of them in both formulas such as V_X, but doing this does not translate back to the general variables well. This choice was made to avoid free occurence of V.

Example 4.22. Let $M = \{6\}$, $N = \{7\}$ and is $M \cup N = \{6, 7\}$.

The following theorem may seem obvious but it is useful to state as we will be using it to prove further theorems.

Theorem 4.23. *An object x is in the set $M \cup N$ iff it is in the set M or the set N. In formula:*

$$x \in M \cup N \Leftrightarrow x \in M \vee x \in N$$

Proof. Using the identity, $M \cup N = M \cup N$ and substituting $M \cup N$ for V in Definition 4.21, we get

$$M \cup N = M \cup N \Leftrightarrow [\forall x \in X((x \in M \cup N) \Leftrightarrow (x \in M \vee x \in N)) \wedge M \cup N]$$

which is a tautology. Hence, the theroem can be infered. □

For working purpose we can write, $M \cup N \coloneqq \{x \in X |\ x \in M \vee x \in N\}$.

Although it is obvious that the operation of union is commutative, associative and idempotent, we will still prove them for the sake of practice, completeness and to justify stating Theorem 4.23.

Theorem 4.24 (Commutativity)**.** $M \cup N = N \cup M$.

Proof. We will prove by cases.

1. **Case:** Suppose $x \in M$. Then, by Theorem 4.23, $x \in M \cup N$. It also holds that $x \in N \cup M$.
2. **Case:** Suppose $x \in N$. Then, by Theorem 4.23, $x \in M \cup N$. It also holds that $x \in N \cup M$.

We get both $x \in M \cup N$ and $x \in N \cup M$ in both cases, hence, by Axiom 4.4, $M \cup N = N \cup M$. □

Theorem 4.25 (Associativity)**.** $(M \cup N) \cup O = M \cup (N \cup O)$.

Proof. We will prove by cases.

1. **Case:** Suppose $x \in M$. Then, by Theorem 4.23, $x \in M \cup N$. It also holds that $x \in N \cup M$.
2. **Case:** Suppose $x \in N$. Then, by Theorem 4.23, $x \in M \cup N$. It also holds that $x \in N \cup M$.

We get both $x \in M \cup N$ and $x \in N \cup M$ in both cases, hence, by Axiom 4.4, $M \cup N = N \cup M$.

$x \in (M \cap N) \cap O \Leftrightarrow x \in (M \cap N) \wedge x \in O$. Which implies, $x \in M \wedge x \in N \wedge x \in O$. Similarily, $x \in M \cap (N \cap O) \Leftrightarrow x \in M \wedge x \in (N \cap O)$. Which implies, $x \in M \wedge x \in N \wedge x \in O$. Hence by Axiom 4.4, $(M \cap N) \cap O = M \cap (N \cap O)$. □

Theorem 4.26 (Idempotence)**.** $M \cup M = M$.

Proof. From Theorem 4.33, $x \in M \cap M \Leftrightarrow x \in M \wedge x \in M$. Since, $x \in M \wedge x \in M \Leftrightarrow x \in M$. Hence, by Definition 4.32, $M \cap M = M$. □

Some other obvious theorems are also stated below.

Theorem 4.27. $M \cup \emptyset = M$

Theorem 4.28. $M \subseteq M \cup N$

Theorem 4.29. $M \subseteq N \Leftrightarrow (M \cup N = N)$

Theorem 4.30. $M \subseteq O \wedge M \subseteq N \Rightarrow M \cup N \subseteq O$

Now, for the intersection of two sets, we will first prove the following theorem.

Theorem 4.31. *There exists a unique set S_X, for every subset M and N of X such that every element $x \in X$ is in S_X iff $x \in M$ and $x \in N$. In formula:*

$$\exists! S_X \forall M \forall N [\forall x \in X((x \in S_X) \Leftrightarrow (x \in M \wedge x \in N))]$$

This theorem now justfies the definition of the intersection of two sets.

Definition 4.32 (Intersection)**.** Let M and N be two subsets of X. Then, the **intersection** of the two sets, $S \coloneqq M \cap N$, is the set containing elements that are in both M and N. In formula:

$$M \cap N = S \Leftrightarrow [\forall x \in X((x \in S) \Leftrightarrow (x \in M \wedge x \in N)) \wedge S]$$

Similar to the union, we will prove the following theorem.

Theorem 4.33. *An object x is in the set $M \cap N$ iff it is in the set M and the set N. In formula:*

$$x \in M \cap N \Leftrightarrow x \in M \wedge x \in N$$

Proof. Using the identity, $M \cap N = M \cap N$ and substituting $M \cap N$ for S_X in Definition 4.32, we get

$$M \cap N = M \cap N \Leftrightarrow [\forall x \in X((x \in M \cap N) \Leftrightarrow (x \in M \wedge x \in N)) \wedge M \cap N]$$

which is a tautology. Hence, the theroem can be infered. □

Again, we can write, $M \cap N \coloneqq \{x \in X |\ x \in M \wedge x \in N\}$.

If the two sets have no element in common, i.e. $M \cap N = \emptyset$ then they are called **disjoint sets**.

Theorems asserting commutativity, associativity and idempotence of intersection are proved below.

Theorem 4.34 (Commutativity)**.** $M \cap N = N \cap M$.

Proof. From Theorem 4.33, $x \in M \cap N \Leftrightarrow x \in M \wedge x \in N$. Similarily, $x \in N \cap M \Leftrightarrow x \in N \wedge x \in M$. We have both $x \in M$ and $x \in N$ in both cases, hence by Axiom 4.4, $M \cap N = N \cap M$. □

Theorem 4.35 (Associativity)**.** $(M \cap N) \cap O = M \cap (N \cap O)$.

Proof. From Theorem 4.33, $x \in (M \cap N) \cap O \Leftrightarrow x \in (M \cap N) \wedge x \in O$. Which implies, $x \in M \wedge x \in N \wedge x \in O$. Similarily, $x \in M \cap (N \cap O) \Leftrightarrow x \in M \wedge x \in (N \cap O)$. Which implies, $x \in M \wedge x \in N \wedge x \in O$. Hence by Axiom 4.4, $(M \cap N) \cap O = M \cap (N \cap O)$. □

Theorem 4.36 (Idempotence)**.** $M \cap M = M$.

Proof. From Theorem 4.33, $x \in M \cap M \Leftrightarrow x \in M \wedge x \in M$. Since, $x \in M \wedge x \in M \Leftrightarrow x \in M$. Hence, by Definition 4.32, $M \cap M = M$. □

Some more theorems are stated below.

Theorem 4.37. $M \cap \emptyset = \emptyset$

Theorem 4.38. $M \cap N \subseteq M$

Theorem 4.39. $M \subseteq N \Leftrightarrow (M \cap N = M)$

We now state two important distribution theorems of intersection and union. These theorems are analagoues to the distribution theorems of conjunction and disjunction.

Theorem 4.40. *Let A, B and C be three subsets of set Ω. Then,*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap B)$$

.

Theorem 4.41. *Let A, B and C be three subsets of set Ω. Then,*

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup B)$$

.

The difference of two sets will be formulated in the same manner as the other two. Staring with the justifying theorem.

Theorem 4.42. *There exists a unique set D_Ω, for every subset M and N of Ω such that every element $x \in \Omega$ is in D_Ω iff $x \in M$ and $x \notin N$. In formula:*

$$\exists! D_\Omega \forall M \forall N [\forall x \in \Omega((x \in D_\Omega) \Leftrightarrow (x \in M \wedge x \notin N))]$$

Definition 4.43 (Difference)**.** Let M and N be two subsets of Ω. Then, the **difference** of the two sets, $D \coloneqq M \backslash N$, is the set containing elements that are in M but not in N. In formula:

$$M \backslash N = D \Leftrightarrow [\forall x \in \Omega((x \in D) \Leftrightarrow (x \in M \wedge x \notin N)) \wedge D]$$

Theorem 4.44. *An object x is in the set $M \backslash N$ iff it is in the set M but not in the set N. In formula:*

$$x \in M \backslash N \Leftrightarrow x \in M \wedge x \notin N$$

In set-builder notation, $M \backslash N \coloneqq \{x \in \Omega |\ x \in M \wedge x \notin N\}$.

Unlike the other two, the difference operation is not idempotent because of the existence of the empty set.

Theorem 4.45 (Non-indempotence of difference)**.** $M \backslash M = \emptyset$.

Some theorems on all three operations are now listed.

Theorem 4.46. $M \backslash (M \cup N) = M \backslash N$.

Theorem 4.47. $M \cap (M \backslash N) = M \backslash N$.

Theorem 4.48. $(M \backslash N) \cup N = M \cup N$.

Theorem 4.49. $(M \cup N) \backslash N = M \backslash N$.

Theorem 4.50. $(M \cap N) \backslash N = \emptyset$.

Theorem 4.51. $(M \backslash N) \cap N = \emptyset$.

Finally the de Morgan's in set theory.

Theorem 4.52. *Let A, B and C be three subsets of set Ω. Then,*

$$A \backslash (B \cup C) = (A \backslash B) \cup (A \backslash B)$$

.

Theorem 4.53. *Let A, B and C be three subsets of set Ω. Then,*

$$A \backslash (B \cap C) = (A \backslash B) \cap (A \backslash B)$$

.

We can define another operation based on the justifying theorem of difference.

Definition 4.54 (Symmetric Difference)**.** Let M and N be two subsets of Ω. Then, the **symmetric difference** of the two sets, $M \triangle N$, is the set containing elements that are in *exactly one* of M or N. In formula:

$$M \triangle N = D_S \Leftrightarrow [\forall x \in \Omega((x \in D_S) \Leftrightarrow \neg(x \in M \Leftrightarrow x \in N)) \wedge D]$$

It can also be defined by the identity: $M \triangle N = (M \backslash N) \cup (N \backslash M)$. For convienice, we can write, $M \triangle N = \{x \in \Omega | \ \neg(x \in M \Leftrightarrow x \in N)\}$

Sometimes the difference is also called the *relative compliment* of M in N. When the *universe* Ω is clear from the context, we write, $M^c \coloneqq \Omega \backslash M$, and call M^c the compliment of M in the universe of discourse. A compliment in the absolute universe V (class of all sets) in von Neumann set theory does exists. But this is not possible in Zermelo-Fraenkel set theory

Theorem 4.55. *There exists a unique set C_Ω, for every subset M of Ω such that every element $x \in \Omega$ is in C_Ω iff $x \notin M$. In formula:*

$$\exists! C_\Omega \forall M[\forall x \in \Omega(x \in C_\Omega \Leftrightarrow x \notin M)]$$

Definition 4.56 (Compliment)**.** Let M be a subsets of Ω. Then, the **compliment** of the set, $M^c \coloneqq \Omega \backslash M$, is the set containing elements that are in Ω but not in M. In formula:

$$\Omega \backslash M = M^c \Leftrightarrow [\forall x \in \Omega((x \in M^c) \Leftrightarrow x \notin M) \wedge M^c]$$

Theorem 4.57. $(A^c \cap B^c)^c = A \cup B$.

Theorem 4.58. $(A^c \cup B^c)^c = A \cap B$.

4.2.3 Pairing Axiom

Right now, we have only introduced just one set, the empty set, which is the simplest set. So, we introduce another set containing only one element x, $\{x\}$ is called the **singleton set**. Note that the object x and $\{x\}$ are two differet objects and should be confused with each other. We will now introduce the *pairing axiom* which allows us to create sets from two given objects.

Axiom 4.59 (Pairing axiom)**.** $\exists A \forall z(z \in A \Leftrightarrow z = x \vee z = y)$

The union axiom could be made part of the axiom system if we replace the pairing axiom with the weaker singleton set axiom. The pairing axiom would then follow from pairs of these singleton sets. However, we want to show that the union axiom is redundant and can be derived from the pairing axiom and the sum axiom.

Pair sets can be defined after proving the following theorem.

Theorem 4.60. $\exists! A \forall z(z \in A \Leftrightarrow z = x \vee z = y)$.

Definition 4.61 (Pair sets)**.** $\{x, y\} = w \Leftrightarrow \forall z(z \in w \Leftrightarrow z = x \vee z = y) \wedge (w \text{ is a set})$

Again, for convienice of writing proofs we state the following lemma.

Lemma 4.62. $z \in \{x, y\} \Leftrightarrow z = x \vee z = y$.

Now, we can prove that elements in a set are unordered.

Lemma 4.63. $\{x, y\} = \{u, v\} \Rightarrow (x = u \wedge y = v) \vee (x = v \wedge y = u)$.

Remark 4.64. Now, the singleton sets can now be defined as a special case of the pair set $\{x\} = \{x, x\}$. It should also be obvious that for every objcets x there exists only one unique singleton set because of the axiom of extentionality. Using the pairing axiom we can also define the triplet sets, quadruplet sets and so on as: $\{x, y, z\} = \{x, y\} \cup \{z\}$ and $\{x, y, z, w\} = \{x, y\} \cup \{z, w\}$.

An obvious corrolary that follows is.

Corollary 4.65. $\{x\} = \{y\} \Leftrightarrow x = y$.

We can now define *ordered pairs* in terms of singleton sets and unordered pair sets. This definition given by Kuratowski (1921) was historically important in reducing the theory of relations to the theory of sets, but the earliest definition permitting this reduction is to be found in Wiener [1914].

Definition 4.66 (Ordered Pair)**.** For every two objects a and b there exists an object, the **ordered pair** $(a, b) = \{\{a\}, \{a, b\}\}$.

The following theorem ensures orderness of an ordered pair.

Lemma 4.67. $\forall x \forall y \forall u \forall v((x, y) = (u, v) \Leftrightarrow x = u \wedge y = v)$

Remark 4.68. Unlike the set $\{x, y\} = \{y, x\}$, the equality $(x, y) = (y, x)$ holds only if $x = y$. In general, for a given list of objects $x_1, \ldots, x_n$ there exists an ordered **n-tuple** $(x_1, \ldots, x_n)$, where the order matters of course. A 3-tuple (x_1, x_2, x_3) is also called a **triple**, and 4-tuple (x_1, x_2, x_3, x_4) is also called a **quadruple**.

Remark 4.69. The objects x_1 and x_2 are called the first and second **components** respectively of the ordered pair $(x = x_1, x_2)$. More generally, for $x = (x_1, \ldots, x_n)$, x_j is the j^{th}-component of the n-tuple for $1 \leq j \leq n$. We also define $\text{pr}_1(x) \coloneqq x_1$, $\text{pr}_2(x) \coloneqq x_2$, and so on up to $\text{pr}_n(x) \coloneqq x_n$. $\text{pr}_j(x) \coloneqq x_j$ is called the j^{th} **projection** of x.

We prove the general property for n-tuples:

Lemma 4.70. $\forall x_1, \ldots, x_n \forall y_1, \ldots, y_n((x_1, \ldots, x_n) = (y_1, \ldots, y_n) \Leftrightarrow x_1 = y_1 \wedge \cdots \wedge x_n = y_n)$.

4.2.4 Definition by Abstraction

4.2.5 Sum Axiom

4.2.6 Power Set Axiom

Axiom 4.71 (Power Set Axiom)**.** *For every set M, there exists a set $\mathcal{P}(M)$ called the **power set** of M, which contains all subsets of M. In formula:*

$$\exists \mathcal{P}(M) \forall N(N \in \mathcal{P}(M) \Leftrightarrow N \subseteq M)$$

We write, $\mathcal{P}(M) \coloneqq \{N |\ N \subseteq M\}$. Sometimes we also write, 2^M for the power set of M.

The following lemmas are an obvious corrolary of the power set axiom.

Lemma 4.72. *Let M be a set. Then,*

- $\emptyset \in \mathcal{P}(M)$ *and* $M \in \mathcal{P}(M)$.
- $x \in M \Leftrightarrow \{x\} \in \mathcal{P}(M)$.
- $N \subseteq M \Leftrightarrow N \in \mathcal{P}(M)$.
- $\mathcal{P}(M) \neq \emptyset$.

4.2.7 Cartesian Product

Definition 4.73 (Cartesian Product)**.** If M and N are sets, then there exist a set $M \times N$ called the **Cartesian product** of M and N, which is the set of all ordered pairs (x, y) with $x \in M$ and $y \in N$. In formula:

$$M \times N = \{z |\ \exists x \in M \exists y \in N : z = (x, y)\}$$

in short notation: $M \times N = \{(x, y) |\ x \in M, y \in N\}$.

Example 4.74. For $M \coloneqq \{x, y\}$ and $N \coloneqq \{*, \Psi, \oplus\}$, the Cartesian product is, $M \times N = \{(x, *), (x, \Psi), (x, \oplus), (y, *), (y, \Psi), (y, \oplus)\}$.

Lemma 4.75. *Let M and N be two sets.*

1. $M \times N = \emptyset \Leftrightarrow (M = \emptyset) \vee (N = \emptyset)$
2. *In general,* $M \times N \neq N \times M$

Proof. 1. We prove the implication "$\Rightarrow$" by contradiction. Suppose $M \times N = \emptyset$. Suppose $\neg((M = \emptyset) \vee (N = \emptyset))$. It follows by De Morgan's law, $(M \neq \emptyset) \wedge (N \neq \emptyset)$. This implies, $\exists x \in M$ and $\exists y \in N$, and hence $\exists (x, y) \in M \times N$. Therefore, by definition, $M \times N \neq \emptyset$, which contradicts $M \times N = \emptyset$. Therefore, $(M = \emptyset) \vee (N = \emptyset)$ and hence, $M \times N = \emptyset \Rightarrow (M = \emptyset) \vee (N = \emptyset)$.

We prove the converse "$\Leftarrow$" by contraposition. Suppose $M \times N \neq \emptyset$. This implies, $\exists x \in M$ and $\exists y \in N$, and hence $\exists (x, y) \in M \times N$. Therefore, by definition, $M \times N \neq \emptyset$. Consequently. $(M \neq \emptyset) \wedge (N \neq \emptyset)$. Then, $\neg((M = \emptyset) \vee (N = \emptyset))$ follows by De Morgan's law. Hence, $M \times N = \emptyset \Leftarrow (M = \emptyset) \vee (N = \emptyset)$.

2. We prove by contradiction. Suppose, $\neg[\exists M \exists N(M \times N \neq N \times M)]$. Equivalentely, $\forall M \forall N(M \times N = N \times M)$. We provide a counterexample. Let $M = \{a\}$, and $N = \{*, \phi\}$. Then, $M \times N = \{(a, *), (a, \phi)\}$ and $N \times M = \{(*, a), (\phi, a)\}$. Clearly, by Lemma 4.67 $(a, *) \neq (*, a)$, hence, by Axiom 4.4, $M \times N \neq N \times M$ contradicting $\forall M \forall N(M \times N = N \times M)$. Hence, $\exists M \exists N(M \times N \neq N \times M)$.

□

Remark 4.76. The Cartesian product of three sets, M_1, M_2 and M_3 is defined by,

$$M_1 \times M_2 \times M_3 \coloneqq (M_1 \times M_2) \times M_3$$

We can repeat this process for n sets, $M_1, \ldots, M_n$ to define the **n-fold Cartesian product**,

$$M_1 \times \cdots \times M_n \coloneqq (M_1 \times \ldots M_{n-1}) \times M_n$$

For an n-tuple, x in $M_1 \times \cdots \times M_n$, we write $(x_1, \ldots, x_n)$ instead of $(\ldots((x_1, x_2), x_3), \ldots, x_n)$ and call x_j the j^{th} component of x for $1 \leq j \leq n$. The element x_j is the projection $\text{pr}_j(x) \coloneqq x_j$ of the general n-tuple x.

Remark 4.77. In the special case where all $M_1, \ldots, M_n$ are the same set M, i.e., $M_1 = \cdots = M_n = M$, we can write, $M^n = M_1 \times \cdots \times M_n$.

4.2.8 Axiom of Regularity

4.3 Countability

4.3.1 Permutations

4.3.2 Equenumerous Sets

4.3.3 Countable Sets

4.3.4 Infinite Products

4.3.5 Uncountable Sets

4.4 Axiom of Choice

5 Relation and Functions

5.1 Relations

We talk about a binary relation,

Definition 5.1 (Relation)**.** Let M and N be two sets. A subset $R \subseteq M \times N$ is called a **(binary) relation** on M and N. We write xRy or $x \underset{R}{\sim} y$ for $(x, y) \in R$.

Remark 5.2. More generally, subsets R of an n-fold Cartesian product $M_1 \times \cdots \times M_n$ are called **n-ary relations** on $M_1, \ldots, M_n$.

Definition 5.3 (Inverse relation)**.** Let $R \subseteq M \times N$ be a binary relation on sets M and N. The **inverse relation** is defined as,

$$R^{-1} \coloneqq \{(y, x) \in N \times M |\ (x, y) \in R\}$$

Definition 5.4 (Composition)**.** Let, L, M, N be three sets. Let $R \subseteq L \times M$ and $S \subseteq M \times N$ be two binary realtions. Then, there exists a **composition** of R followed by S, $S \circ R \subseteq L \times N$ defined by,

$$S \circ R \coloneqq \{(x, z) \in L \times N |\ \exists y \in M : (x, y) \in R \wedge (y, z) \in S\}.$$

Lemma 5.5. *Let $R \subseteq X \times Y$, $S \subseteq Y \times U$ and $T \subseteq U \times V$ be relations. Then the compositions are associative.*

$$(T \circ S) \circ R = T \circ (S \circ R) = T \circ S \circ R \subseteq X \times V.$$

The above lemma can be generalised for several relations.

For large number of relations it is frequently useful to Compositions in a diagram. In a diagram, we write $X \xrightarrow{R} Y$ in place of $R \subseteq X \times Y$. The diagram,

$$\begin{array}{ccc} X & \xrightarrow{R} & Y \\ & \searrow^{A} & \downarrow{\scriptstyle S} \\ & & U \end{array}$$

is called a **commutative diagram**, where $S \circ R = T$. Similarily, the below diagram is commutative if $V \circ U = S \circ R$

$$\begin{array}{ccc} X & \xrightarrow{R} & Y \\ {\scriptstyle A}\downarrow & & \downarrow{\scriptstyle S} \\ U & \xrightarrow{B} & V \end{array}$$

For many arrows, such diagrams are commutative if the following is true: If X and Y are sets in the diagram and one can get from X to Y via two different paths following the arrows, for example,

$$X \xrightarrow{R_1} A_1 \xrightarrow{R_2} A_2 \xrightarrow{R_3} \ldots \xrightarrow{R_n} Y \qquad \text{and} \qquad X \xrightarrow{S_1} A_1 \xrightarrow{S_2} A_2 \xrightarrow{S_3} \ldots \xrightarrow{S_n} Y$$

then the relations $R_n \circ R_{n-1} \circ \cdots \circ R_1$ and $S_n \circ S_{n-1} \circ \cdots \circ S_1$ are equal. For example, the diagram

$$\begin{array}{ccc} X & \xrightarrow{R} & Y \\ {\scriptstyle C}\downarrow & \swarrow^{A}\ \searrow_{B} & \downarrow{\scriptstyle S} \\ V & \xleftarrow[T]{} & U \end{array}$$

is commutative if $A = S \circ R$, $B = T \circ S$ and $C = T \circ S \circ R = T \circ A = B \circ R$, which is the associativity statement of Lemma 5.5.

5.2 Equivalence Relations

Definition 5.6. A relation R on M is **reflexive** if xRx for all $x \in M$, i.e., if R contains the diagonal

$$\Delta_M \coloneqq \{(x, x) |\, x \in M\}.$$

Definition 5.7. A relation R on M is **symmertic** if it holds:

$$\forall x \in M \forall y \in M (xRy \Rightarrow yRx).$$

Remark 5.8. Obviously, R is symmetric if and only if $R^{-1} = R$. A relation is called **antisymmetric** if $\forall x \in M \forall y \in M (xRy \wedge yRx \Rightarrow x = y)$. A relation is called **asymmetric** if $\forall x \in M \forall y \in M (xRy \Rightarrow \neg(yRx))$. And finally, a relation is called **total** if $\forall x \in M \forall y \in M (xRy \vee yRx)$.

Definition 5.9. A relation R on M is **transitive** if it holds:

$$\forall x \in M \forall y \in M \forall z \in M (xRy \wedge yRz \Rightarrow xRz).$$

Definition 5.10. Let N be a nonempty subset of M and R a relation on M. Then the set $R_N \coloneqq (N \times N) \cap R$ is a relation on N called the **restriction** of R to N.

Remark 5.11. Obviously, $xR_N y$ if and only if $x, y \in N$ and xRy. Usually we write R instead of R_N when the context makes clear the set involved.

Definition 5.12 (Equivalence Relation)**.** A relation on M which is reflexive, transitive and symmetric is called an **equivalence relation** on M and is usually denoted $\sim$. For each $x \in M$, the set

$$[x] \coloneqq \{y \in X; y \sim x\}$$

is the **equivalence class** of (or, containing) x, and each $y \in [x]$ is a **representative** of this equivalence class.

Finally,

$$X/\sim \coloneqq \{[x] |\, x \in X\},$$

"X modulo $\sim$", is the set of all equivalence classes of X. Clearly $X/\sim$ is a subset of $\mathcal{P}(X)$.

Definition 5.13 (Partition)**.** A **partition** of a set X is a subset $\mathcal{A} \subseteq \mathcal{P}(X) \backslash \{\emptyset\}$ with the property that, for each $x \in X$, there is a unique $A \in \mathcal{A}$ such that $x \in A$. That is, $\mathcal{A}$ consists of pairwise disjoint subsets of X whose union is X.

Lemma 5.14. *Let $\sim$ be an equivalence relation on X. Then $X/\sim$ is a partition of X.*

Proof. Since, $\forall x \in X (x \in [x])$. Then, $X = \bigcup_{x \in X} [x]$. Suppose, $z \in [x] \cap [y]$. Then, $z \sim x$ and $z \sim y$, and hence $x \sim y$. Hence, $[x] = [y]$. Hence, two equivalence classes are either disjoint or equal. □

It follows immediately from the definition that the function

$$p \coloneqq p_X : X \to X/\sim, \quad x \to [x]$$

is a well defined surjection, the (canonical) **quotient function** from X to $X/\sim$.

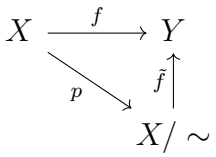
Example 5.15. Let X be the set of inhabitants of Munich. Define a relation on X by $x \sim y :\Leftrightarrow (x$ and y have the same parents). This is clearly an equivalence relation, and two inhabitants of Munich belong to the same equivalence class if and only if they are siblings.

Example 5.16. The "smallest" equivalence relation on a set X is the diagonal Δ_X, that is, the equality relation.

Example 5.17. Let, $f : X \to Y$ be a function. Then

$$x \sim y :\Leftrightarrow f(x) = f(y)$$

is an equivalence relation on X. The equivalence class of $x \in X$ is $[x] = f^{-1}(f(x))$. Moreover, there is a unique function $\tilde{f}$ such that the diagram



is commutative. The function $\tilde{f}$ is injective and $\mathrm{im}(\tilde{f}) = \mathrm{im}(f)$. In particular, f is bijective if f is surjective.

Example 5.18. If $\sim$ is an equivalence relation on a set X, and Y is a nonempty subset of X, then the restriction of $\sim$ to Y is an equivalence relation on Y.

5.3 Order Relation

Definition 5.19 (Partial Order)**.** A relation (R, X, X) on a set X is called an **ordered relation**, **partial order**, or **ordered relation**, if it is reflexive, antisymmetric and transitive.

Remark 5.20. Order relations are usually denoted by $\leq$, $\preceq$ or $\subseteq$. So, If $\preceq$ is a partial order on X, then the pair $(X, \preceq)$ is called a partially ordered set. If the partial order is clear from context, we write simply X for $(X, \preceq)$ and say X is a partially ordered set.

Remark 5.21. If the relation is aditionally total, i.e.,

$$\forall x, y \in X : (x \preceq y) \vee (y \preceq x)$$

it's called a **total ordering** or total order. The pair, $(X, \preceq)$ is then called a **totally ordered set**. We can express the following true statements for an order relation on X that for all $x, y, z \in X$:

- Reflexive: $x \preceq x$
- **Antisymmetric**: $(x \preceq y) \wedge (y \preceq x) \Rightarrow x = y$

- Transitive: $(x \preceq y) \wedge (y \preceq z) \Rightarrow x \preceq z$

The idea of an order relation is that the element of a set with respect to a specific property. in a total order, every element is comparable with every other elements. In a partial order, this is not necessarily the case.

Example 5.22 (Partial order and total order)**.**

1. The identity relation id_X is a partial order on every set X.
2. The universal relation U_X is not a partial order on any set X that contains at least two elements.
3. The relation $\leq$ is a total order on every subset of $\mathbb{R}$.
4. The inclusive relation $\subseteq$ is a partial order on the power set $\mathcal{P}(X)$ of any set X. And a total order if and only if X contains either no elements or exactly one element.
5. The division relation $|$ is a partial order on $\mathbb{N}$.

Lemma 5.23. *If $\preceq$ is a partial order on a set X, then the inverse $\succeq$, is also a partial order on X.*

Proof. DIY! XP □

Lemma 5.24. *If $\preceq$ is a total order on a set X, then $\succeq$, is also a total order on X.*

Proof. DIY! XP □

Definition 5.25. Let X with the relation $\preceq$ be a partially ordered set.

5.4 Operations

5.5 Functions

5.6 Compositions and Commutative Diagrams

5.7 Injections, Surjections and Bijections

5.8 Inverse Functions

5.9 Set-Valued Functions and Fibres

6 Numbers

Characterization of the integers, rational numbers and real numbers has been a central problem for the classical researches of Weierstrass, Dedekind, Kronecker, Frege, Peano, Russell, Whitehead, Brouwer, and others.

In this chapter we will rigorously construct the number systems from first principles in the hierarchy of the number sets $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$, i.e. the *natural*, *integers*, *rational*, *real*, and then the *complex* numbers. Each successive number set can be constructed from the previous one, and we'll see how.

We'll see both the *constructive* and *axiomatic* approaches, but our primary focus would be on the constructive approcah which has the advantage over the *axiomatic* formulation of the real numbers of David Hilbert, that the entire structure of mathematics can be built up from a few foundation stones coming from mathematical logic and axiomatic set theory.

6.1 Structure of Numbers Systems

Even though in analysis courses we construct the number systems from natural numbers to complex in the hierarchy discussed above, this is not the way they were historically developed. Apart from the natural numbers other numbers such as fractions and irrational numbers such as π and square roots were also being used long before, even in ancient civillisations.

To really understand how numbers emerged and evolved to the modern system, let's go all the way back. You might wonder, "*Is this a history class? Why can't we just start with the maths?*"

Don't worry, I won't spend too much time on the history, the goal is to see how number systems evolved and what was the philosophy behind it. Although you can skip this section if you want.

6.1.1 Historical Development

Natural numbers emerged "naturally" due to the need of ancient humans to count *real discrete* objects in the world such as apples, trees and animals. You can find use of notches on bones and marks on the walls of caves as "symbols" for numbers even in the *early stone age*.

But the first systemic use of numbers in civillisations were found in the *valley of the Nile, Euphrates and Tigris*. *Egyptians* for instance made use of **hieroglyphs** for powers of 10 such as the numbers 1, 10, 100, 1000, 10 000, 100 000 and 1 000 000 [9].

The *Babylonians* used **cuneiform symbols** on clay tablets, based on a mixed decimal and sexagesimal position notation which was more complicated than the Egyptians. These symbols looks completely different to the ones we use today but they were enough to form any natural number (except zero) and they could also do addition, subtraction, multiplication and divisions with them, and not only that, they even had rules to form fractions, do operations and solve equations with them.

Babylonians in particular were quite skillful and had sophisticated techniques for arithmetic and algebra, they had considerable influence on the development of these

[9] found on a mace of **King Narmer**, of the first Egyptian dynasty (circa 3000 BC)

areas.[10].

The *earlier system* of the *Greeks* made use of symbols in a decadic number system which could be used to represent numbers such as 1, 5, 10, 50, 100, 500, and other multiples of them.

You have probably recognised the pattern here, every civillisation is creating a system which assigns some set of symbols to *some* numbers which could be used to build *other* numbers, but so far we haven't seen unique symbols for every number between 1-9 like the modern system.

The *later Greek system* made use of the 24 letters of the standard Greek alphabet and three more from the oriental tradition to represent every number between 1-9, 10-90 (multiples of ten: 10, 20,...), 100-900, 1000-9000[11] and 10,000. This is sometimes also called the Ionic system. It was probably the first time when writing numbers wasn't so tedious but calculations in the Greek system was still quite complicated since it wasn't purely positional.

6.1.2 Philosophy of Numbers

Even though I used the modern numbers (1, 2, 3, etc.) as a means to explain what these systems represented, you should not think of those symbols as representing "one", "two", "three" and so one, it's much better to think of it as a *one-to-one correspondence* of representing some "amount" of discrete objects by those symbols.

This could be hard to get your head around since we are so used to connecting the idea of the words "one", "two", "three" to the idea of counting. So, let me give you a famous example: A flock of four sheep and a grove of four trees are related to each other in a way in which neither is related to a pile of three stones or a grove of seven trees. Of course, you know the relationship here is about the 'number' of objects but we don't need to invoke numbers to actually understand the relationship. The relationship which is being refered to is the concept of **cardinal numbers**. Without counting the sheep or the trees, we can pair them with each other, for example by tethering each (and only one) sheep to the trees, so that each sheep and each tree belongs to exactly one of the pairs. Such a pairing between the members of two sets of objects is called a **one-to-one (1-1) correspondence**.

This is why even animals and toodlers can tell the difference between the amount of discrete objects, numaricity is a pre-linguistic concept and we don't need human language for it [12].

You can find a representation of numbers by counters by the Greeks (such as the beads of an abacus, pebbles and so on), which was a means by which arithmetical theorems were discovered.

So, the first step is to separate the idea of Numerosity, the concept of "how many" in quantity, from the concept of Numeral, the symbolic representation.

[10] for a detailed account see H.-D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, J. Neukirch, A. Prestel, R. Remmert, *Zahlen*

[11] strictly speaking the same symbols for 1-9 were used with a subscript accent on the left for 1000-9000.

[12] see *Approximate Number System* for a detailed account of the study on how animals and toddlers also have an idea of cardinal numbers

6.1.3 Formalism of Numbers

The technical term for this concept of 1-1 correspondence is **cardinality**, it was used by Frege and Cantor to define natural numbers as "finite potencies" and "finite cardinal numbers" respectively. A similar concept was used by B. Russell and N. Bourbaki as well. We'll see how the cardinality of sets can be used to formulise numbers in later sections.

If you are thinking that, "*OK, so numbers are just 1-1 correspondence like mapping.*" Then, it's not really true. Some people argue that the concept of cardinality or any sort of correspondence is not really necessary to define numbers, it can be constructed purely from formal axioms as purely abstract idea.

This purely abstract idea of numbers is fine if you just want to do maths. In fact, both of these perspective have thier pros and cons depending on what you are working on, the axiomatic formulation does not care about what the natural numbers are, it just only cares about what are the properties they and what you can do with them. This might seem philosophically dull and it sort of is, but it is actually quite helpful when moving from standard to non-standard numbers and we'll discuss it in further sections how.

You may be a bit annoyed, thinkinh, "*Why can't mathematicians just have one formulation? Why does everyone have different opinions on numerbs? Isn't maths supposed to be objectively true?*"

I get what you feel, it does seem like mathematicians don't seem to agree even on numbers, this is mostly because different people came up with different formulations of number systems but the good part is that all of them are actually equivalent.

It was not until the 19th century that mathematicians gave formal definitions of the concept of number, and their foremost consideration was initially to provide *secure foundations for analysis*, which is why we are discussing about numbers so much. It was not until after Dedekind and Cantor (and others) had defined real numbers by means of sets of rational numbers that the classical definitions of the natural numbers in terms of logic and set theory then followed. The realization that the extensions of the natural numbers to the integers and the rationals could still essentially be regarded as a topic of algebra was closely bound up with the introduction of the fundamental algebraic ideas of ring theory and field theory

6.2 Natural Numbers

We already have an intuitive idea of what the natural naturals are: elements of the set $\mathbb{N} := \{0, 1, 2, 3, \dots\}$[13]. But this definition is not adequate, we don't precisely know what the set $\mathbb{N}$ *is*. What I mean by that is we are constructing the set of natural numbers by using the natural numbers *themselves*.

This definition of the natural numbers is just like saying "start from 0 and count forward indefinitely". But this idea begs the question, "How do we know we can count forward *indefintely* without ending up with some largest number? Or circuling back to some other number say 0 itself?" This question might seem a bit silly but we'll see that if you don't choose your axioms properly then circuling back to 0 could be possible.

[13] in some texts the natural numbers start at 1 instead of 0, but this is just a matter of convention. I believe that it philosophically makes sense to include zero to have a concrete notion of "nothing" in quantities.

6.2.1 The Peano Axioms

The goal is to define the natural numbers using the most primitive principles. A standard way to construct the natural numbers is the based on the axioms of Guiseppe Peano, who formalises the idea that given any natural number, there is always a next largest natural number.

Historically this is not how natural numbers were formulated the first time but we want to be able to derive the rules of arithematics just from these axioms and it is easier for beginners to understand in this way.

Unlike Dedekind, Peano was not interested in set theoretical construction of natural numbers but in axiomatisation in formal language. But, I will still make extensive use of set theory and logic just for the sake of rigour, so make sure you have the necessary background.

From the conclusion of our previous discussions, it seems like the natural number consist of the set $\mathbb{N}$, with 0 as a distinguised element and we can get every other natural number by counting forward. "Counting forward" can be defined as an increment operation or more precisely a **successor function** $\nu : \mathbb{N} \to \mathbb{N}$, thus for $n \in \mathbb{N}$, the element $\nu(n)$ is called the **successor** of n. So, we have the objects: 0, $\nu(0)$, $\nu(\nu(0))$, $\nu(\nu(\nu(0)))$,... as the elements of $\mathbb{N}$. Of course, writing the elements of $\mathbb{N}$ in this way can get very unweirdy, hence we'll adapt a different convention, and write in the familiar notation: $1 := \nu(0), 2 := \nu(1), 3 := \nu(2), \ldots$

If we start writing the axioms for natural numbers it seems like we only need these two axioms to define natural numbers formally:

Axiom 6.1. *Zero is a natural number.* In formula*:* $0 \in \mathbb{N}$*.*

Axiom 6.2. *If n is a natural number, then $\nu(n)$ is the successor of n and it is also a natural number.* In formula: $n \in \mathbb{N} \implies \nu(n) \in \mathbb{N}$

And now we prove for any property of the natural numbers, let's take up an example.

Proposition 6.3. *4 is a natural number.*

Proof. We have from Axiom 2.1, $0 \in \mathbb{N}$, and $n \mapsto \nu(n)$ from Axiom 2.2. To show is $4 \in \mathbb{N}$. From Axiom 2.1 and Axiom 2.2, we get $0 \mapsto \nu(0) = 1$, $1 \mapsto \nu(1) = 2$, $2 \mapsto \nu(2) = 3$, and $3 \mapsto \nu(3) = 4 \Rightarrow 4 \in \mathbb{N}$, which was to be shown. □

It might seem like these two axioms are enough but there are some problems. In previous sections I said that if you don't choose your axioms properly than it is possible to circle back at some number. Since, I did not say whether $4 := \nu(3)$ it is fair to ask how I could have come to this conclusion and not get something like 1 or 2, since the function ν is not injective it would not contradict Axiom 2.1 and 2.2 if $\nu(3) = 1$.

To avoid this issue, we simply define ν to be an injective function and propose this axiom:

Axiom 6.4. *If $n, m \in \mathbb{N}$ are two different natural numbers, i.e. $n \neq m$ then $\nu(n) \neq \nu(m)$.* In formula*:*

$$\forall n, m \in \mathbb{N} : (\nu(n) = \nu(m)) \Rightarrow (n = m)^{14}.$$

[14] this is by using the contraposition of the implication, see *Foundations of Mathematics* for basics of logic.

Now, two numbers cannot have the same successor. Now, to address the other issue, by following our axioms, we could get $\nu(n) = 0$ for some $n \in \mathbb{N}$, for example if we had $\nu(4) = 0$ then we would be circuling from 0,1,2,3,4, to 0 again, which would lead to 0,1,2,3,4,0,1,2,3,... which is exactly what I meant by circling back to 0.

To avoid this we shall say that zero is not a successor of any number. Formally:

Axiom 6.5. *No natural number has 0 as the successor.* In formula*:*

$$\forall n \in \mathbb{N} : \nu(n) \neq 0.$$

Equivalently*:* $0 \notin \nu[\mathbb{N}]$.

Another way to deal with this issue is to redefine the successor function itself as $\nu : \mathbb{N} \to \mathbb{N}^{\times}$, where $\mathbb{N}^{\times} := \mathbb{N}\backslash\{0\}$. This makes ν a bijection and restricts 0 from being a successor of any other number since it's not even in the codomain of the mapping ν, our explicit axiom does this thing in a much simpler way and we can prove any proposition related to natural numbers.

Proposition 6.6. *1 is not equal to 5.*

Proof. To show is $\neg(1 = 5) \Leftrightarrow [(1 = 5) \Rightarrow \bot]$. Let, $1 = 5$. It follows from Axiom 2.2 and 2.3

$$(1 = 5) \Rightarrow (\nu(0) = \nu(4)) \Rightarrow (0 = 4).$$

From Axiom 2.4, $[(1 = 5) \Rightarrow (0 = 4)] \Leftrightarrow \bot$. Hence, $1 \neq 5$, which was to be shown. □

So, it seems like now we've fixed all problems and now our axioms describe the behaviour of natural numbers perfectly. But there is still one issue, we are allowing some Terms[15] which may not be of the Type $\mathbb{N}$. Meaning there may be other "rogue" elements in our number system which are not of the form 0, 1, 2, 3,... Because I did not write out the "..." in Axiom 6.2, we don't know if the "pattern" of the symbols 0, 1, 2, 3, would *for sure* continue in the fashion we want. You could have 0, 1, 2, 3, ϕ, a, π and whatever, and it would still satify all our axioms.

What we want *precisely* is that the 1-1 correspondence between the *objects* 0, $\nu(0)$, $\nu(\nu(0))$, $\nu(\nu(\nu(0)))$,... (which were in $\mathbb{N}$ due to the successor function) and the *symbols* 0, 1, 2, 3,... (which were *not* in $\mathbb{N}$, except 0) should continue, but we did not formally guarantee it for all objects in $\mathbb{N}$[16].

This could be solved by introducing type theory concepts but we want to stick to the historical formalism, we will explore this modern idea later on.

This seems like a very difficult task to do, we would have to write out all the correspondence explicitly and that is very impractical. So, we will use a simple but powerful technique:

Axiom 6.7 (Induction Schema)**.** *For every property $\varphi(n)$ over any natural number n it holds that: if $\varphi(0)$ is true, and if for all $n \in \mathbb{N}$, the property $\varphi(\nu(n))$ follows from $\varphi(n)$, then $\varphi(n)$ is true for all $n \in \mathbb{N}$.*

In formula*:*

$$\Big[\varphi(0) \wedge \forall n \in \mathbb{N} : (\varphi(n) \Rightarrow \varphi(\nu(n)))\Big] \;\Rightarrow\; \forall n \in \mathbb{N} : \varphi(n)$$

[15] see the notes on *Foundations of Mathematics* for basic concepts of Type Theory

[16] you may wonder that this would mean that the deduction in the proofs might be invalid but that's not true because withing our axiom system the conclusions do not create any contradicions.

Remark 6.8. This axiom is stated as a *schema* because it represents an infinite family of axioms - one for each property φ of the natural numbers. More generally, you can say that the induction schema is a pattern of formulas that yield infinitly many axioms having the formula $[\varphi(0) \wedge \forall x : (\varphi(x) \Rightarrow \varphi(\nu(x)))] \Rightarrow \forall x \varphi(x)$ each formula $\varphi(x)$ in the language of arithematics.[17]

This schema says that if some property $\varphi(n)$ holds for 0, meaning $\varphi(0)$ is and if $\varphi(n)$ is true for every natural number, then $\varphi(\nu(n))$ is also true for all natural numbers, and if both of them are true then $\varphi(n)$ holds for all natural numbers. Hence, we can get:

$$\varphi(1) \text{ because } \varphi(0) \Rightarrow \varphi(\nu(0))$$
$$\varphi(2) \text{ because } \varphi(1) \Rightarrow \varphi(\nu(1))$$
$$\varphi(3) \text{ because } \varphi(2) \Rightarrow \varphi(\nu(2))$$
$$\vdots$$

and after recursive application of the induction schema for infinitly many steps, we get - $\forall n \in \mathbb{N} : \varphi(n)$. This avoids any "rogue" objects to appear.

The Axioms 2.1-2.5 are known as the *Peano axioms* for the natural numbers. Historically Peano originally formulated nine axioms (with 1 as the distinguised element)[18], and as I already said, he did not use set theory rather used second-order logic. The first five were the same as ours and the next four were *logical axioms* about equality.

6.2.2 Definition of Natural Numbers

The Peano axiomatisation postulates axioms that any model of the natural number must satify. So far, we have only chosen axioms based on informal reasoning, but we have not provided a formal proof that such a model actually exists.

We will now explore Dedekind's approach and other investigations into proving the existence proof.

To do this, we first define the natural numbers in a modern set-theoretic way. As mentioned before, this definition is equivalent to the Peano Axioms:

Definition 6.9. The **natural numbers** are defined as a triple $(\mathbb{N}, 0, \nu)$, consisting of a set $\mathbb{N}$, with a distinguished element $0 \in \mathbb{N}$, together with a successor function $\nu : \mathbb{N} \to \mathbb{N}$ which satisfy the following axioms:

- ν is injective
- $0 \notin \nu[\mathbb{N}]$
- If $N \subset \mathbb{N}, 0 \in N$, and $\forall n \in \mathbb{N} : \nu(n) \in N$ then $N = \mathbb{N}$

The third property is the set theoretic formulation of the *principle of complete induction*, its equivalent to the Axiom 2.5. can be seen if you replace the property φ by the subset N. The principle of induction is not some new kind of syllogism of mathematicians set apart from the ordinary rules of inference in logic; it is merely the use of the third axiom to prove that certain statements are valid for all natural numbers.

[17] This is just one of the forms of the **induction schema**.

[18] see, Guiseppe Peano, *Arithmetices Principia nova methodo exposita*, 1889

This is an axiomatic definition of the natural numbers, meaning we haven't yet *constructed* the familiar {0,1,2,3,...}, we have only stated that there exists a system $(\mathbb{N}, 0, \nu)$ in which the Peano axioms hold. You could use either the Indo-Arabic system {0,1,2,3,...} or the Roman system {O,I,II,III,...} and both will satify the Peano axioms, in fact these systems are not really different except for the use of different symbols, so we have an **isomorphism** between these two systems and certainly any system $(\mathbb{N}, 0, \nu)$ which satifies the Peano axioms will be isomorphic to any of these two, we will show this isomorphism rigorously in a moment.

First let's consider whether we can show that there exists a **model** for the natural numbers. Clearly if the natural numbers exists they must form an *infinite system: A set M is called an infinite system, if there is an injective mapping $f : M \to M$ such that $f[M] \subset M$.*

This definition expresses the fact that only infinite sets can be mapped injectively onto one of their proper subsets. Historically this was the definition given by Dedekind, though instead of injective mappings, he used the term "ähnliche Abbildungen" (similarity mappings)[19]. The significance of such system is the following theorem proved by Dedekind:

Theorem 6.10. *Any infinite system contains a model $(\mathbb{N}, 0, \nu)$ for the natural numbers.*In other words: *There exists an infinite set if and only if there exists a system $(\mathbb{N}, 0, \nu)$.*

Proof. Let A be an infinite system. Then by definition there is an injective mapping $f : A \to A$ with $f[A] \subset A$. It follows that $0 \in A \implies f(0) \notin A$. Let I be the class of all sets $M \subset A$ with $(0 \in M) \wedge (f[M] \subset M)$. By hypothesis, $I \neq \emptyset$. Thus we can define the intersection $\bigcap_{M \in I} M$. This set satifies the axioms for $(\mathbb{N}, 0, \nu)$ if one takes $f|M$ as the mapping φ. □

Thus to prove the existence of a model we just need to prove the existence of an infinite system.

Dedekind gave an existence proof which implicitly used the **unrestricted comprehension principle** introduced by G. Frege in 1893: *For all property φ of sets, the set $M_\varphi := \{x | \varphi(x)\}$ exists.* But as you may know, B. Russell found that this axiom leads to contradictions. Similar unsuccessful attempt was made by Bolzano[20].

So to prove existence of naturanl numbers in the framework of axiomatic set theory, we assume the *restricted comprehension principle*, and we need the *Infinity axiom: An Inductive set exists.* Here an **inductive set** is a set N which contains $\emptyset$, such that for all $z \in N$, $z \cup \{z\}$ is also in N. We can thus use this inductive to form an infinite system:

$$\mathbb{N} := \{m | m \text{ is an inductive set}\},$$

now we define our successor mapping $\nu : \mathbb{N} \to \mathbb{N}$ by $\nu(z) := z \cup \{z\}$. Now, we'll define $0 := \emptyset$. This construction can now shown that $\mathbb{N}$ is itself an inductive set and our system $(\mathbb{N}, 0, \nu)$ satisfies the Peano axioms. Thus $(\mathbb{N}, 0, \nu)$ is a model for the natural numbers. Below I'll give an informal explanation of how this works:

We start with $0 := \emptyset$, then we use the successor mapping and get

$$1 := \nu(0) = 0 \cup \{0\} = \emptyset \cup \{\emptyset\} = \{\emptyset\} \text{ so } 1 := \{0\},$$
$$2 := \nu(1) = 1 \cup \{1\} = \{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\} \text{ so } 2 := \{0, 1\},$$

[19] see, Richard Dedekind, *Was sind und was sollen die Zahien?*

[20] see, Blozano, *Paradoxien des Unendlichen*

$$3 := \nu(2) = 2 \cup \{2\} = \{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \text{ so } 3 := \{0, 1, 2\},$$

And since we have already said that $\mathbb{N}$ is an inductive set, every natural number is defined as the set of numbers smaller than it. Basically, we already assumed that infinite sets exist and in our proof, $\mathbb{N}$ was the smallest infinite set.

6.2.3 Arithematics of Natural Numbers

6.2.4 The Division Algorithm

6.2.5 Complete Induction

6.2.6 Recursive Definition

6.3 Integers and Rationals

6.3.1 Integers

6.3.2 Rational Numbers

6.3.3 Rational Zeros of Polynomials

6.3.4 Absolute Value, Exponentials, and Square roots

6.3.5 Gaps in Rational Numbers