William Fourie

THE EDIFICE OF TRUTH

December 2016

Contents

Introduction

[This introduction is a work in progress and a place holder. It may or may not be used in the final document.] [Current tips: change the tone a bit and stop using bold face when you are not defining things.]

This is a book about mathematics. In a lazier world, that alone might suffice for an introduction, but in this world perhaps a bit more is required.

The purpose of this book, this project, is to try to build up as much of mathematics as possible, starting from the subject's foundations. Why on earth would one want to do this? Well, simply, in order to explain things.

One day, when meeting with friends, I'd tried to explain the topic of Galois theory to them.[1] Galois theory is a beautiful interplay between the theories of groups and fields, two types of mathematical structures that are often encountered at higher levels.

At the time, these friends did not know what groups were, nor fields, so I tried my best to explain those to them. It took several lunches in front of a whiteboard, and more than a little hand waving at times, but, in the end, I feel as though I got some of the message across. Through it all, I really would have liked to have done better.

This is my attempt to do better.

Originally, Galois Theory was to be the central aim of this project. This would be a very focused text, drilling a hole through the lemmas, definitions, and theorems to explain what the Galois correspondence[2] is, exactly, and impart some understanding of why this is so important.

Originally, the project was to start merely from naive set theory - where one only uses intuitive understandings of the terms 'set' and 'collection' - together with 'obvious' properties of numbers which needed no exposition.

Over time, the aims got broader, and the foundations got deeper, until eventually we ended up with what you see before you.

The hope is that this text will allow any reader of modest ability to learn any topic in mathematics, without relying on any vague prerequisites. This text is intended to be almost entirely self-contained, so as to be especially suitable for autodidacts. One need only skim over the chapters on logic and set theory if they wish, as they are there for

[1] This sort of thing was a regular occurrence at the time, and I can still be a bit insufferable on occasion.

[2] The basic result of Galois theory.

completeness. They are not intended to be a wall barring the less avid readers from entry.

The sheer breath of scope that this project has obtained necessitates a somewhat unusual chapter layout. The book is divided into several numbered **parts**, and within each part, there are **chapters**. Chapters within a part depend only on chapters from earlier parts, and not from chapters within the same or later parts. At the beginning of each chapter (apart from the first) the prerequisites are made clear.

It would be folly to try to read this text 'from cover to cover', and far better to simply choose a topic, and go through the dependencies to get towards it. A graph of which chapters depend on which is provided. [WIP at present]

One could, in principle, follow the original path towards Galois theory, without stopping to see the sights along the way, for example.

The strictly hierarchical nature of this project is why is has been given the name it has. On display here is a towering edifice of **theorems**, **definitions**, **propositions**, **lemmas**, **corollaries**, and more, all starting from a small set of **axioms**.[3] In short, this is a towering edifice of truth.

The inspiration for what this text has become is an ancient textbook dating from the classical era, when Greek geometry dominated European mathematics. Around ???bce, a scholar by the name of Euclid, from the city of Alexandria, authored a set of volumes which came to be known as *The Elements*.

The Elements was astoundingly successful, and in the twenty three centuries since it was written, only The Bible has gone through more editions. *The Elements*, more than any other book, established that mathematics is about **proof**.

What set this work apart from the innumerable other mathematics scrolls which must have been written then or since was Euclid's strict adherence to what we now call the **axiomatic method**.

In an axiomatic method, rather than attempt to make deduction from evidence, one simply states a set of assumptions, or *axioms*, and then proceeds to logically deduce other truths that follow from those axioms. Some of Euclid's assumptions were, for example, that one may extend any finite straight line indefinitely, that one may draw a circle with any center and radius, and that, given a straight line L_1 and a point P not on L_1, there is exactly one line L_2 through P which does not intersect with L_1.[4]

[3] Briefly, an axiom is a statement we assume to be true. A theorem is a statement that we can logically prove is true, given a set of axioms. Propositions, lemmas, and corollaries are just terms for different sorts of theorem - propositions are 'minor theorems', lemmas are technical results used to prove more important theorems, and corollaries are theorems that immediately follow from others.

[4] This axioms in particular is known as the **parallel postulate**. Many mathematicians since Euclid believed, but could not prove, that the parallel postulate could in fact be proved from the other assumptions. As it turned out, this was not the case, as the **non-Euclidean** geometries discovered by Gauss, Bolyai, and Lobachevsky demonstrated.

The axiomatic method allows one to systematically build up a body of knowledge without needing later modification, as opposed to the scientific method, which must by its very nature constantly change. The theorems Euclid proved so long ago are still correct.[5]

Since the classical period, mathematical rigour has waxed and waned - most notably being almost absent during the time of Newton, when the subject of calculus was plagued with unjustified infinitesimals and infinities - but it was only in the nineteenth and twentieth centuries that the foundations of mathematics as we know them today were created. At the time of writing, it is possible to derive almost all of modern mathematics from a set of axioms known as ZFC[6], with a few augmentations, on occasion.

In this text, we shall try not to forget the history of the subject - as glimpsed at, above. The majority of historical notes shall be relegated to biographies given at the end of the book, and introductions to chapters, but a lot of references are still made in the main text and in the numerous footnotes. Hopefully readers will leave this book with not only knowledge of various results, but also with an understanding of the context in which some of these results were proved.

Of course, mere knowledge of results and history is not enough to excel in mathematics at any level; one must practice techniques. In this book, at the end of every chapter, are a select number of **exercises** meant to allow readers to test their understanding of the material, practice techniques - mostly of proof - shown in the chapter, and perhaps explore some extra material which could not find an appropriate place within the chapter itself.

There may not be quite as many questions as in some textbooks, but this is primarily because a great effort has been made into having as extensive an answers section as possible. There is never just one way to approach a problem, but the truly stumped reader can always turn to the back of the book to find at least one approach investigated.[7]

Most of these solutions are the author's own, so there are likely several errors in the answers. It is to be hoped that most of these will be spotted and brought to the author's attention by readers, so that over time the answers section improves.

Finally, it should be noted that this text is best read in electronic form, where the extensive hyperlinking may be used to navigate the document. There are no major issues with reading this text on paper, it is just that one may have a better experience reading the pdf version. For example, whenever we reference a theorem from another chapter - using the full format of 'Theorem [Chapter].[Number]' - the reference is

[5]His results are still correct - save the fact that he used a few assumptions he didn't write down - but we use different axioms, now, so some of his theorems have turned into definitions, such as the Pythagorean theorem turning into the definition of distance.

[6]ZFC is an abbreviation for **Zermelo-Fraenkel with the Axiom of Choice**. The axioms were named after Ernst Zermelo, and Abraham Fraenkel.

[7]Do not be too tempted to do this at the first sign of difficulty, though, for some problems are meant to be blankly stared at for a few hours.

itself a link to the theorem, and one can easily navigate back to where they were using the links back to the table of contents at the top and bottom of every page.

Part 1

TOC

Chapter 1A

Zermelo-Fraenkel Set Theory

1A.1 Necessary Logical Prerequisites

Definition 1A.1. The **alphabet of ZFC+U** consists of the symbols $\forall$, $\exists$, $\wedge$, $\vee$, $\Rightarrow, \Leftrightarrow, \neg, =$, and $\in$, together with parentheses - we will sometimes write square brackets in their place when it aids readability - commas, and variables - which are generally just lower case letters such as $x_1, x_2, \ldots, y_1, y_2, \ldots$ but we may use other letters to represent variables so long as we indicate explicitly or by context that that is what they are.

The symbols $\forall$ and $\exists$ are known as the **universal and existential quantifiers**, respectively, and $\vee, \wedge, \Leftrightarrow, \Rightarrow$, and $\neg$ are known as the **sentential connectives**. Together, these symbols form the **alphabet of first order logic**. We call $\wedge$ and $\vee$ the **disjunction** and **conjunction** symbols, respectively. We call $\Rightarrow$ and $\Leftrightarrow$ the **implication** and **logical equivalence** symbols. We call $\neg$ the symbol of **negation**.

In informal use, we generally read $\forall$ as 'for all', $\exists$ as 'there exists', $\wedge$ as 'and', $\vee$ as 'or', $\Rightarrow$ as 'implies', $\Leftrightarrow$ as 'if and only if', and $\neg$ as 'it is not the case that'. Synonymous phrases are also allowable.

As the reader should know, $=$ is the **equality** sign, and we tend to read $=$ as 'equals', or similar. The symbol $\in$ is known as the **membership** symbol, and generally one reads $\in$ as 'is in' or 'is a member of', or similar. Ξ

Definition 1A.2. A **term** is a symbol which is either a variable or a constant - where constants are new symbols which may be added to our theory according to rules which will be given near the end of this section.[1] Ξ

[1] Some other first order theories such as Peano Arithmetic also start with several *function symbols* - for example, the addition sign - and in those cases the application of a function symbol to a collection of terms is also a term. ZFC+U is capable of defining functions internally, however, so these are not needed.

TOC

Definition 1A.3. Well formed formulas, also known as wffs[2], are certain expressions of the above symbols defined by the following rules:

1. If t, s are terms, then $(t = s)$ and $(t \in s)$ are well-formed formulas.
2. If ϕ, ψ are wffs, then $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $(\phi \Rightarrow \psi)$, $(\phi \Leftrightarrow \psi)$, and $(\neg\phi)$ are also formulas.
3. If ϕ is a formula, and x is a variable, then $((\forall x)\phi)$ and $((\exists x)\phi)$ are formulas. We say that ϕ is the **scope** of the quantifiers $(\forall x)$ and $(\exists x)$ in the above two expressions.
4. No expression is a well-formed formula unless its being so follows from the above three rules.

We sometimes call wffs 'propositions', a word meaning 'statements to be proved'. Ξ

Convention 1A.4. If ϕ is a well-formed formula of the form $(\neg\psi)$, where ψ is some other wff, then we may omit the outer parentheses of ϕ.

If $\phi, \psi, \beta, \gamma$ are wffs, and ϕ is of the form $(\psi \wedge \beta)$ or $(\psi \vee \beta)$, then we may omit the outer parentheses of ϕ so long as it is not part of a larger expression of the form $\neg(\phi)$, $(\phi \wedge \gamma)$, or $(\phi \vee \gamma)$.

If x is a variable, and $\phi, \psi, \beta, \gamma$ are wffs, then if ϕ is of the form $((\exists x)\psi)$ or $((\forall x)\psi)$, then we may omit the outer parentheses of ϕ when it is not part of a larger expression of the form $(\neg\phi)$, $(\phi\wedge\gamma)$, or $(\phi \vee \gamma)$.

If ϕ, ψ, β are wffs, and ϕ is of the form $(\psi \Rightarrow \beta)$ or $(\psi \Leftrightarrow \beta)$, then we may omit the outer parentheses of ϕ so long as it is not part of some larger expression.

We say that $\neg$ has **greater binding strength** than $\wedge$ and $\vee$, which have equal binding strength greater than that of any quantifier $(\forall x)$ or $(\exists x)$. Quantifiers all have equal binding strength greater than that of $\Rightarrow$ and $\Leftrightarrow$, which both have the weakest binding strength of any logical symbol.

Definition 1A.5. If x is a variable, and β is a formula, then we say that an *occurrence* of x in β is **bound** in β if and only if it is either an occurrence of x lying in the scope of a quantifier $(\forall x)$ or $(\exists x)$, or if it occurs next to a quantifier. If an occurrence of x is not bound, we say that it is **free**. Ξ

Definition 1A.6. A variable x is said to itself be **bound** (resp. **free**) in a wff if it has a bound (resp. free) occurrence in that wff.[3] Ξ

[2] We have decided to defer to the more commonly used term 'formulas' as opposed to the more correct 'formulae'. We shall, however, use 'wffs' in favour of the more usual'wfs'.

[3] Note that if a variable does not occur in the formula then it is neither free nor bound, and if it occurs more than once it may be both free and bound.

Definition 1A.7. If β is a wff, t is a term, and y is an individual variable, then we say that t is **free for** y **in** β if and only if no free occurrence of y in β lies within the scope of a quantifier of the form $(\forall x_i)$, where x_i is a variable occurring in t. Ξ

Definition 1A.8. Let $x_1, \ldots, x_n$ be variables, let β be a wff, and let $t_1, \ldots, t_n$ be terms such that t_i is free for x_i in β for each i. We shall write $\beta[t_1, \ldots, t_n/x_1, \ldots, x_n]$ to denote the wff obtained by substituting every free occurrence of x_i in β for t_i, for each i.

If all of the free variables of a wff β occur in the list $x_1, \ldots, x_n$,[4] then we may sometimes write β as $\beta(x_1, \ldots, x_n)$, and then write $\beta(t_1, \ldots, t_n)$ to mean $\beta[t_1, \ldots, t_n/x_1, \ldots, x_n]$. Ξ

Definition 1A.9. We call $\vdash$ the **logical consequence** sign. If Γ is an ordered list $\phi_1, \ldots, \phi_n$ of wffs - where n may be 0 - and ψ is a wff, then we informally read $\Gamma \vdash \phi$ as 'ϕ is a logical consequence of Γ'. The expression $\Gamma \vdash \phi$ is called a **sequent**,[5] as is the symbol $\top$. Ξ

Definition 1A.10. If ϕ, ψ, and γ are wffs, t is a term, x, y are variables, and Γ, Δ, and Φ are (finite) ordered lists of wffs, then the following diagrams are known as **rules of inference:**

1. **Structural rules:**

   (a) Assumption:

$$\frac{\top}{\Gamma, \phi, \Delta \vdash \phi}\ \mathsf{Asm}$$

   (b) Weakening:

$$\frac{\Gamma \vdash \phi}{\Gamma, \psi \vdash \phi}\ \mathsf{Wkn}$$

   (c) Swapping:

$$\frac{\Gamma, \phi, \psi, \Delta \vdash \beta}{\Gamma, \psi, \phi, \Delta \vdash \beta}\ \mathsf{Swp}$$

   (d) Repetition Elimination:

$$\frac{\Gamma, \phi, \phi, \Delta \vdash \psi}{\Gamma, \phi, \Delta \vdash \psi}\ \mathsf{Rep{-}Elim}$$

[4]This does not mean that any of the listed variables actually need be free, merely that no free variables occur outside the list. It need not necessarily even be the case that any of the $x_1, \ldots, x_n$ occur in β.

[5]The word is derived from the *Sequent Calculus*, created by Gerhard Gentzen (1909 - 1945), which is a close relative of the *Natural Deduction Calculus* - also invented by Gentzen - which we are using now. Our meaning of 'sequent' is slightly different from his, however, as Genzen allowed multiple formulas to be placed on the right, rather than just one. Our use is the same as that of [5].

2. Axiom rule

(a) If ϕ has been designated as an **axiom** - we shall soon designate some formulas as axioms - then:

$$\frac{\top}{\vdash \phi}\ \mathsf{Axiom}$$

3. Conjunction rules:

(a) Introduction:

$$\frac{\Gamma \vdash \phi \qquad \Delta \vdash \psi}{\Gamma, \Delta \vdash \phi \wedge \psi}\ \wedge\mathsf{-Intro}$$

(b) Elimination left:

$$\frac{\Gamma \vdash \phi \wedge \psi}{\Gamma \vdash \phi}\ \wedge\mathsf{-Elim_L}$$

(c) Elimination right:

$$\frac{\Gamma \vdash \phi \wedge \psi}{\Gamma \vdash \psi}\ \wedge\mathsf{-Elim_R}$$

4. Implication rules:

(a) Introduction:

$$\frac{\Gamma, \phi \vdash \psi}{\Gamma \vdash \phi \Rightarrow \psi}\ \Rightarrow\mathsf{-Intro}$$

(b) Elimination:

$$\frac{\Gamma \vdash \phi \Rightarrow \psi \qquad \Delta \vdash \phi}{\Gamma, \Delta \vdash \psi}\ \Rightarrow\mathsf{-Elim}$$

5. Disjunction Rules

(a) Introduction left:

$$\frac{\Gamma \vdash \phi}{\Gamma \vdash \phi \vee \psi}\ \vee\mathsf{-Intro_L}$$

(b) Introduction right:

$$\frac{\Gamma \vdash \psi}{\Gamma \vdash \phi \vee \psi}\ \vee\mathsf{-Intro_R}$$

(c) Elimination:

$$\frac{\Gamma \vdash \phi \vee \psi \qquad \Delta, \phi \vdash \gamma \qquad \Phi, \psi \vdash \gamma}{\Gamma, \Delta, \Phi \vdash \gamma}\ \vee\mathsf{-Elim}$$

6. Negation rules:

(a) Introduction:

$$\frac{\Gamma, \phi \vdash \psi \vee \neg\psi}{\Gamma \vdash \neg\phi}\ \neg\text{–Intro}$$

(b) Elimination: [6]

$$\frac{\Gamma \vdash \neg\neg\phi}{\Gamma \vdash \phi}\ \neg\text{–Elim}$$

7. Equivalence rules:

(a) Introduction:

$$\frac{\Gamma \vdash (\phi \Rightarrow \psi) \wedge (\psi \Rightarrow \phi)}{\Gamma \vdash \phi \Leftrightarrow \psi}\ \Leftrightarrow\text{–Intro}$$

(b) Elimination:

$$\frac{\Gamma \vdash \phi \Leftrightarrow \psi}{\Gamma \vdash (\phi \Rightarrow \psi) \wedge (\psi \Rightarrow \phi)}\ \Leftrightarrow\text{–Elim}$$

8. Existential Quantifier rules:

(a) (Introduction) If t is free for x in ϕ, then:

$$\frac{\Gamma \vdash \phi[t/x]}{\Gamma \vdash (\exists x)\phi}\ \exists\text{–Intro}$$

(b) (Elimination) If y is free for x in ϕ, and y does not occur in ψ or in any formula of Δ, then:

$$\frac{\Gamma \vdash (\exists x)\phi \qquad \Delta, \phi[y/x] \vdash \psi}{\Gamma, \Delta \vdash \psi}\ \exists\text{–Elim}$$

9. Universal Quantifier rules:

(a) (Introduction) If x does not occur freely in any formula of Γ, then:

$$\frac{\Gamma \vdash \phi}{\Gamma \vdash (\forall x)\phi}\ \forall\text{–Intro}$$

(b) (Elimination) If t is free for x in ϕ, then:

$$\frac{\Gamma \vdash (\forall x)\phi}{\Gamma \vdash \phi[t/x]}\ \forall\text{–Elim}$$

[6] This rule is sometimes called *The Law of Double Negation.* Some other systems of logic, such as ?'s *intuitionistic* logic, do not use this rule.

A sequent $\Gamma \vdash \phi$ is said to **directly follow from** a collection of other sequents *by virtue of a given rule of inference* if and only if an instance of that rule of inference has $\Gamma \vdash \phi$ as bottom sequent and has the others as top sequents.

A sequent is said to **follow from** a collection of others if and only if it either directly follows from them by virtue of some rule of inference, or it directly follows from another collection of sequents, and each member of that collection follows from the given one.

We say that a wff ϕ **follows from** a collection of sequents if and only if $\vdash \phi$ follows from that collection.

If a sequent follows from $\top$ alone, then we say that that sequent is **valid**. Ξ

Definition 1A.11. If ϕ is a wff, then we say that ϕ is a **theorem** if and only if $\vdash \phi$ is valid. If no application of Axiom is needed to show this, then we call ϕ a **logical truth**. Ξ

Theorem 1A.12. If ϕ is a wff, then $\phi \Rightarrow \phi$.

Proof. We draw a diagram to show that $\phi \Rightarrow \phi$ follows from $\top$, below:

$$\dfrac{\dfrac{\top}{\phi \vdash \phi}\ \mathsf{Asm}}{\vdash \phi \Rightarrow \phi}\ \Rightarrow -\mathsf{Intro}$$

so demonstrating that $\phi \Rightarrow \phi$ is a theorem, and moreover is a logical truth. OEΔ

Theorem 1A.13. If ϕ, ψ are wffs, then $\phi \land \neg\phi \Rightarrow \psi$.

Proof. Again, we draw a diagram, this time a bit taller:

$$\dfrac{\dfrac{\dfrac{\dfrac{\top}{\phi \land \neg\phi, \neg\psi \vdash \phi \land \neg\phi}\ \mathsf{Asm}}{\phi \land \neg\phi \vdash \neg\neg\psi}\ \neg-\mathsf{Intro}}{\phi \land \neg\phi \vdash \psi}\ \neg-\mathsf{Elim}}{\vdash \phi \land \neg\phi \Rightarrow \psi}\ \Rightarrow -\mathsf{Intro}$$

which shows the desired relationship. OEΔ

Theorem 1A.14. If α, β are wffs, then $(\alpha \lor \beta) \land \neg\beta \Rightarrow \alpha$.

Proof. Again we draw a diagram - this time showing that the given formula follows from $\top$ by showing that it follows from sequents which follow from $\top$ - below:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\top}{(\alpha \lor \beta) \land \neg\beta \vdash (\alpha \lor \beta) \land \neg\beta}\ \mathsf{Asm}}{(\alpha \lor \beta) \land \neg\beta \vdash \alpha \lor \beta}\ \land-\mathsf{Elim_L} \quad \vdash \alpha \Rightarrow \alpha \quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\top}{\beta \vdash \beta}\ \mathsf{Asm} \quad \dfrac{\dfrac{\top}{(\alpha \lor \beta) \land \neg\beta \vdash (\alpha \lor \beta) \land \neg\beta}\ \mathsf{Asm}}{(\alpha \lor \beta) \land \neg\beta \vdash \neg\beta}\ \land-\mathsf{Elim_R}}{\beta, (\alpha \lor \beta) \land \neg\beta \vdash \beta \land \neg\beta}\ \land-\mathsf{Intro}}{(\alpha \lor \beta) \land \neg\beta, \beta \vdash \beta \land \neg\beta}\ \mathsf{Swp} \quad \vdash \beta \land \neg\beta \Rightarrow \alpha}{(\alpha \lor \beta) \land \neg\beta, \beta \vdash \alpha}\ \Rightarrow -\mathsf{Elim}}{(\alpha \lor \beta) \land \neg\beta \vdash \beta \Rightarrow \alpha}\ \Rightarrow -\mathsf{Intro}}{(\alpha \lor \beta) \land \neg\beta, (\alpha \lor \beta) \land \neg\beta \vdash \alpha}\ \lor-\mathsf{Elim}}{(\alpha \lor \beta) \land \neg\beta \vdash \alpha}\ \mathsf{Rep}-\mathsf{Elim}}{\vdash (\alpha \lor \beta) \land \neg\beta \Rightarrow \alpha}\ \Rightarrow -\mathsf{Intro}$$

As you can see, these diagrams can get quite large. OEΔ

From these rules and definitions one can derive any number of logical truths, and a great deal are given in the exercises for the reader to prove if they so wish.[7]

From this point onwards, we will implicitly assume all of these truths, as well as several other obvious ones that we may have forgotten to put in the exercises. We will also, for the most part, avoid using the diagrammatic proofs of above in favour of the more usual informal style of argumentation.

Later, we may return to formal logic, in order to study logical systems such as these from a more mathematical point of view, such as in *model theory*, but for now we end this section by stating two more important definitions, as well as the Axioms of Equality.

Definition 1A.15. If $\beta(x)$ is a well-formed formula, and x is a variable, then we write $(\exists! x)\beta(x)$ as an abbreviation of the formula $(\exists x)[\beta(x) \wedge ((\exists y)\beta(y) \Rightarrow (x = y))]$. We read $(\exists! x)\beta(x)$ as **'there exists a unique x such that $\beta(x)$'** and call $\exists!$ the **unique existential quantifier**. The unique existential quantifier has the same binding strength as the other quantifiers. Ξ

Definition 1A.16. If $\beta(x)$ is a well-formed formula, x is a variable, and $(\exists! x)\beta(x)$ is a theorem, then the symbol $\iota x\beta(x)$ is called a **constant**.[8] If $\iota x\beta(x)$ is a constant, then we take $\beta(\iota x\beta(x))$ as an axiom. We read $\iota x\beta(x)$ as **'the unique x such that $\beta(x)$'**. One cannot omit the outer parentheses of β in the expression $\iota x\beta$. Ξ

Axioms of Equality. If β is a well-formed formula, x, y are variables such that y is free for x in β, and β' is the result of substituting some, but not necessarily all, of the occurrences of x in β for y, then the following wffs are known as **Axioms of Equality:**

1. $(\forall x)(x = x)$
2. $(\forall x)(\forall y)((x = y) \Rightarrow (\beta \Rightarrow \beta'))$

As one might guess, we will take every one of the Axioms of Equality to be an axiom. The Axioms of Equality, together with those introduced by constants, are known as our **logical axioms** - this is to distinguish them from the **mathematical axioms**, which we will introduce in the next section.

We will implicitly assume many obvious properties of equality from now on, such as that $(x = y) \Leftrightarrow (y = x)$. The reader may prove such properties in the exercises, if they wish. Ξ

[7] I promise I will eventually add exercises.

[8] Other theories may also start with a several *undefined* constants, such as 0 in Peano Arithmetic, but ZFC+U does not need such things.

1A.2 The Axioms of ZF Set Theory

Convention 1A.17. We shall call terms of set theory - that is, every term encountered from now on[9] **sets**. If x, y are sets, then recall that we read $(x \in y)$ as 'x is a member of y' or 'x is an element of y' or similar.

Definition 1A.18. We introduce the following abbreviations: if A is a set, x is a variable, and β is a formula, then we shall write $((\forall x \in A)\beta)$ to mean $((\forall x)[(x \in A) \Rightarrow \beta])$, and $((\exists x \in A)\beta)$ to mean $(\exists x)[(x \in A) \wedge \beta]$. We call these **bounded quantifiers**, and read the expressions above as 'for all x in A' and 'there exists x in A', respectively.

Very similar examples, such as $(\exists! x \in A)$, $(\forall x > y)$[10], and so on will be used later without comment. It is to be hoped that the reader would be able to parse these expressions from context.

We shall also in general decree that a diagonal slash in a symbol representing a relation between two objects should indicate negation, for example we will write $(x \neq y)$ to mean $\neg(x = y)$ and $(x \notin y)$ for $\neg(x \in y)$. The general meaning of 'relation' shall be covered in a later chapter. Ξ

Our first axiom shall set out precisely when we want two sets to be equal. As the reader would hopefully expect, we shall say that two sets are equal if and only if they have the exact same members.

Axiom of Extensionality. We take the following formula to be an axiom:

$$(\forall A)(\forall B)[(\forall x)(x \in A \Leftrightarrow x \in B) \Rightarrow (A = B)]$$

We may read this as 'for all sets A and for all sets B, if for all x we have that x is in A if and only if x is in B, then $A = B$' or less formally 'if two sets have exactly the same members, then they are equal'. Ξ

Convention 1A.19. If x is a variable and $\beta(x)$ is a formula, then instead of $\iota y((\forall x)((x \in y) \Leftrightarrow \beta(x)))$ we will write $\{x \mid \beta(x)\}$. One may read this as 'the set of all x such that $\beta(x)$'. This is called **set builder notation**.

Theorem 1A.20. If x is a variable and $\beta(x), \gamma(x)$ are formulas such that $(\forall x)(\beta(x) \Leftrightarrow \gamma(x))$ - i.e. such that $(\forall x)(\beta(x) \Leftrightarrow \gamma(x))$ is a theorem - then if we have $(\exists! A)(\forall x)(x \in A \Leftrightarrow \beta(x))$ and $(\exists! B)(\forall x)(x \in B \Leftrightarrow \gamma(x))$, then $\{x \mid \beta(x)\} = \{x \mid \gamma(x)\}$.

Similarly, $\{x \mid \beta(x)\} = \{y \mid \beta(y)\}$ if y is some other variable.

[9] That is, every term encountered in every chapter which depends on this one, and every further term in this chapter.

[10] When we have an appropriate definition of $>$.

Proof. If $z \in \{x \mid \beta(x)\}$, then $\beta(z)$, so as $\beta(z) \Leftrightarrow \gamma(z)$ we have $\gamma(z)$, so $z \in \{x \mid \gamma(x)\}$. Similarly, if $z \in \{x \mid \gamma(x)\}$, then $z \in \{x \mid \beta(x)\}$, so as z was arbitrary we have the desired result from the Axiom of Extensionality.

Similarly, we have that both $z \in \{x \mid \beta(x)\} \Leftrightarrow \beta(z)$, and $z \in \{y \mid \beta(y)\} \Leftrightarrow \beta(y)$, so again by extensionality we have what was to be proved. OEΔ

Now, we might want to set as an axiom that $\{x \mid \beta(x)\}$ exists for any $\beta(x)$, so that we can easily obtain any set we desire. Perhaps we could call this assumption something like the *Axiom Schema of Comprehension*, where we say ‘schema’ to mean a bundle of axioms - one for each formula, here. Unfortunately, as the following theorem shows, this sort of assumption leads only to ruin.

Theorem 1A.21 (Russell’s Paradox[11]). The assumption that, if $\beta(x)$ is any formula, then

$$(\exists B)(\forall x)(x \in B \Leftrightarrow \beta(x))$$

leads to a contradiction.

Proof. Take $\beta(x)$ to be the formula $x \notin x$, then we have by the assumption that $(\exists B)(\forall x)(x \in B \Leftrightarrow x \notin x)$, whence, by extensionality, $\{x \mid x \notin x\}$ is a set. Writing R for $\{x \mid x \notin x\}$, we have $R \in R \Leftrightarrow R \notin R$,[12] which is of course impossible, giving us the desired contradiction. OEΔ

We shall instead adopt a weaker axiom schema:

Axiom Schema of Separation. For each formula $\beta(x)$ not containing the variable B, we shall take the following as an axiom:

$$(\forall A)(\exists B)(\forall x)(x \in B \Leftrightarrow (x \in A) \wedge \beta(x))$$

In other words, so long as β does not contain B, we can form the set $\{x \mid (x \in A) \wedge \beta\}$ for any set A. We tend to write $\{x \in A \mid \beta\}$ instead of $\{x \mid (x \in A) \wedge \beta\}$, for convenience. Ξ

Definition 1A.22. We write $A \subseteq B$ to mean $(\forall x)(x \in A \Rightarrow x \in B)$. If $A \subseteq B$, then we say that A is a **subset** of B, that A is **contained** in B, or similar.

We write $B \supseteq A$ to mean $A \subseteq B$, and in this case say that B is a **superset** of A. We write $A \subset B$ to mean $(A \subseteq B) \wedge (A \neq B)$ and in this case call A a **proper subset** of B.[13] Ξ

[11] Discovered by Bertrand Russell in 1903.

[12] We have omitted the parentheses around $R \in R$, and $R \notin R$ as it is clear what is meant.

[13] Some authors write $\subset$ to mean $\subseteq$, and write $\subsetneq$ for $\subset$, so be aware of this when reading other texts. We prefer to have symmetry with the $\leq$ and $<$ signs.

Theorem 1A.23. If $\beta(x)$ is any suitable formula, and A is any set, then $\{x \in A \mid \beta(x)\} \subseteq A$.

Proof. This follows from the fact that $(x \in A) \wedge \beta(x) \Rightarrow x \in A$, for any x. OEΔ

We now introduce three further axioms which allow us to build more sets we would like to have exist.

Axiom of the Empty Set. There exists a set which does not have any elements. Symbolically:

$$(\exists B)(\forall x)(x \notin B)$$

We could have equivalently written this as $(\exists B)(\forall x)(x \notin B)$ so that it is clearer that extensionality gives unique existence. Ξ

Axiom of Pairs. If x, y are sets, then there exists a set B containing x and y only as members. Symbolically, we take:

$$(\forall x)(\forall y)(\exists B)(\forall z)(z \in B \Leftrightarrow (z = x) \vee (z = y))$$

as an axiom. Ξ

Axiom of Power Sets. For any set A, there exists a set B whose members are precisely the subsets of A. Symbolically, we take:

$$(\forall A)(\exists B)(\forall x)(x \in B \Leftrightarrow x \subseteq A)$$

as an axiom. Ξ

Now, extensionality clearly gives unique existence for the sets specified above, so we may form the following accompanying definitions:

Definition 1A.24. We write $\varnothing$ for the **empty set**, that is, we write $\varnothing$ to mean the set $\{x \mid x \neq x\}$. Ξ

Definition 1A.25. If x, y are sets, then we write $\{x, y\}$ for the **pair set**[14] $\{z \mid (z = x) \vee (z = y)\}$; if $x = y$ then we write $\{x\}$ for $\{x, x\}$, and call $\{x\}$ a **singleton**. Ξ

Definition 1A.26. If A is a set, then we write $\mathcal{P}(A)$ for the **power set** of A - that is, for $\{x \mid x \subseteq A\}$. Ξ

It may be the case, on occasion, that we wish to combine several sets together in some way. Ideally, this would form a new set which has as members all those of the constituent sets. We could perhaps add as an axiom that $\{x \mid (x \in A) \vee (x \in B)\}$ exists for any sets A, B, but such an axiom would only allow us to combine finitely many sets - which is not sufficient for all that a mathematician may want to do.

The following axiom allows us to combine as many sets as we want, so long as we can gather them all as members of some other set:[15]

[14] This is also known as a **doubleton**.

[15] It would be prudent at this point to note that membership is not generally transitive - i.e. if $x \in y$ and $y \in z$, then it is not necessarily the case that $x \in z$. For example, $\varnothing \in \{\varnothing\}$ and $\{\varnothing\} \in \{\{\varnothing\}\}$, but $\varnothing \notin \{\{\varnothing\}\}$.

Axiom of Unions. If A is a set, then there exists a set B consisting of precisely the members of the members of A. That is, symbolically, we take:

$$(\forall A)(\exists B)(\forall x)(x \in B \Leftrightarrow (\exists C)((x \in C) \land (C \in A))$$

as an axiom. Ξ

As usual, extensionality allows us to make an accompanying definition:

Definition 1A.27. If A is a set, then we write $\bigcup A$ to mean the set $\{x \mid (\exists C \in A)(x \in C)\}$. Alternatively, we may write $\bigcup_{a \in A} a$ to mean the same thing.

We call $\bigcup A$ the **union** of all of the elements of A; in particular, if $A = \{x, y\}$ then we call $\bigcup A$ the **union of** x **and** y and write $x \cup y$ for this set. Ξ

Definition 1A.28. If x, y, z are sets, then we define $\{x, y, z\}$ to be $\{x, y\} \cup \{z\}$, and similarly one can define four-sets $\{x, y, z, w\}$, five-sets $\{x, y, z, w, v\}$, and so on. If given $x_1, \ldots, x_n$, we tend to write $x_1 \cup \cdots \cup x_n$ for $\bigcup\{x_1, \ldots, x_n\}$. Ξ

Thus, the above, in principle, allows us to form the union of infinitely many sets, if we have an infinite set of sets. Unfortunately,[16] the axioms so far listed do not allow us to form an infinite set - where at this point we only use the word 'infinite' intuitively. The following definition and axiom addresses this problem:

Definition 1A.29. We say that a set W is **inductive** if and only if $\varnothing \in W$ and, whenever $x \in W$, we have $x \cup \{x\} \in W$. Ξ

Axiom of Infinity. There exists at least one inductive set. Ξ

The reader should see why an inductive set would meet most reasonable definitions of 'infinite'.

Note that the above does not assert the existence of a unique set, as many of the other axioms do. This makes sense, as one would not expect there to be a unique inductive set. Later, we will be able to use the above to construct an explicit inductive set.

We finish off this section with the final two axioms of ZF set theory. They are somewhat harder to motivate than the others, but they are necessary for sets to work the way we want. The first states that if we can assign a set A_i to every member i of some set I, then we can form $\{A \mid (\exists i \in I)(A = A_i)\}$, and the second allows us to prove that there can be no circular chains of membership - i.e. we cannot have $x \in x$ or $x \in y \in x$, or similar.

[16] Unless you are a finitist, at least.

Axiom Schema of Replacement. If $\beta(x, y)$ is a formula which does not contain B and is functional in x - by which we mean that $(\forall x)(\forall y)(\forall z)(\beta(x, y) \wedge \beta(x, z) \Rightarrow (y = z))$ - then there exists a set B containing the 'image' of any set A under this formula - i.e. the set $\{y \mid (\exists x \in A)\beta(x, y)\}$.

More formally, if we write $fn(\beta)$ to mean 'β is functional', then for any formula $\beta(x, y)$ not containing B, we take the following:

$$(\forall A)(fn(\beta) \Rightarrow (\exists B)(\forall y)[(y \in B) \Leftrightarrow (\exists x \in A)\beta(x, y)])$$

as an axiom. Ξ

Axiom of Regularity. If A is a set which is not $\varnothing$, then there exists a member of A which does not share any members with A. Symbolically, we take:

$$(\forall A)(A \neq \varnothing \Rightarrow (\exists a \in A)(\forall x)(x \in a \Rightarrow x \notin A))$$

as an axiom. Ξ

We recommend the reader try now to use the Axiom of Regularity to show that no set can be a member of itself, so that it becomes clearer why the axiom is stated as it is.

1A.3 The Algebra of Sets

In this section we shall show how one can define various operations on sets, and how these operations interact both with each other, and with the subset relation. Most of the material here is not difficult to prove, and it is likely the reader knows these results already, so many proofs are left to the reader.

Definition 1A.30. If A, B are sets, then we write $A \cap B$ for the set $\{x \in A \mid x \in B\}$, and $A \neg B$ for $\{x \in A \mid x \notin B\}$. These definitions are valid by the Axiom Schema of Separation.

$A \cap B$ is called the **intersection** of A and B, and $A \neg B$ is the **relative complement** of B in A. Ξ

We wold like to extend the above definition of intersection to arbitrary sets, in a similar way to how $\bigcup$ extends $\cup$. Thankfully, no new axioms are needed for this.

Theorem 1A.31. If A is a *non-empty* - i.e. not equal to $\varnothing$ - set, then there exists a set B such that, for all x, $(x \in B) \Leftrightarrow (\forall a \in A)(x \in a)$.

Proof. As A is non-empty, there exists some $c \in A$, whence the set $N = \{x \in c \mid (\forall a \in A)(x \in a)\}$ exists by the Axiom Schema of Separation. However, $(x \in c) \wedge (\forall a \in A)(x \in a)$ is equivalent to $(\forall a \in A)(x \in a)$, as $c \in A$, so N is our desired set. OEΔ

Definition 1A.32. If A is a non-empty set, then we write $\bigcap A$ for the set $\{x \mid (\forall a \in A)(x \in a)\}$, and call $\bigcap A$ the **intersection of all the elements of** A. One may also write $\bigcap_{a \in A} a$ for $\bigcap A$. Ξ

One might ask why we need the clause that A is non-empty in the above. Simply put, this is because we cannot reasonably extend the notion of intersection over a set to $\varnothing$, because for any x it is vacuously true that x is a member of every member of $\varnothing$ - as $\varnothing$ has no members.

This would give $\bigcap \varnothing$ as being the set of all sets, which cannot exist as then we could use it together with the Axiom Schema of Separation to create Russell's paradox. Thus, we leave $\bigcap \varnothing$ undefined.

We now list some elementary consequences of the above notions.

Theorem 1A.33. If A, B, C are sets, then we have:

(a) $A \cup B = B \cup A$;

(b) $A \cap B = B \cap A$;

(c) $A \cup (B \cup C) = (A \cup B) \cup C$;

(d) $A \cap (B \cap C) = (A \cap B) \cap C$;

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

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

(g) $C \neg (A \cup B) = (C \neg A) \cap (C \neg B)$;

(h) $C \neg (A \cap B) = (C \neg A) \cup (C \neg B)$.

Proof. These all follow from the logical truths that the reader proved in the exercises of 1A.1 (this implies that the exercises must be at the back and must be individual for each section, then). For example, we have:

$$\begin{aligned}
x \in C \neg (A \cap B) \quad &\Leftrightarrow \quad (x \in C) \wedge (x \notin (A \cap B)) \\
&\Leftrightarrow \quad (x \in C) \wedge \neg(x \in A \wedge x \in B) \\
&\Leftrightarrow \quad (x \in C) \wedge (x \notin A \vee x \notin B) \\
&\Leftrightarrow \quad [(x \in C) \wedge (x \notin A)] \vee [(x \in C) \wedge (x \notin B)] \\
&\Leftrightarrow \quad (x \in C \neg A) \vee (x \in C \neg B) \\
&\Leftrightarrow \quad x \in (C \neg A) \cup (C \neg B)
\end{aligned}$$

which proves (h). OEΔ

Theorem 1A.34. If A, B, C are sets, then:

(a) $A \cup \varnothing = A$;

(b) $A \cap \varnothing = \varnothing$;

(c) $A \cap (B \neg A) = \varnothing$;

(d) $A \cup (B \neg A) = A \cup B$.

Proof. Again, these are not difficult; $x \in A \cup \varnothing$ if and only if $x \in A$ or $x \in \varnothing$, but $x \notin \varnothing$, so this is equivalent to $x \in A$. The others are similar. OEΔ

Theorem 1A.35. If A, B, C are sets, then:

(a) $A \subseteq B \Rightarrow A \cup C \subseteq B \cup C$;

(b) $A \subseteq B \Rightarrow A \cap C \subseteq B \cap C$;

(c) $A \subseteq B \Rightarrow \bigcap A \subseteq \bigcap B$;

(d) $A \subseteq B \Rightarrow (C \neg B) \subseteq (C \neg A)$;

(e) If $A \neq \varnothing$, then $A \subseteq B \Rightarrow \bigcap B \subseteq \bigcap A$;

Proof. These are all straightforward. For example, in (e), if $x \in \bigcap B$, then for all $b \in B$ we have $x \in B$, whence as $A \subseteq B$ we have for all $a \in A$ that $x \in a$, so $x \in \bigcap A$. As x was arbitrary, this gives $A \neq \varnothing \Rightarrow (A \subseteq B \Rightarrow \bigcap B \subseteq \bigcap A)$. OEΔ

Now it is occasionally the case that we may wish to apply one of the above set operations to every element of a set at once, and form a new set out of the results. The below shows that this is possible.

Theorem 1A.36. If A, B are sets, then there exists C (different potentially for each part) such that for all x:

(a) $x \in C \Leftrightarrow (\exists X \in B)(x = A \cap X)$;

(b) $x \in C \Leftrightarrow (\exists X \in B)(x = A \cup X)$;

(c) $x \in C \Leftrightarrow (\exists X \in B)(x = A \neg X)$;

(d) $x \in C \Leftrightarrow (\exists X \in B)(x = \mathcal{P}(X))$.

Proof. Letting X be some arbitrary element of B, we have:

(a) We know that $A \cap X \subseteq A$, so let C be the subset of $\mathcal{P}(A)$ defined to be equal to $\{x \in \mathcal{P}(A) \mid (\exists X \in B)(x = A \cap X)\}$, which exists by the power set and separation axioms.

(b) We know that $A \cup X \subseteq A \cup (\bigcup B)$, so as before we can let $C = \{x \in \mathcal{P}(A \cup (\bigcup B)) \mid (\exists X \in B)(x = A \cup X)\}$, which exists by the union, pairing, power set, and separation axioms.

(c) This follows as in (a), noting that $A \neg X \subseteq A$.

(d) As $X \subseteq \bigcup B$, we know that $\mathcal{P}(X) \subseteq \mathcal{P}\mathcal{P}(\bigcup B)$, whence by the separation axioms we can form

$$\{x \in \mathcal{P}\mathcal{P}(\bigcup B) \mid (\exists X \in B)(x = \mathcal{P}(X)\}$$

and this is our desired C. OEΔ

Definition 1A.37. Given sets A, B, we write $\{A \cap X \mid X \in B\}$, $\{A \cup X \mid X \in B\}, \{A \neg B \mid X \in B\}$, and $\{\mathcal{P}(X) \mid X \in B\}$ for the sets of the above theorem. We will write similar expressions $\{t(X) \mid X \in B\}$ for $\{x \mid (\exists X \in B)(x = t(X))\}$, where $t(X)$ is a term for each X, if we can prove the required existence theorems.[17] Ξ

Theorem 1A.38. If A, B are sets, then we may extend the results of Theorem 1A.33 as follows:

(a) If $B \neq \varnothing$, then $A \cup \bigcap B = \bigcap\{A \cup X \mid X \in B\}$;

(b) $A \cap \bigcup B = \bigcup\{A \cup X \mid X \in B\}$;

(c) If $B \neq \varnothing$, then $A \neg \bigcap B = \bigcup\{A \neg X \mid X \in B\}$;

(d) If $B \neq \varnothing$, then $A \neg \bigcup B = \bigcap\{A \neg X \mid X \in B\}$.

Proof. We prove (d), and leave the rest to the exercises. If $B \neq \varnothing$, then we have for any x:

$$\begin{array}{lll} x \in A \neg \bigcup B & \Leftrightarrow & (x \in A) \wedge (x \notin \bigcup B) \\ & \Leftrightarrow & (x \in A) \wedge \neg(\exists b \in B)(x \in b) \\ & \Leftrightarrow & (x \in A) \wedge (\forall b \in B)(x \notin b) \\ & \Leftrightarrow (!) & (\forall b \in B)(x \in A \neg b) \\ & \Leftrightarrow & x \in \bigcap\{A \neg X \mid X \in B\} \end{array}$$

where the reverse implication of (!) only holds if B is non-empty (why?). OEΔ

1A.4 Further results

In this final section, we shall list a few more results of interest that one can deduce immediately from the axioms, such as that there can be no circular chains of membership.

First we shall construct an explicit inductive set, as promised earlier.

Definition 1A.39. We will sometimes write 0 to mean $\varnothing$, 1 to mean $0 \cup \{0\}$, 2 to mean $1 \cup \{1\}$, and so on.[18] We will of course call 0 **zero**, 1 **one**, and so on. Ξ

Note that, intuitively, each 'n' here has n elements. Later, we will in fact use the above to *define* what it means for a set to 'have n elements'.

Definition 1A.40. If n is a set, then we say that n is a **natural number** if and only if n is a member of every inductive set. Ξ

Example 1A.41. $0, 1, 2, 3, \ldots$ are all natural numbers. Ξ

[17] In practice, this is usually very tedious to do, and the results very obvious.

[18] We cannot explicitly say that we are using a base 10 representation, yet, but that is not necessary to use these symbols as labels.

Theorem 1A.42. There exists a set N whose elements are precisely the natural numbers.

Proof. By the Axiom of Infinity, there exists at least one inductive set, say B, whence by the Axiom Schema of Separation we may form $\{x \in B \mid (\forall A)(A \text{ is inductive } \Rightarrow x \in A)\}$. Write N for this set. We have, for any n, that:

$$\begin{aligned} n \in N &\Leftrightarrow n \in B \text{ and } n \text{ is a member of every inductive set} \\ &\Leftrightarrow n \text{ is a member of every inductive set} \\ &\Leftrightarrow n \text{ is a natural number} \end{aligned}$$

so that N is the desired set. OEΔ

Definition 1A.43. We write ω for the aforementioned set N, and for obvious reasons call it the **set of natural numbers**. Ξ

The reader may have been expecting us to write $\mathbb{N}$ for the above set, and that notation will also be used, but really $\mathbb{N}$ is more appropriately written to represent the algebraic structure consisting of the natural numbers.[19]

It will turn out that there are many possible underlying sets for that structure; ω is used to indicate that the above set is in fact the first *transfinite ordinal*. We will discuss ordinals in detail in later chapters.[20]

We now show that two of our axioms[21] are superfluous, and can actually be deduced as theorems from the others.

Theorem 1A.44. The Axiom Schema of Separation may be derived from the Axiom Schema of Replacement.

Proof. Let $\beta(x)$ be some wff not containing the variable B. Let $\gamma(x, y)$ be the formula $\beta(x) \wedge (x = y)$. It should be clear that if $\gamma(x, y_1)$ and $\gamma(x, y_2)$, then $y_1 = y_2$, so that $\gamma(x, y)$ is functional.

We can thus apply the Axiom Schema of Replacement to deduce:

$$(\forall A)(\exists B)(\forall y)[(y \in B) \Leftrightarrow (\exists x \in A)(\gamma(x, y))]$$

which is equivalent to

$$(\forall A)(\exists B)(\forall y)[(y \in B) \Leftrightarrow ((y \in A) \wedge \beta(y)))]$$

so that we obtain the instance of the Axiom Schema of Separation given by $\beta(x)$. As $\beta(x)$ was arbitrary this gives the desired result. OEΔ

[19] By this, we mean the structure that will arise after we have defined addition and multiplication.

[20] Briefly, *ordinals* are special sets which are well ordered - a term which will be formally defined in chapter ??? - by the $\in$ relation. $0, 1, 2, 3, \ldots$ are all ordinals. Not only can the elements of each ordinal be well ordered by $\in$, but so can the (class of all) ordinals, and under that order, it happens that ω is the first 'infinite' ordinal.

[21] In a sense, infinitely many.

Theorem 1A.45. The Axiom of Pairs is derivable from the Axiom of Power Sets, the Axiom of Extensionality, the Axiom of the Empty Set, and the Axiom Schema of Replacement.

Proof. In the Axiom Schema of Replacement let A be the set $\mathcal{P}\mathcal{P}(\varnothing)$, which exists by the Empty Set, Power Set, and Extensionality axioms. Let x, y be any two sets. Let $\phi(u, v)$ be the formula

$$(u = \varnothing \wedge v = x) \vee (u = \{\varnothing\} \wedge v = y)$$

which 'sends' $\varnothing$ to x, and $\{\varnothing\}$ to y.

It is clear that $\phi(u, v_1) \wedge \phi(u, v_2) \Rightarrow v_1 = v_2$ for any v_1, v_2, as only one of $u = \varnothing$ or $u = \{\varnothing\}$ must be true.

Thus, if one applies the Axiom Schema of Replacement to A and $\phi(u, v)$, then they get:

$$(\exists B)(\forall z)(z \in B \Leftrightarrow (\exists a \in A)((a = \varnothing \wedge z = x) \vee (a = \{\varnothing\} \wedge z = y)))$$

and so as there exists $a_1 \in A$ such that $a = \varnothing$, and $a_2 \in A$ such that $a = \{\varnothing\}$, we have that the above is equivalent to[22]

$$(\exists B)(\forall z)(z \in B \Leftrightarrow (z = x) \vee (z = y))$$

so that B is our desired pair set. OEΔ

Finally, we show that one cannot have any circular chains of membership in ZF, using the Axiom of Regularity.

Theorem 1A.46. If $a_1, \ldots, a_n$ are sets, then it cannot be the case that $a_1 \in a_2 \in \cdots \in a_n \in a_1$.

Proof. If it were the case that $a_1 \in \cdots \in a_n \in a_1$, then - writing a_{n+1} for a_1 if necessary - we have for every a_j that $a_j \in \{a_1, \ldots, a_n\} \cap a_{j+1}$. This contradicts the Axiom of Regularity, as that states that at least one of the a_{j+1} must have an empty intersection with $\{a_1, \ldots, a_n\}$, so in fact such circular membership chains cannot be. OEΔ

In particular, though the reader already knows this, we cannot have $x \in x$ for any x, so if the Russell set $\{x \mid x \notin x\}$ did actually exist, then it would be the set of all sets.

This concludes this introduction to Zermelo-Fraenkel Set Theory. ZF suffices for most applications, but not all, and we will need to adopt two further axioms in later chapters in order to develop our theory fully.

These first of these assumptions is known as the *Axiom of Choice*, which the reader will meet in chapter 2A - this gives us the 'default' working set theory known as ZFC.

[22] We also need the fact that, in general, $(\exists x \in A)(\psi_1(x, y) \vee \psi_1(x, y))$ is equivalent to $((\exists x \in A)\psi_1(x, y)) \vee ((\exists x \in A)\psi_2(x, y))$, but the reader has, of course, already shown this in the exercises.

The second assumption is called the *Axiom of Universes*, and this will be necessary in later chapters on *category theory*, which requires us to have some very large sets.

We may investigate further axioms in later chapters, but this will mostly be purely for the purpose of investigating their consequences, rather than for the purpose of developing a separate mathematical theory.

Part 2

Chapter 2A

Relations, Orders, and Maps

2A.1 Ordered Pairs and Product Sets

Definition 2A.1. Given sets a, b, we define the **(Kuratowski) ordered pair** (a, b) to be the set $\{\{a\}, \{a, b\}\}$.[1] Ξ

Theorem 2A.2. For any a, b, c, d, we have that $(a, b) = (c, d)$ if and only if $a = c$ and $a = d$.

Proof. If $a = c$ and $b = d$ then of course $(a, b) = (c, d)$; it is the converse statement which is interesting.

If $(a, b) = \{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\} = (c, d)$, then either $\{a\} = \{c\}$ or $\{a\} = \{c, d\}$.

If $\{a\} = \{c, d\}$, then $a = c = d$, so $b = c = d = a$.

If $\{a\} = \{c\}$, then $a = c$ and also either $\{a, b\} = \{c\}$ or $\{a, b\} = \{c, d\}$. In the first case, we have $a = b = c = d$. In the second, either $b = c = a = d$ or $b = d$.

Thus, in all cases, $a = c$ and $b = d$, as was to be shown. OEΔ

Definition 2A.3. Given an ordered pair (a, b), we call a the **first (or left) component** of the pair, and call b the **second (or right) component**. This is well defined by Theorem 2A.2. Ξ

Theorem 2A.4. Given any two sets A, B, their exists a set whose members are precisely all the ordered pairs with first component from A and second component from B.

Proof. We can use the Axiom Schema of Separation to obtain the set $P = \{(a, b) \in \mathcal{P}\mathcal{P}(A \cup B) \mid (a \in A) \wedge (b \in B)\}$. For all $a \in A$,

[1] This definition was first given by Kazimierz Kuratowski in 1921, but it is not the first definition of (a, b) ever devised; N. Wiener gave the first definition - namely letting (a, b) be $\{\{\{a\}, \varnothing\}, \{\{b\}\}\}$ - in 1914. One may prove in the exercises that this definition satisfies Theorem 2A.2 and also prove that several other potential definitions do or do not. In the end, it is immaterial which definition we choose to use; Kuratowski's is preferred for its simplicity.

for all $b \in B$, we have that $(a, b) \in \mathcal{P}\mathcal{P}(A \cup B)$, so this P is our desired set. OEΔ

Definition 2A.5. Given sets A, B, we define their **Cartesian Product**[2] $A \times B$ to be the set of all ordered pairs (a, b), where $a \in A$ and $b \in B$. Ξ

[2]Named after René Descartes.

Appendices

TOC

Chapter A

Biographies

A.1 János Bolyai (1902 - 1860)

A.2 René Descartes (1596 - 1650)

A.3 Euclid of Alexandria (c300bce)

A.4 Abraham Fraenkel (1891 - 1965)

A.5 Carl Friedrich Gauss (1777 - 1855)

A.6 Gerhard Gentzen (1909 - 1945)

A.7 Kazimierz Kuratowski (1896 - 1980)

A.8 Nikolai Lobachevsky (1792 - 1856)

And who deserves the credit? And who deserves the blame? Oh, Nikolai Ivanovich Lobachevsky is his name!

A.9 Isaac Newton (1642 - 1726)

There is a bit of uncertainty over what dates to put for Newton's birth and death. This uncertainty does not state from a lack of knowledge as to the point in history that he was born and died, but rather due to the calendars in use at the time.

Using the Julian Calendar - the older calendar, in use in Britain, at the time - Newton was born on Christmas Day of 1642, and died on 20 March 1626.

Under the Gregorian calendar - the one we currently use - Newton was born on the fourth of January, 1643, and died on the thirty-first of March, 1627.[1]

[1] One might wonder why the apparently large discrepancy between his birth dates and his death dates. At the time of Netwon's birth, the Gregorian calendar was ten days ahead of the Julian calendar. At the time of Newton's death, the gap had

TOC

A.10 Bertrand Russell (1872 - 1970)

A.11 Norbert Wiener (1894 - 1964)

A.12 Ernst Zermelo (1871 - 1953)

increased to 11 days. Why, then, does there appear to be almost a year's difference between the death dates? Well, simply put, this was because the new year on the Julian calendar is on the 25th of March, rather than the 1st of January. For more information, see [1].

Chapter B

Exercises

B.1 Chapter 1A

B.1.1 Section 1A.1

Chapter C

Answers for the Exercises

C.1 Chapter 1A

C.1.1 Section 1A.1

Bibliography

[1] Thony Christie, *Calendrical confusion or just when did Newton die?*, `https://thonyc.wordpress.com/2015/03/20/calendrical-confusion-or-just-when-did-newton-die/`, The Renaissance Mathematics, 2015.

[2] Matti Eklund, *On How Logic Became First Order*, Nordic Journal of Philosophical Logic, Vol. 1, No. 2, pp. 147-167, Scandinavian University Press, 1996.

[3] H. B. Enderton, *Elements of Set Theory*, Academic Press, 1977.

[4] Leila Haaparanta, *The Development of Modern Logic*, Oxford University Press, 2009.

[5] M. Huth, and M. Ryan, *Logic in Computer Science: Modelling and Reasoning about Systems*, Cambridge University Press, 2004.

[6] Stephen C. Kleene, *Introduction to Metamathematics*, North-Holland Publishing Co, 1952.

[7] Elliott Mendelson, *Introduction to Mathematical Logic*, 5th edition, CRC Press, 2010.

[8] U.C. Merzbach and C. B. Boyer, *A History of Mathematics*, 3rd edition, Wiley, 2011.

[9] R. M. Robinson, *An Essentially Undecidable Axiom System*, Proceedings of the International Congress of Mathematicians, pp. 729 - 730, 1950.