TOPICS IN SET THEORY AND LOGIC

MATH40000

An MMath thesis on various topics
in (combinatorial) set theory and logic

Manchester, from September 2017 to April 2018

BY
LINUS RICHTER

SUPERVISOR
GARETH JONES

2018
THE UNIVERSITY OF MANCHESTER

Abstract

This report is designed to serve as an introduction to several topics in set theory.

The presentation is threefold: firstly, we introduce cofinality of ordinals. As we shall see, cofinality gives rise to several interesting notions within combinatorial set theory, which, in turn, enable us to develop new techniques to characterise cardinals (regular and singular cardinals will be prevalent in that section and shall be studied rigorously).

Secondly, we introduce several notions of combinatorial set theory, such as clubs and hence stationary sets. As we shall see, the idea of stationary sets is highly applicable, and we will study applications thereof to cardinal arithmetic. Further, we will illustrate the versatility of stationary sets by presenting Silver's theorem, a seminal result on inequalities of cardinal exponentiation. As opposed to the original proof, we will describe a line of reasoning presented by Baumgartner and Prikry which is of a purely combinatorial nature and makes heavy use of stationary sets. As the Generalised Continuum Hypothesis is concerned, by applying Silver's theorem we will even be able to provide a remarkable equality of the continuum function for singular cardinals, which was unanticipated until and highly surprising upon its discovery in 1975.

In the third and final part of this report, which is somewhat disjoint from the previous two, we define the constructible universe and show that it is an inner model of ZF. We then go on to study its properties and, in particular, the characteristics of the Axiom of Constructibility. This report will be concluded by our proof of the relative consistency of AC and ZF, which will follow from our reasoning showing that AC holds in the constructible universe.

Contents

1 Introduction

Given the fact that set theory in its current form is merely a century old, it is remarkable how varied and diverse the field has become. Developed with the idea of defining the founding blocks of all mathematical notions, set theory has quickly evolved; nowadays, it is both utterly interesting in its own right (with many diverse branches of active research) but also a valued tool to mathematicians that permits connections to virtually all areas of mathematics.

This report is designed to give an overview of select topics in combinatorial set theory and the constructible universe. As prerequisites, the reader should possess a solid understanding of ordinals, cardinals, the implications of the Axiom of Choice, as well as cardinal arithmetic. Necessary notions will be defined rigorously in the preliminaries section 2 and, if necessary, recapped throughout.

The first part (section 3) covers multiple topics in combinatorial set theory. We begin by introducing *cofinality* which will provide us with a deeper understanding of ordinals and their characteristics. Cofinality will enable us to obtain further inequalities of cardinal arithmetic, which in turn allow us to characterise cardinals in a particular way indicating their internal structure (*regular and singular cardinals* will play a major role throughout that and the subsequent sections). We will also investigate the application of cofinality to partially ordered sets, as well-founded posets permit a fairly simple adaptation of the theory of cofinality.

We then go on to define several types of constructions of set systems in section 4, each being interesting mathematically in its own right. After introducing *filters and ideals* we introduce a crucial notion of combinatorial set theory: the concept of *stationary sets* will give us powerful tools which we shall use to prove seminal results in combinatorial set theory that combine ideas of regular and singular cardinals as well as stationary sets (such as *Fodor's theorem*). The section will be concluded by *Silver's theorem*, whose combinatorial proof will relate the Generalised Continuum Hypothesis to the aforementioned theory. This will be a prime example of the versatility of stationary sets: the original proof of Silver's theorem from 1975 was based on the idea of *forcing* introduced by Paul J. Cohen a decade earlier. By obtaining a purely combinatorial proof, Baumgartner and Prikry translated the problem from the realm of inner model theory into combinatorial set theory.

The final part (section 5) will be somewhat disjoint from the previous two. Based on Keith Devlin's text book *Constructibility* (2017) (reference [Dev17]), we will define *Gödel's constructible universe* L and investigate its implications on the consistency of axioms with regard to ZF. We will begin by describing the metatheory expressed in the language of set theory. After internalising the theory into the universe of sets, we proceed by considering the notion of definability, which will be crucial throughout. We shall study the idea of *inner models of* ZF and, eventually, by showing that ZF proves that the Axiom of Choice holds in L, we will deduce that AC is relatively consistent with ZF.

Unless stated otherwise, all proofs have been found by the author.

We hope the reader will find the material presented on the ensuing pages as fascinating as the author perceived it at the time of writing.

2 Notation, Conventions and Preliminaries

We use the following notation and conventions. Many well-known results are given below as well for completeness.

- The set of natural numbers is denoted by $\omega = \{0, 1, 2, 3, \dots\}$.
- If X is a set, then $\mathcal{P}(X)$ denotes the power set of X.
- If $A \subset X$, then we denote the complement of A with respect to X by A^c. That is, we define $A^\mathsf{c} := X \setminus A$.
- In this account, we make no notational difference between $\subset$ and $\subseteq$, and $\supset$ and $\supseteq$, respectively. Whenever proper subsets are required, those will be mentioned explicitly.
- If X is a set then

  $$\bigcup X := \bigcup_{S \in X} S.$$

  Similarly, provided X is non-empty, we define

  $$\bigcap X := \bigcap_{S \in X} S.$$

- For a set X, we denote the disjoint union of the elements of X by $\bigsqcup X$.
- Whenever sets are considered, we abbreviate *linearly ordered* by simply writing *ordered* for the sake of improved readability.

Functions will be ubiquitous in this report. Hence the following note is issued: the reader is advised to make themselves familiar with our choice of nomenclature regarding the *range* and the *image* of functions so as to avoid confusion. If f is a function between sets X and Y, then we call Y the range of f. The image of f, denoted by $\mathrm{img}(f)$, is defined to be the set comprising exactly those $y \in Y$ for which there exists $x \in X$ such that $f(x) = y$.

Remark. *Occasionally, we will consider class functions. Clearly, such functions will not be sets of ordered pairs (as traditionally defined in set theory). However, they will still enjoy the same properties set functions possess. We will not go into detail on the logical foundation that allows us to make this assumption. Rather, we will take such properties for granted.*

2.1 Well-ordered Sets, Ordinals, and the Axiom of Choice

Definition 2.1. A set system is a class of sets.

We do not define set systems as subsets of power sets; some set systems we will consider are proper classes, e.g. the class of all ordinals.

The branches of set theory we will deal with require the Axiom of Choice as one of their main tools. We will use it in the following way; that is, whenever the Axiom of Choice is mentioned, we refer to this way of stating it:

Axiom (The Axiom of Choice). *Let* $\{C_\gamma : \gamma \in \Gamma\}$ *be a set where* C_γ *is non-empty for every* $\gamma \in \Gamma$. *Then there exists a function* f *with domain* Γ *such that* $f(\gamma) \in C_\gamma$ *for each* $\gamma \in \Gamma$. *We shall call this function a choice function on the set system.*

We will abbreviate the Axiom of Choice by AC and, further, we will assume AC from now on (occasionally, we may mention the usage of AC explicitly).

Remark. *There are many different ways of stating the Axiom of Choice; that is,* AC *is equivalent to a variety of different statements. As it turns out, such results might be more applicable in different contexts than using the original definition of* AC *as given above. More about this follows later once we have introduced cardinalities and well-orderings.*

We will need the following results in the course of this account.

Definition 2.2. Let X and Y be sets. We say that two sets have the same cardinality if there is a bijection between them. We then write $X \sim Y$. If a function f gives a suitable bijection, we write $X \sim_f Y$.

Remark. *It can easily be shown that this yields an equivalence relation:*

- *Reflexive:* $X \sim_f X$ *(trivially by the identity map* $x \mapsto x$*)*
- *Symmetric:* $X \sim_f Y \Rightarrow Y \sim_{f^{-1}} X$ *(a bijection gives rise to an inverse function)*
- *Transitive:* $X \sim_f Y \wedge Y \sim_g Z \Rightarrow X \sim_{g \circ f} Z$ *(composing bijective functions gives the required bijection)*

Hence cardinality induces an equivalence relation, as required. It is worth mentioning, though, that we have not defined a universe in which the equivalence relation holds. It does, however, hold hereditarily on each set within the universe we choose.

Now we can treat sets as representatives of equivalence classes, and therefore the following definition makes sense:

Definition 2.3. We call two sets X and Y equivalent if they have the same cardinality.

In order to be able to deal with ordinals and cardinals (and, more importantly, the notion of well-ordered sets), both will be defined here:

Definition 2.4. An ordered set $\langle X, \prec \rangle$ is called well-ordered if every non-empty subset $Y \subset X$ has a least element with respect to $\prec$.

Remark. *It can be shown that, using order-isomorphisms, these give rise to an equivalence relation of well-ordered sets. We denote the operation that is compatible with order-isomorphisms by* type*, which gives rise to the notion of order types:*

Definition 2.5. A set α is called an ordinal if the following two conditions hold:

- If β is an element of α, then β is a subset of α. (We say that α is a transitive set.)
- α is strictly well-ordered by membership.
  Here, we define membership on α by $\in_\alpha = \{\langle x, y \rangle : x, y \in \alpha, x \in y\}$.

Remark. *Transitivity of* α *can be interpreted as follows:*

$$\textit{If } \gamma \in \beta \textit{ and } \beta \in \alpha, \textit{ then } \gamma \in \alpha.$$

To show this, consider $\beta \in \alpha$ and assume that $\gamma \in \beta$. Then, by transitivity, $\beta \subset \alpha$, and so $\gamma \in \alpha$, as required.

The notation we use with ordinals is one of the following:

Remark. *For ordinals α and β, we define*

$$\alpha \in \beta \Leftrightarrow \alpha < \beta,$$

and we will use both notations depending on which one is more intuitive in the respective context.

Definition 2.6. We say that a well-ordered set $\langle X, \prec \rangle$ is of order type α if there exists an order-isomorphism from $\langle X, \prec \rangle$ to $\langle \alpha, < \rangle$.
We then write $\text{type}(X, \prec) = \alpha$.

By a standard theorem, one can show that the order types of well-ordered sets are ordinals. Further, note that no ordinal (and in fact no well-ordered set in general) can be order-isomorphic to an initial segment of itself.

Remark. *Whenever the ordering is implicit, it may be omitted and we shall write* $\text{type}(X)$.

Proposition 2.7. *Let A be a set of ordinals. Then A has an upper bound that is an ordinal. The least upper bound is $\bigcup A$. We also denote $\bigcup A$ by* $\sup(A)$.

The proof of this result is omitted, it can be found in [HH99, p. 65].

Remark. *Note that for any set A of ordinals it is clear that* $\sup(A)$ *is a limit ordinal if and only if* $\max(A)$ *exists and is a limit ordinal itself or if* $\max(A)$ *does not exist at all.*

It can be shown easily that the class of all ordinals is a proper class. We will denote the class of all ordinal numbers by **ON**.

In order to develop our understanding of cardinals, we will require the following standard theorem:

Theorem 2.8 (Hartogs' Lemma)**.** *Let A be a set. Then there exists an ordinal α such that there is no injection from α into A.*

Remark. *Assuming the Axiom of Choice, it is clear that Hartogs' lemma is an immediate consequence (A can be well-ordered, hence it is order-isomorphic to a unique ordinal. The fact that there are infinitely many ordinals yields the result). The lemma also holds in* ZF*, though.*

We will now focus on cardinals: as mentioned earlier, we assume AC and we may therefore use the von Neumann cardinal assignment:

Definition 2.9. An ordinal α is said to be a cardinal if there is no bijection from α to any ordinal smaller than α. In symbols,

$$\forall \beta < \alpha (\beta \not\sim \alpha).$$

Using the definition of ordinals, this is the same as saying that a cardinal is not in bijection with any of its initial segments.

In general, we denote cardinals by letters from the middle of the Greek alphabet such as κ, λ, μ, et cetera. Ordinals, in contrast, are denoted by letters from the beginning of the Greek alphabet such as α, β, γ, and so on.

Definition 2.10. Let X be a set and let κ be a cardinal. The cardinality of X is given by κ if X is equivalent to κ.

As promised, we now briefly return to AC: in set theory, it is often much easier to appeal to a result that is equivalent to AC rather than applying the axiom directly. In particular, the *well-ordering principle* and the *trichotomy property of cardinalities* spring to mind. Thus, we present the following result (these have been introduced in MATH43021 Set Theory):

Theorem 2.11 (Equivalences of AC)**.** *The following are equivalent:*

(1): AC

(2): If X is a set, then there exists an ordering $\prec$ such that $\langle X, \prec \rangle$ is a well-ordered set. (This is known as the well-ordering principle.)

(3): Let X, Y be sets. Then there exists either an injection from X to Y or an injection from Y to X.

(4): Let $\langle P, \prec \rangle$ be a non-empty partially ordered set. If every chain in P has an upper bound, then P has a maximal element. (This is Zorn's lemma. Partially ordered sets will be used later on, and formal definitions of all terms used will be given.)

Remark. *Note that the Schröder-Bernstein theorem yields that (3) implies the trichotomy property of cardinalities. (The Schröder-Bernstein theorem states that if X, Y are sets and there exist injections both ways, then the cardinalities of X and Y coincide.) Of course, the Schröder-Bernstein theorem is immediate when* AC *is assumed as every set can be well-ordered, and it can be proven that between well-ordered sets there exists either an order-isomorphism or one set is order-isomorphic to an initial segment of the other.*

Further, as we assume AC*, the trichotomy property of cardinals implies that Hartogs' lemma can be interpreted as a statement about cardinals: if X is a set, then there exists an ordinal α such that $|X| < |\alpha|$.*

We can make use of these equivalences in order to show that every set has a cardinality:

Remark. *Using the well-ordering principle, and therefore* AC*, we can show that every set is in bijection with a unique cardinal: let X be a set. Now X can be well-ordered by $\prec$, say. Hence $\operatorname{type}(\langle X, \prec \rangle) = \alpha$ for some ordinal α (the order type exists, in particular). Using Hartogs' lemma, and using the fact that the class of ordinals is well-ordered, we can find the least ordinal λ that is in bijection with α. Now λ is a cardinal by definition 2.9, and hence it is the cardinality of X. Also, by definition 2.9, we must have that the cardinality is unique.*

Also, after introducing the cardinality operation, we would like to stress the following point:

Remark. *One can see here that the concept of ordinals is a refinement of cardinality. Consider two well-ordered sets $\langle X, \prec \rangle$ and $\langle Y, \prec \rangle$. Then $\operatorname{type}(X, \prec) = \operatorname{type}(Y, \prec)$ implies $|X| = |Y|$, but $|X| = |Y|$ does not imply $\operatorname{type}(X, \prec) = \operatorname{type}(Y, \prec)$. A simple example is given by the sets ω and $\omega + 1$ with the natural ordering of ordinals (they clearly have the same cardinality but not the same order type).*

Definition 2.12. Let λ and κ be cardinals. Then we define ${}^{\lambda}\kappa$ to be the set of all functions from λ to κ. The cardinality of ${}^{\lambda}\kappa$ is defined to be κ^{λ}.

Remark. *Note that our definition of cardinal arithmetic just given obeys the properties* $\kappa^0 = 1$ *and* $0^\kappa = 0$ *for all non-zero cardinals* κ*.*

This first section is concluded by noting one more crucial convention (it is assumed that the reader is familiar with the notions of limit and successor ordinals):

Remark. *We do not consider the zero-ordinal a limit ordinal. That is, we consistently use the trichotomy property of ordinal types: an ordinal* α *is either a successor, a limit, or the zero ordinal.*

2.2 Constructing Cardinals and Cardinal Arithmetic

As the concept of successors is vital in the class of ordinals, the naturally arising question is whether we can extend the notion of successors to the subclass of cardinals; it turns out that we can indeed:

Definition 2.13. Let κ be a cardinal. Then we define the successor of κ by

$$\kappa^+ = \min\{\lambda \in \mathrm{Card}\colon \kappa < \lambda\}.$$

Remark. *Formally, we need to verify that the cardinal successor always exists. We can go about doing so as follows: we know that the collection of cardinals forms a proper class and is therefore infinite. We will appeal to Hartogs' lemma: if* $\kappa \in \mathrm{Card}$ *then there exists an ordinal* α *such that* $|\alpha| > \kappa$*. Consider the set (note that this is indeed a set)*

$$K = \{\beta \in Ord : \kappa < |\beta| \leq |\alpha|\}.$$

Note that $\alpha \in K$ *and hence* K *is non-empty. Further,* K *is a set of ordinals and hence well-ordered. Thus we may define*

$$\kappa^+ := |\min(K)|$$

which provides us with the cardinal successor. Further, it follows directly that there is no cardinal between κ *and* κ^+*, as required.*

In section 2, we defined the notions of ordinals and cardinals and gave a few examples. Now, we will recap a vital notion of cardinals that can be constructed using transfinite induction and recursion.

Definition 2.14. We can define the class of cardinals using transfinite recursion as follows:

- The base case is given by $\omega_0 = \omega$.
- Let α be an ordinal. Using transfinite recursion we then define ω_α by

$$\omega_\alpha = \min\{\lambda\colon \lambda \text{ is a cardinal} \wedge \forall\beta < \alpha(\omega_\beta < \lambda)\}.$$

This gives rise to a well-ordered class of infinite cardinals. Crucially, it contains all the infinite cardinals, and uses definition 2.13 (it is clear that $(\omega_\alpha)^+ = \omega_{\alpha+1}$ for any α, and $\omega_\beta = \sup\{\omega_\gamma : \gamma < \beta\}$, for any limit ordinal β).

Remark. *According to our definition, we may identify the ordinal* ω_α *with the cardinal* $\aleph_\alpha$*. We will choose the notation depending on the context, and will therefore not resort to using only one. In detail, whenever the ordering is necessary, we will write* ω*. If we refer to the cardinality only, we will use alephs.*

The hypothesis that $\aleph_1 = 2^{\aleph_0}$ is known as the *Continuum Hypothesis*. Generalising yields the *Generalised Continuum Hypothesis* that states that for all $\alpha \in \mathbf{ON}$ we have $\aleph_{\alpha+1} = 2^{\aleph_\alpha}$.

Remark. *We note that for ordinals α and β, definition 2.14 yields*

$$\alpha \leq \beta \Rightarrow \omega_\alpha \leq \omega_\beta,$$

with equality if and only if $\alpha = \beta$.

The following result is of utmost importance (the proof is omitted, details can be found in [HH99, p. 80]):

Theorem 2.15 (Fundamental Theorem of Cardinal Arithmetic)**.** *If κ is an infinite cardinal, then $\kappa^2 = \kappa$.*

We prefer to use the following corollary for its simpler applicability. The proof is immediate from theorem 2.15.

Corollary 2.16. *Let $\lambda, \kappa \in \mathrm{Card}$. Then*

$$\lambda + \kappa = \max(\lambda, \kappa)$$

whenever $\aleph_0 \leq \max(\lambda, \kappa)$. If we also have $0 < \min(\lambda, \kappa)$, then

$$\lambda \cdot \kappa = \max(\lambda, \kappa).$$

Further, it is well-known that usual integer addition and multiplication, and addition and multiplication of finite cardinals coincide.

The following two definitions will be immensely useful when considering cardinal arithmetic (both can be found in [Kun80, p. 45]):

Definition 2.17. Let α be a non-zero ordinal and consider the set system $\{\kappa_\xi : \xi < \alpha\}$, in which every κ_ξ is a cardinal. Then we define the sum and product of all κ_ξ to equal

$$\begin{aligned}\sum_{\xi<\alpha} \kappa_\xi &= \left|\bigcup\{\{\xi\} \times \kappa_\xi : \xi < \alpha\}\right| \\ &= \left|\bigcup_{\xi<\alpha} \{\xi\} \times \kappa_\xi\right|\end{aligned}$$

and

$$\prod_{\xi<\alpha} \kappa_\xi = |\{f : \mathrm{dom}(f) = \alpha \text{ and } \forall \xi \in \alpha (f(\xi) \in \kappa_\xi)\}|$$

respectively.

We would like to make one final remark:

Remark. *Although this report covers set theory, we will not be too busy working with the axioms of* ZFC *directly. We rather take most consequences of* ZFC *for granted, and refer to them occasionally.*

3 Combinatorial Set Theory

This section covers multiple topics in combinatorial set theory. we begin by introducing *cofinality* which will allow us to obtain a deeper understanding of ordinals and their characteristics. Cofinality will enable us to obtain further inequalities of cardinal exponentiation, which in turn allow us to characterise cardinals in a particular way indicating its internal structure (*regular and singular cardinals* will play a major role throughout this and the subsequent sections).

In section 4, we then go on to define several types of constructions of set systems, each being interesting mathematically in its own right. After defining *filters and ideals* we introduce a crucial notion of combinatorial set theory: the concept of *stationary sets* will provide us with powerful tools with which we shall use to prove seminal results in combinatorial set theory that combine ideas of regular and singular cardinals as well stationary sets (such as *Fodor's theorem*). The section will be concluded by *Silver's theorem*, whose combinatorial proof will relate the Generalised Continuum Hypothesis to the aforementioned theory.

3.1 Cofinality

As introduced in the first section, we have a good understanding of ordinals and cardinals and how they are interlaced. However, it would be of use to know more about ordinals with respect to their intrinsic ordering.

Take $\alpha \in \mathbf{ON}$. For what ordinals β exist subsets of that order type which are not bounded in α? Without making this idea rigorous, we can see immediately that α is unbounded in α. But are there any subsets of order type strictly less than α which are also unbounded?

The notion of cofinality introduced in this subsection will help us make sense of how different subsets of ordinals (and hence indeed of any well-ordered set) behave under certain operations concerning unboundedness. We begin by stating the following definition:

Definition 3.1. Let $\langle A, \prec \rangle$ be an ordered set, and let $B \subset A$. Then we call B a cofinal subset of A (for brevity we might say B is cofinal in A) if

$$\forall x \in A \, \exists y \in B (x \preceq y). \tag{$*$}$$

We call $(*)$ the cofinality property.

An informal description could be the following: a cofinal subset is some type of subset of "trump cards" compared to the underlying set; every element in the underlying set can be trumped (or at least equalled) by an element of the cofinal subset.

Remark. *Assume the order type of A is a limit ordinal. In the literature, subsets B of A that have the cofinality property are also called unbounded sets in A. Similarly, subsets that do not satisfy the cofinality property may be called bounded sets in A.*

As we will almost exclusively consider the case of limit ordinals, we will use the terms cofinal and unbounded interchangeably.

Example 1. *It immediately follows that A is cofinal in A, for any $x \in A$ satisfies the cofinality property itself.*

Example 2. *We can see that ω is cofinal in $\langle \mathbb{Q}, < \rangle$ by the following: let $x \in \mathbb{Q}$. Consider $y = [x] + 1 \in \omega$, where $[\cdot]$ denotes the integer part function. As such an element y exists for any $x \in \mathbb{Q}$, we see that ω satisfies the cofinality property in $\mathbb{Q}$, as required.*

It is clear that the cofinality property behaves well on sets that are both subsets of the underlying set and supersets of the respective cofinal subset. The following result gives a rigorous representation of this result:

Lemma 3.2. *Let $\langle A, \prec \rangle$ be an ordered set. If $B \subset A$ is cofinal in A, then so is every $C \subset A$ for which $B \subset C \subset A$.*

The proof is straightforward:

Proof. By definition 3.1 and example 1, we can deduce that the cases $B = C$ and $A = C$ are trivial. Hence, assume that $B \subset C \subset A$, all subsets being proper. Let $x \in A$. By cofinality of B, there exists $y \in B$ such that $x \preceq y$. As $B \subset C$, we have that $y \in C$, and hence C is cofinal in A. □

Example 3. *Let $\Omega = \{\omega\}$. Then Ω is cofinal in $\omega + 1$ with respect to the well-ordering $\in_{\omega+1}$. Indeed, take any element $a \in \omega$, then the condition is fulfilled automatically since $a \in \omega$ (this is the ω being an element of Ω) and $\omega \in \Omega$. Similarly, take ω itself, then the condition is fulfilled by virtue of definition 3.1 and by $\leq$ allowing equality and hence the trichotomy property of ordinals.*

For the next example, the following notation is introduced:

Definition 3.3. Let Γ be a set of ordinals. We write $\mathrm{Lim}(\Gamma)$ to denote the set of all limit ordinals contained in Γ. That is,

$$\mathrm{Lim}(\Gamma) = \{\beta \in \Gamma : \beta \text{ is a limit ordinal}\}.$$

Example 4. *Take the ordinal ω_1 and consider $\mathrm{Lim}(\omega_1)$. We will show that $\mathrm{Lim}(\omega_1)$ is unbounded in ω_1. Fix any infinite $\alpha \in \omega_1$ (in the case of finite α the result is trivial). Now consider the set $J := \mathrm{Lim}(\omega_1 \setminus \alpha)$.*

Claim: The set J is non-empty.
Proof: Assume it is empty, then $\omega_1 = \sup(\{\alpha + 1, \alpha + 2, \ldots\})$. But as α is an element of ω_1, it is clear that α is countable. Hence ω_1 is a countable union of countable sets, and therefore also countable. Contradiction. ■

Now pick the ordinal $\min(J)$ to verify the cofinality.

Remark. *Of course, ω_1 is not special in the example above. In fact, for any uncountable cardinal κ, the set $\mathrm{Lim}(\kappa)$ is cofinal in κ. The proof is identical to the previous example with κ substituted for ω_1. (When we consider the step $\kappa = \sup(\{\alpha+1, \alpha+2, \ldots\}) = \alpha+\omega$, then note that $\kappa = |\alpha + \omega| = |\alpha| < \kappa$, which yields the contradiction.)*

Using the previous results, we would like to relate the order type to the cardinality of a set using the following result, which will be useful in future proofs:

Proposition 3.4. *Let A be a set. Then $|A| \leq \mathrm{type}(A, \prec)$ for any well-ordering $\prec$ on A.*

Proof. By definition 2.10, $|A|$ is the least ordinal which is in bijection with the set A, and as an order-isomorphism is, in particular, a bijection, we obtain a contradiction otherwise. □

In order to progress, however, we need this very important result:

Theorem 3.5 (Hausdorff Cofinality Theorem)**.** *Let $\langle A, \prec \rangle$ be an ordered set. Then there exists a cofinal subset $B \subset A$ such that B is well-ordered and $\mathrm{type}(B, \prec) \leq |A|$.*

This proof is taken from [HH99, pp. 85-6], further explanations have been added accordingly. The reader's attention is drawn to the fact that this proof requires AC.

Proof. We begin by considering a well-ordering $\prec_1$ of A such that $\text{type}(A, \prec_1) = |A|$. Such a well-ordering exists by AC. Now define B by

$$B = \{y \in A : \forall z \in A(y \prec z \Rightarrow y \prec_1 z)\} \subset A.$$

We will show that B is non-empty.

Fix $x \in A$ and consider

$$y = \min_{\prec_1}\{z \in A : x \preceq z\}.$$

Such an element y exists since $\prec_1$ is a well-ordering. Note that y depends on x, and that $y \geq x$. Hence, if we show that $y \in B$, then we have shown that B is cofinal in $\langle A, \prec\rangle$.

We verify the definition in order to show that y is an element of B: take any $z \in A$ and suppose $y \prec z$. Now $x \prec z$. To show this, assume for contradiction that $z \in A$, $y \prec z$ and $z \preceq x$. By the definition of y we have that $x \preceq y$. Hence

$$(z \preceq x \wedge x \preceq y) \Rightarrow z \preceq y,$$

which contradicts our assumption that $y \prec z$. Hence $x \prec z$. Moreover, $x \prec z$ implies that $y \prec_1 z$. We can show this using the following: assume $z \preceq_1 y$, then

$$z \preceq_1 \min_{\prec_1}\{z \in A : x \preceq z\} = y$$

and hence we must have that $z \prec x$ (if $z \in \{w \in A : x \preceq w\}$, then $z = y$, but we cannot have equality since we assumed $y \prec z$ above). However, this contradicts that $x \prec z$. Therefore $y \prec_1 z$ and so $y \in B$ and hence B is non-empty. By the remark above, we are done.

In order to finish the proof, we are required to show that the orderings $\prec$ and $\prec_1$ coincide on the set B. Consider $x, y \in B$ with $x \preceq y$. Then the definition of B immediately yields that $x \preceq_1 y$. For the other direction, consider $w, z \in B$ with $w \preceq_1 z$. For contradiction, assume that $z \prec w$. But w and z are elements of B, and hence $z \prec w \Rightarrow z \prec_1 w$ by the definition of B. Contradiction. Therefore, the orderings $\prec$ and $\prec_1$ coincide on B and hence $\prec$ well-orders B. Therefore

$$\text{type}(B, \prec) = \text{type}(B, \prec_1) \leq \text{type}(A, \prec_1) = |A|,$$

where the first equality follows from the reasoning above and the inequality follows since B is a subset of A. □

Remark. *Note that the ordering of B is the same as the ordering of A; this is crucial and renders the theorem so useful. (Clearly, using* AC*, we could find a well-ordering on B without the above axiom. However, as there is no guarantee that the given well-ordering preserves the original ordering of A, we would not gain any insight into the cofinality of A.)*

Theorem 3.5 is a very powerful and useful result. It will be frequently applied throughout the following sections.

In order to illustrate its versatility, take any ordered set $\langle A, \prec\rangle$. We may now assume the existence of a cofinal well-ordered subset of A and we can determine its order type.

As the class of ordinals is linearly ordered, we may consider the set of cofinal well-ordered subsets of A and order it by their respective order types (an upper bound is given by $|A|$ as seen in example 1). Using the fact that the class of ordinals is well-ordered allows us to go one step further: we can now determine the least ordinal α for which there is a cofinal subset of A with order type α. It is clear that this ordinal must be unique.

This idea gives rise to the following definition:

Definition 3.6. Let $\langle A, \prec \rangle$ be an ordered set and let α be an ordinal. Then α is said to be the cofinality of A if it is the least ordinal for which $\langle A, \prec \rangle$ has a well-ordered subset of order type α which is cofinal in A. The cofinality of A with respect to $\prec$ is denoted by $\mathrm{cf}(\langle A, \prec \rangle)$.

When introducing the notion of cofinality in definition 3.1, we mentioned that another term used frequently for what we call the cofinality property is the word *unbounded*. This is what the reader should think of when considering the cofinality of an ordered set: the cofinal subset is not bounded by any element in the original set. Note, however, that cofinal subsets of successor ordinals are indeed bounded. Hence this notation only applies to limit ordinals. (As we shall not spend too much time considering successor ordinals, we will almost always be able to use the terms *cofinal* and *unbounded* interchangeably.)

Remark. *As done previously, we omit the ordering if it is implicit and simply write* $\mathrm{cf}(A)$ *in such cases. Further it is useful to note the following: if α is an ordinal and $A \subset \alpha$, then A is cofinal in α if and only if* $\sup(A) = \alpha$.

As always, we would like to consider ordinals only. That is, rather than working with sets and their respective well-orders, dealing with their ordinal representatives (i.e. their order types) is much more convenient. The following remark allows us to do so:

Remark. *If f is an order-isomorphism between ordered sets $\langle A, \prec \rangle$ and $\langle A', \prec' \rangle$, we obtain that their cofinalities must coincide. If they did not, definition 3.6 yields that f is not a bijection, which contradicts the isomorphism-property of f. Therefore the cofinality operation is well-defined. (Informally, we may say that cofinality "survives" order-isomorphisms, which indeed follows immediately from the definition.)*

As well-ordered sets are in particular linearly ordered, we can now deduce the following: whenever working with a well-ordered set, $\langle X, \prec \rangle$ say, and aiming to determine its cofinality, we are allowed to choose a different representative from its order-type equivalence class. We will choose the unique ordinal that is order-isomorphic to $\langle X, \prec \rangle$.

Remark. *Another way of looking at this definition is the following: take any linearly ordered set $\langle A, \prec \rangle$. We wold like to find the least ordinal α for which there exists a cofinal well-ordered subset $\langle B, \prec \rangle$ with order type α. We look at all cofinal well-ordered subsets of A with the same ordering $\prec$, and hence consider the set*

$$\mathrm{CF}(A, \prec) = \{B \subset A : \langle B, \prec \rangle \text{ is cofinal and well-ordered in } A\}.$$

By Hausdorff's Cofinality Theorem, the set $\mathrm{CF}(A, \prec)$ *is non-empty. Now we may consider the order type of every set $B \in CF(A, \prec)$, and hence look at*

$$\mathrm{TYPE}(A, \prec) = \{\mathrm{type}(B, \prec) : B \in \mathrm{CF}(A, \prec)\}.$$

By our definition, it follows directly that the cofinality of $\langle A, \prec \rangle$ equals the least element of $\mathrm{TYPE}(A, \prec)$. *Hence we define*

$$\mathrm{cf}(\langle A, \prec \rangle) = \min(\mathrm{TYPE}(A, \prec)),$$

which yields an equivalent definition, as required.

As we have shown that the cofinality operation is well-defined and since theorem 3.5 gives us a well-ordered cofinal subset of any ordered set, the following result is quite handy:

Proposition 3.7. *Let $\langle A, \prec \rangle$ be an ordered set and suppose $B \subset A$. If B is cofinal in A, then* $\mathrm{cf}(A) = \mathrm{cf}(B)$.

Proof. Note that B is a subset of A and B is cofinal in A, and thus $\mathrm{cf}(B) \leq \mathrm{cf}(A)$, for any cofinal subset in B is also cofinal in A.

Aiming for a contradiction, suppose $\mathrm{cf}(B) < \mathrm{cf}(A)$. That is, there exists $C \subset B$ (where C is well-ordered and cofinal) such that $\mathrm{type}(C, \prec) = \mathrm{cf}(B) < \mathrm{cf}(A)$. Note that the existence of such a subset C is guaranteed by definition of cofinality and the remark above. But if C is cofinal in B and B is cofinal in A, then C is cofinal in A. Thus

$$\mathrm{type}(C) = \mathrm{cf}(B) < \mathrm{cf}(A) \leq \mathrm{type}(C),$$

where the last inequality follows since the order type of C constitutes an upper bound for the cofinality of A. This provides the required contradiction.

Thus $\mathrm{cf}(A) = \mathrm{cf}(B)$, as required. □

Remark. *This result is crucial and very useful as it allows the following reasoning: whenever we aim to calculate the cofinality of a linearly ordered set, we may now consider any well-ordered cofinal subset (whose existence is guaranteed by theorem 3.5), determine its order type, and calculate the cofinality of the order type instead.*

Hence we can deduce the following:

Theorem 3.8. *Let $\langle X, \prec \rangle$ be an ordered set. Suppose that A is a cofinal well-ordered subset of X, where* $\mathrm{type}(A, \prec) = \alpha$. *Then* $\mathrm{cf}(\langle X, \prec \rangle) = \mathrm{cf}(\alpha)$.

Proof. By proposition 3.7, the cofinalities of X and A coincide. By our previous remark and the order isomorphism from any well-ordered set to its order type, the cofinality of A and α coincide. □

We can use the previous remark to give the following alternative definition of cofinal subsets (showing equivalence is straightforward):

Definition 3.9. Let α be an ordinal. Then the increasing ordinal sequence $(\gamma_\nu)_{\nu \in \theta}$ is called <u>cofinal in α</u> if the set $\{\gamma_\nu : \nu \in \theta\} \subset \alpha$ has the cofinality property. The least ordinal θ for which such a cofinal sequence exists is called the <u>cofinality of α</u>, denoted by $\mathrm{cf}(\alpha)$.

We will use both definition 3.1 and definition 3.9 depending on which is more intuitive in the respective context.

Remark. *From definition 3.9 it is straightforward to show that a sequence $(\gamma_\nu)_{\nu \in \theta}$ is cofinal in α if and only if*

$$\bigcup_{\nu \in \theta} \gamma_\nu = \alpha,$$

as can be found in [Dev93, p. 88]. The proof is simple:

($\Rightarrow$): If $(\gamma_\nu)_{\nu \in \theta}$ is cofinal in α, then, by definition, for any $\beta \in \alpha$ there is a $\theta' \in \theta$ such that $\beta \leq \gamma_{\theta'}$. Thus $\beta \subset \gamma_{\theta'}$. As β was chosen arbitrarily and since α is an ordinal we obtain

$$\bigcup_{\nu \in \theta} \gamma_\nu \supset \alpha.$$

From definition 3.9 we know that $\gamma_\nu \in \alpha$ for any $\nu \in \theta$, and so

$$\bigcup_{\nu \in \theta} \gamma_\nu \subset \alpha,$$

as required.

($\Leftarrow$): Follows immediately from definition 3.9.

The remaining part of this section is devoted to useful (and hopefully interesting) results regarding cofinalities of sets and ordinals. In order to further develop the notion of cofinality, we need examples and intuition on how the cofinality operation behaves on different order types; this is what this subsection will do. And indeed, as we will see, there are connections between the ordinal types and their respective cofinalities.

Eventually, we will derive a few results which will make the calculations of cofinalities of arbitrary ordered sets much easier. But for now, we will try and find them "by hand", that is, by finding a suitable unbounded sequence of ordinals and proving its minimality.

We begin with the following:

Theorem 3.10. *Let $\langle A, \prec \rangle$ be an ordered set. Then*

$$\operatorname{cf}(\langle A, \prec \rangle) = 1 \Leftrightarrow \langle A, \prec \rangle \textit{ has a last element.}$$

Proof.

($\Rightarrow$): If $\operatorname{cf}(\langle A, \prec \rangle) = 1$ then there is a well-ordered cofinal subset $B \subset A$ with $\operatorname{type}(B) = 1$. Hence B is a singleton. Assume that $B = \{b\}$ is cofinal in $\langle A, \prec \rangle$. But we assumed that A does not have a last element, hence there exists $c \in A$ for which $b \prec c$, a contradiction to the cofinality property of B.

($\Leftarrow$): Assume $\langle A, \prec \rangle$ has a last element, b say. Then b trivially satisfies the cofinality property. Hence the singleton $\{b\}$ is cofinal in A, and $\operatorname{type}(\{b\}, \prec) = |\{b\}| = 1$, as required. ☐

We can use our newly acquired knowledge on successor ordinals as they, by definition, contain a last element. The following corollary follows directly:

Corollary 3.11. *Let $\langle A, \prec \rangle$ be an ordered set. If $\operatorname{type}(A, \prec) = \alpha + 1$, that is, a successor ordinal, then $\operatorname{cf}(\langle A, \prec \rangle) = 1$.*

Proof. Assume f is the order-isomorphism from $\langle A, \prec \rangle$ to $\langle \alpha + 1, < \rangle$. Then take $B = f^{-1}(\alpha)$ and the result follows immediately from the proof of theorem 3.10. ☐

Remark. *The two results above visualise the difference between the ideas of bounded, unbounded, and cofinal subsets. It is clear now that every subset of a limit ordinal is cofinal if and only if it is unbounded. For successor ordinals, however, a singleton (which is of course bounded) suffices. Due to the trivial nature of cofinality in successor ordinals, we will direct our attention to limit ordinals from now on.*

After finding a connection between successor ordinals and their cofinality, the naturally arising question is whether there are similar relationships with limit ordinals. We would expect such relationships to be more complicated since limit ordinals involve the notion of supremum and hence cannot be decomposed as easily as successor ordinals. However, the trichotomy property of ordinal types notably simplifies the proofs. As an example, we begin with the following result, which will help us significantly in further proofs:

Proposition 3.12. *Let α be an ordinal. If α is a limit ordinal, then* $\mathrm{cf}(\alpha)$ *is also a limit ordinal.*

We will go on and prove this in due course, but we require one further result:

Proposition 3.13. *Let α be an ordinal. Then*

$$\alpha = 0 \Leftrightarrow \mathrm{cf}(\alpha) = 0.$$

Proof.
($\Rightarrow$) By definition 3.1, $\mathrm{cf}(0) \subset 0$, and as 0 is the least ordinal, the only possible cofinality is 0 itself. Clearly, as 0 has no elements, the condition in definition 3.1 is satisfied automatically and the result follows.

($\Leftarrow$) We argue for a contradiction, so assume not. Then $\mathrm{cf}(\alpha) = 0$ and $\alpha > 0$. Thus there is a cofinal well-ordered subset $B \subset \alpha$ which satisfies $\mathrm{type}(B) = 0$. Hence $B = 0$. Thus, by definition of cofinality,

$$\forall \gamma \in \alpha(\gamma \leq 0),$$

and so $\gamma = 0$ for all γ less than α. As we assumed α to be non-empty we have $\alpha = \{0\} = 1$. But 1 is a successor ordinal, and so corollary 3.11 yields a contradiction. □

We can now continue and prove proposition 3.12:

Proof of proposition 3.12. We aim for a contradiction, hence we assume otherwise. By the trichotomy property of ordinal types, we have either $\mathrm{cf}(\alpha) = 0$ or $\mathrm{cf}(\alpha) = \beta + 1$ for some ordinal β. In the former case, proposition 3.13 implies that $\alpha = 0$, a contradiction. In the latter case, there exists a cofinal subset $\Gamma \subset \alpha$ (which is clearly well-ordered as it is a set of ordinals) such that $\mathrm{type}(\Gamma) = \mathrm{cf}(\alpha) = \beta + 1$. Now Γ has a last element, namely $\{\beta\}$. But α is a limit ordinal, hence it does not have a last element. Thus Γ cannot be cofinal in α. Contradiction. □

Corollary 3.14. $\mathrm{cf}(\omega) = \omega$*.*

Proof. We know that ω is the least limit ordinal, and hence the result follows by proposition 3.12. □

What can we say about the converse of proposition 3.12? As can be shown easily, this holds as well:

Proposition 3.15. *For any ordinal α, if* $\mathrm{cf}(\alpha)$ *is a limit ordinal, then so is α.*

Proof. By proposition 3.13, we know that $\alpha > 0$. If α is a successor ordinal, then $\mathrm{cf}(\alpha) = 1$ by theorem 3.10. □

Hence we can reformulate and state the following theorem:

Theorem 3.16. *Let α be an ordinal. Then*

$$\alpha \textit{ is a limit ordinal} \Leftrightarrow \mathrm{cf}(\alpha) \textit{ is a limit ordinal.}$$

Remark. *It follows directly from theorem 3.10, proposition 3.12, and from proposition 3.13 that there exists an ordinal α for which there is no set A with* $\mathrm{cf}(\langle A, \prec\rangle) = \alpha$*; If $\alpha \neq 1$ is a successor ordinal, then there is no such set since* $\mathrm{cf}(\langle A, \prec\rangle) = 1$ *whenever A has a last element, and* $\mathrm{cf}(\langle A, \prec\rangle)$ *is a limit ordinal whenever A is order-isomorphic to a limit ordinal.*

We have now found a couple of results on ordinals, but we have merely acknowledged the subclass of cardinals whilst discussing cofinalities. To put this right, we will now deduce some rather interesting results about cardinals in combination with the cofinality property and our previously derived results.

One more corollary to proposition 3.12 will be useful (the proof is immediate from proposition 3.12):

Corollary 3.17. *Let α be a limit ordinal. Then* $\mathrm{cf}(\alpha) \geq \omega$.

When dealing with infinite cardinals, we will make frequent use of definition 2.14. As it turns out, we can reduce many statements about infinite cardinals, i.e. a cardinal of the form ω_α, to a problem about α, and then use what we have derived previously.
The following example (which can be found in [HH99, p. 86]) illustrates this point:

Example 5. *We want to show that* $\mathrm{cf}(\omega_\omega) = \omega$. *In order to do so, it suffices to note that* $\omega_\omega = \sup\{\omega_n : n \in \omega\}$. *Hence the set* $\{\omega_n : n \in \omega\}$ *is cofinal in* ω_ω. *As* ω *is the least limit ordinal, the result follows immediately.*

Another example of a calculation of cofinality follows.

Lemma 3.18. *Let k be a finite cardinal. Then* $\mathrm{cf}(\omega \times k) = \omega$.

Proof. Consider the subset $\Gamma = \{\omega \times (k-1) + n : n \in \omega\}$. This is cofinal in $\omega \times k$, and we will show this below.

The set Γ is obviously a subset of $\omega \times k$, hence we only prove the cofinality property: Let $\alpha \in \omega \times k$. Then $\alpha = \omega \times m + n$, where $0 \leq m < k$ and $n \in \omega$ (this holds since $\omega \times k$ is a limit ordinal). If $m < k-1$, then $\omega \times (m+1) \leq \omega \times (k-1) \in \Gamma$ works; If $m = k-1$, take $\omega \times (k-1) + (n+1)$, and hence Γ satisfies the cofinality property. We note that $\mathrm{type}(\Gamma) = \omega$, and hence, by proposition 3.12 and the fact that ω is the least limit ordinal, the result follows. □

Remark. *This previous proof was included for illustrative purposes only. Note that the fact that* $\omega \times (n+1) = \sup\{\omega \times n + k : k \in \omega\}$ *implies* $\mathrm{cf}(\omega \times (n+1)) = \omega$ *immediately, as required.*

So far, we have mostly been working with specific examples of ordinals; in the majority of cases, we have calculated the cofinality of such ordinals explicitly. In theorem 3.16, we have shown that an ordinal is a limit ordinal if and only if its cofinality is also a limit ordinal. The examples we examined then painted a similar picture that allowed us to go even one step further: the cofinalities of the limit ordinals we considered turned out to be cardinals. If we could say more about the cofinality of ordinals, and especially about their order type, we could make finding cofinalities significantly simpler.

The following theorem will make rigorous what we have suspected already:

Theorem 3.19. *Let α be an ordinal. Then* $\mathrm{cf}(\alpha)$ *is a cardinal.*

Proof. If α is finite or an infinite successor ordinal, then $\mathrm{cf}(\alpha) = 1$, which is obviously a cardinal. Now assume that α is a limit ordinal, and further assume that $\mathrm{cf}(\alpha) = \theta$, where θ is not a cardinal, i.e. $|\theta| < \theta$. Then there is a sequence $(\gamma_\beta)_{\beta\in\theta}$ that is cofinal in λ. However, according to the definition, there exists a bijection f from $|\theta|$ to θ. Hence the sequence $(\gamma_{f(\mu)})_{\mu\in|\theta|}$ is also cofinal in λ. This contradicts the minimality of θ. □

Theorem 3.19 is a very strong result: knowing that the cofinality of any ordinal is a cardinal allows us to draw our attention to cardinals and their respective properties and operations only (cardinal arithmetic, for instance). We will see in later sections what this means practically.

However, there is one crucial point to remember: by definition, the cofinality of any ordinal is the least *ordinal* α, for which there exists a cofinal subset of *order type* α. We still refer to the order type, *not* the cardinality. (Informally, when considering the cofinality, we are still interested in the "ordering" of any cofinal sequence, not just the "size".) We will see in an example later why loosening the requirement from order-isomorphism to bijections must be avoided by all means. (See proposition 4.27 and its related remark.)

Corollary 3.20. *Every countable limit ordinal has cofinality* ω.

The next result constitutes an interesting connection between the ordinals and the infinite cardinals (this problem can be found in [HH99, p. 91 (3)]):

Proposition 3.21. *Assume* α *is a limit ordinal. Then* $\mathrm{cf}(\omega_\alpha) = \mathrm{cf}(\alpha)$.

We will need a few more tools to proceed with the proof:

Lemma 3.22. *Let* α *be an ordinal. Then the following three statements hold:*

(1): $\mathrm{cf}(\omega_\alpha)$ *is a limit ordinal.*

(2): $\alpha \leq \omega_\alpha$

(3): $\mathrm{cf}(\alpha) \leq \mathrm{cf}(\omega_\alpha)$

Proof.

(1): This proof is given by simply quoting a sequence of results we have proved before: we know that ω_α is a cardinal and hence a limit ordinal, and thus proposition 3.12 yields the result.

(2): Assume otherwise, then there exists $\alpha \in \mathbf{ON}$ such that $\omega_\alpha < \alpha$. Take the least such ordinal and denote it by α'. However, definition 2.14 yields that whenever γ and δ are ordinals such that $\gamma < \delta$ then $\omega_\gamma < \omega_\delta$. This implies that $\omega_{\omega_{\alpha'}} < \omega_{\alpha'} < \alpha'$, which contradicts the minimality of α'.

(3): If α is a successor, then $\mathrm{cf}(\alpha) = 1$, but (1) implies that $\mathrm{cf}(\omega_\alpha)$ is a limit ordinal, which provides the required inequality.
If α is a limit ordinal, then note that by the construction of every infinite cardinal (using definition 2.14) we can find a cofinal sequence in ω_α comprising only infinite cardinals. Let $(\omega_\beta)_{\beta \in \mathrm{cf}(\omega_\alpha)}$ be such a sequence. For this sequence to be unbounded in ω_α, we must have that $\beta < \alpha$ for all $\beta \in \mathrm{cf}(\omega_\alpha)$. As $\omega_\alpha = \sup\{\omega_\delta : \delta < \alpha\}$ it follows directly that the sequence $(\beta)_{\beta \in \mathrm{cf}(\omega_\alpha)}$ is cofinal in α, and so $\mathrm{cf}(\alpha) \leq \mathrm{cf}(\omega_\alpha)$ as required.

□

We now have everything we need to prove proposition 3.21:

Proof of Proposition 3.21. Assume $\mathrm{cf}(\alpha) = \kappa$. Theorem 3.19 implies that κ must be a cardinal, and by theorem 3.16 we know that κ is infinite. By definition 3.9 there exists a cofinal sequence $(\gamma_\nu)_{\nu \in \kappa}$ in α. Now, using definition 2.14, the sequence $(\omega_{\gamma_\nu})_{\nu \in \kappa}$ is cofinal in ω_α. Thus $\mathrm{cf}(\omega_\alpha) \leq \mathrm{cf}(\alpha)$.

Part (3) of lemma 3.22 now yields the required equality. □

Remark. *It is clear that the condition of α being a limit ordinal is necessary; if α were a successor ordinal,* $\mathrm{cf}(\alpha) = 1$*, but ω_α is a limit ordinal. Then proposition 3.12 gives a contradiction.*

3.2 Regular and Singular Cardinals

In the previous section, we have introduced the notion of cofinal subsets and the cofinality operation. We have seen a few examples alongside several interesting facts and relationships between ordinal types and cofinalities. However, we have discovered that some ordinals have the peculiar property of equalling their cofinality. All of those ordinals have something in common: they are in some sense "irreducible" with respect to their order type under the cofinality operation. Hence they deserve a closer look:

Definition 3.23. Let α be an ordinal. We call α a regular ordinal if

$$1 < \mathrm{cf}(\alpha) = \alpha.$$

We may abbreviate this by saying that α is regular.
We call α a singular ordinal, or, for short, singular, if

$$1 < \mathrm{cf}(\alpha) < \alpha.$$

Corollary 3.24. *Every regular ordinal is a cardinal.*

Proof. Follows immediately from theorem 3.19. □

We will see shortly that the distinction between regular and singular ordinals is vital when considering the structure of ordinals.

Example 6. *A few examples follow:*

- *ω is regular.*
- *For all $k \in \omega \setminus \{0, 1\}$, the ordinal $\omega \times k$ is singular.*
- *For all $\kappa \in \omega \setminus \{0\}$, the ordinal $\omega + k$ is neither regular nor singular. In fact every successor ordinal is neither regular nor singular as its cofinality is* 1*.*

If we employ our alternative definition of cofinality in 3.9 then we can give an elegant restatement of the definition of regular cardinals: a cardinal κ is regular if for every set system of cardinals $\{\kappa_\gamma : \gamma \in \delta\}$ (where $\delta \in \mathbf{ON}$) for which $\kappa_\gamma < \kappa$ and $|\delta| < \kappa$ then $|\bigcup_{\gamma\in\delta} \kappa_\gamma| < \kappa$. Informally, this says we cannot attain κ by considering fewer than κ-many sets that are all smaller in cardinality than κ.

This is indeed equivalent: assume the set system $\{\kappa_\gamma : \gamma \in \delta\}$ is as above. Then

$$\begin{aligned}
\left|\bigcup_{\gamma\in\delta} \kappa_\gamma\right| &\leq \left|\bigcup_{\gamma\in\delta} \{\gamma\} \times \kappa_\gamma\right| \\
&= \sum_{\gamma\in\delta} \kappa_\gamma \\
&\leq \prod_{\gamma\in\delta} \kappa_\gamma \\
&\leq |\delta| \cdot \kappa && (*) \\
&< \kappa \cdot \kappa && (**) \\
&= \kappa
\end{aligned}$$

where deduction $(*)$ holds since each κ_γ is smaller than κ, and step $(**)$ follows as $|\delta| < \kappa$. Hence *both* hypotheses must be satisfied. The reasoning is completed by application of the Fundamental Theorem of Cardinal Arithmetic.

Note that the first inequality holds via the injection

$$\bigcup_{\gamma\in\delta} \kappa_\gamma \to \bigcup_{\gamma\in\delta} \{\gamma\} \times \kappa_\gamma$$
$$\alpha \mapsto \langle \min\{\gamma \in \delta : \alpha \in \kappa_\gamma\}, \alpha\rangle$$

and the second inequality is true by virtue of the injection

$$\langle \alpha, \beta\rangle \mapsto \text{function } f \text{ with domain } \delta \text{ for which } f(\alpha) = \beta \text{ and } f(\alpha') = 0 \text{ for } \alpha' \neq \alpha,$$

assuming κ_γ is non-zero for all $\gamma \in \delta$. These two properties of cardinal arithmetic will be used later on.

The advantage of this definition is clear from the following examples (see [BP77, p. 108] for a similar discussion):

Example 7. *It is clear that $\aleph_{\omega+1}$ is regular as for any system of cardinals as described above the cardinality of any union will at most attain $\aleph_\omega$. Similarly, $\aleph_1$ is regular as, otherwise, it would be a countable union of countable sets and hence countable. This reasoning can be easily extended to the general case for $\aleph_{\alpha+1}$ for any ordinal α.*

The cardinal $\aleph_{\omega_1}$, however, is singular; by its definition, we have

$$\aleph_{\omega_1} = \sup_{\alpha\in\omega_1} \aleph_\alpha.$$

Clearly, $|\omega_1| = \aleph_1 < \aleph_{\omega_1}$ and further $\aleph_\alpha < \aleph_{\omega_1}$ for every $\alpha < \omega_1$, as required to verify singularity. Note that this agrees with our previously stated theorem 3.16, and furthermore provides an alternative proof to said result.

A few results linking this section to the idea of cofinalities follow. Among others, we will show the idempotence property of the cofinality operation, which follows next (its statement without proof can be found in [Dev93, p. 89], for example):

Theorem 3.25. *Let α be a limit ordinal. Then the cofinality operation is idempotent on α. That is,* $\mathrm{cf}(\alpha)$ *is a regular cardinal.*

Proof. It is clear by the previous result that $\mathrm{cf}(\alpha)$ is a cardinal. Hence only regularity remains to be proven.

We use theorem 3.19 and construct a contradiction: assume there exists an ordinal α such that

$$\mathrm{cf}(\mathrm{cf}(\alpha)) < \mathrm{cf}(\alpha) \leq \alpha,$$

where the rightmost inequality follows from the definition of cofinality. Hence there exists a cofinal subset $A \subset \alpha$ with $\mathrm{type}(A) = \mathrm{cf}(\alpha)$ (the well-ordering follows immediately). By our assumption, there exists a cofinal subset $B \subset A$ where $\mathrm{type}(B) = \mathrm{cf}(\mathrm{cf}(\alpha))$. Note that the existence of such subsets is guaranteed by the definition of cofinality.

Hence B is a subset of α and as it is cofinal in $\mathrm{cf}(\alpha)$, it is also cofinal in α. Thus

$$\mathrm{type}(B) = \mathrm{cf}(\mathrm{cf}(\alpha)) < \mathrm{cf}(\alpha) \leq \mathrm{type}(B),$$

a contradiction.

Hence $\mathrm{cf}(\alpha)$ is a regular cardinal, as required. □

We shall call this property the idempotence property of the cofinality operation.

Example 8. *Consider the cardinal $\aleph_{\omega_\omega}$. By our previous result, we know that* $\mathrm{cf}(\aleph_{\omega_\omega}) = \mathrm{cf}(\omega_\omega) = \mathrm{cf}(\omega) = \omega$. *Hence, one might informally say that $\aleph_{\omega_\omega}$ is "very singular". (We will investigate this property in the aside at the end of this section.)*

Remark. *Kunen gives a rather elegant proof of theorem 3.25 in [Kun80, pp. 32-3]. We say that f is a cofinal map from α into β (where $\alpha, \beta \in$ **ON**) if the image of f is cofinal in β. Using this notion we can define the cofinality of β as the least $\alpha \in$ **ON** such that there exists a function $f \colon \alpha \mapsto \beta$ whose image is cofinal in β. Indeed, it can be shown that there exists a cofinal strictly increasing map from* $\mathrm{cf}(\beta)$ *into β and that the existence of a cofinal strictly increasing map between any two limit ordinals implies equality of respective cofinalities. Then the result follows directly.*

3.2.1 Regular Limit Cardinals

When examining lemma 3.22, the reader might wonder whether there is a cardinal κ such that $\kappa = \omega_\kappa$. This is the same as saying that the $\aleph$-function has fixed points. (Note that such a cardinal κ must be infinite since $n < \omega_n$ for all $n \in \omega$.) The answer is affirmative; we will construct an example below which proves that the $\aleph$-function does indeed have fixed points. However, one might want to go one step further and ask whether it is possible for such a cardinal κ to be regular as well.

Recall the following: a cardinal ω_α is called a limit cardinal if α is a limit ordinal.

Now consider a limit cardinal ω_κ. We will examine what conditions will be satisfied provided ω_κ is regular. (Note that the limit cardinals we have examined so far, ω_ω for example, failed to be regular.)

Assume that ω_κ is a regular limit cardinal. Then

$$\mathrm{cf}(\omega_\kappa) = \omega_\kappa.$$

By our previous result, proposition 3.21, it follows that

$$\mathrm{cf}(\omega_\kappa) = \mathrm{cf}(\kappa).$$

Hence we may deduce the following: the fact that $\mathrm{cf}(\omega_\kappa) = \mathrm{cf}(\kappa)$ implies

$$\omega_\kappa = \mathrm{cf}(\kappa)$$

(by the fact that ω_κ is assumed to be regular). If we assume that κ is singular, then we obtain

$$\mathrm{cf}(\kappa) = \omega_\kappa < \kappa,$$

which is a contradiction. Hence, if ω_κ is a regular limit cardinal, then $\omega_\kappa = \kappa$, and is therefore an $\aleph$-fixed point.

Using the previous reasoning, one might wonder whether it is sufficient to assume that α is an $\aleph$-fixed point in order to obtain a regular limit cardinal. The following counterexample as presented in [Kun80, p. 34] explains why it does not suffice.

We define a sequence of cardinals as follows:

- Put $\sigma_0 = \omega$.
- For $n \in \omega$, define $\sigma_{n+1} = \omega_{\sigma_n}$.

- Now define $\kappa = \sup\{\sigma_n : n \in \omega\}$.

But now notice that

$$\begin{aligned}\kappa &= \sup\{\sigma_n : n \in \omega\} \\ &= \sup\{\sigma_{n+1} : n \in \omega\} \\ &= \sup\{\omega_{\sigma_n} : n \in \omega\} \\ &= \omega_{\sup\{\sigma_n : n \in \omega\}} \\ &= \omega_\kappa\end{aligned}$$

where the second line is clearly valid as $\sigma_0 = \omega < \sigma_1 = \omega_\omega$ and hence $\sup\{\sigma_n : n \in \omega\} > \sigma_0$. This proves the fixed point property of κ. In fact, it can be shown that κ is the least $\aleph$-fixed point. However, it is clear that the cofinality of κ equals ω; that is, κ is anything but regular. This failure turns out not to be an exception.

Regular limit cardinals are called weakly inaccessible. The notion, when considered in ZFC, comes with a caveat, though. It has been proven that, assuming ZFC is consistent, we cannot deduce the existence of weakly inaccessible cardinals in ZFC (cf. [Kun80, pp. 34, 177]). The field of so called Large Cardinals studies the properties and implications of Large Cardinal Axioms which postulate the existence of additional cardinals. More details and a historical introduction on this topic can be found in Kanamori's text [Kan03, p. 16].

As the nomenclature suggests, the notion of so called *strongly inaccessible cardinals* exists as well. We state it here for completeness:

Definition 3.26. An uncountable regular cardinal κ is called strongly inaccessible if $2^\lambda < \kappa$ for every cardinal $\lambda < \kappa$.

Strongly and weakly inaccessible cardinals can be likened to limit and successor cardinals: in the same way in which we cannot attain a limit cardinal by considering successor cardinals repeatedly, so is it impossible to reach strongly inaccessible cardinals by applying the continuum function to smaller cardinals. This justifies the choice of terminology "inaccessible".

It is clear that every strongly inaccessible cardinal is also weakly inaccessible: assume κ is strongly inaccessible. If κ were not a limit cardinal, then $\kappa = \omega_{\beta+1}$ for some ordinal β. But then Cantor's theorem implies that $2^{\omega_\beta} > \omega_\beta$. As $(\omega_\beta)^+ = \omega_{\beta+1}$, we obtain $2^{\omega_\beta} \geq \kappa$, which is a contradiction. Regularity was assumed.

We conclude this aside with an interesting result that combines our newly introduced notions with the Continuum Hypothesis:

Theorem 3.27. *When* GCH *is assumed, then the notions of weakly inaccessible and strongly inaccessible cardinals coincide.*

Proof. Above, we have already shown that every strongly inaccessible cardinal is also weakly inaccessible (this is a theorem of ZFC and hence also holds in ZFC + GCH).

Suppose κ is weakly inaccessible (i.e. κ is a regular limit cardinal). In view of a contradiction, suppose there exists a cardinal $\lambda < \kappa$ for which

$$2^\lambda \geq \kappa.$$

However, as we assumed GCH we have

$$2^\lambda = \lambda^+ \geq \kappa,$$

and thus it follows that

$$\lambda^+ = \kappa$$

as we cannot have $\lambda < \kappa$ and $\lambda^+ > \kappa$. This provides the required contradiction as it implies that κ is a successor cardinal, and so the theorem holds. □

3.3 Applications

After using cofinality in order to derive results in cardinal arithmetic, we shall now turn to a rather more exotic application of the theory. In the example below, we will apply the notion of cofinality to well-founded partially ordered sets. As it turns out, many of the results about cofinality on well-ordered sets can be adapted to partially ordered sets by altering the definitions slightly. At the end of this section, we will derive a combinatorial result that relates the cofinality of a partially ordered set to its antichains.

3.3.1 Partially Ordered Sets

In this subsection, we will use many of the results we have derived in the previous section.

Definition 3.28. Let $\langle P, \prec \rangle$ be a partially ordered set. Then $A \subset P$ is said to be a cofinal subset of P if

$$\forall p \in P\, \exists q \in A(p \preceq q).$$

The cofinality of P is defined to be the least cardinal κ for which there is a cofinal subset $Q \subset P$ with $|Q| = \kappa$.

Remark. *Note that, in this definition, we do not consider the order-type (which is not defined as P not linearly ordered and hence, in particular, not well-ordered) but the cardinality of such a cofinal subset.*

It is worth mentioning in what sense this definition differs from the definition on linearly ordered sets: we drop the necessity of Q being well-ordered. However, there is an analogue on partial orderings:

Definition 3.29. A set $\langle X, \prec \rangle$ is called well-founded if every non-empty subset $Y \subset X$ contains a $\prec$-minimal element. That is, for every $\emptyset \neq Y \subset X$,

$$\exists u \in Y\, \forall v \in Y(\neg(v \prec u)).$$

Remark. *It is worth emphasising that a partial ordering does not obey the trichotomy property; being incomparable is a legitimate possibility for any pair of elements, which therefore invalidate the trichotomy property. In particular, provided $\prec$ is not a linear order, we may not deduce from $\neg(v \prec u)$ that $u \preceq v$.*

In linearly ordered sets these notions coincide, as can be verified easily. In fact, being well-ordered is a special case of being well-founded, as follows from the definition.

A rather well-known definition follows:

Definition 3.30. Let $\langle P, \prec \rangle$ be a partially ordered set. Then $C \subset P$ is called a chain in P if every pair of elements $x, y \in C$ is comparable. That is, for all such pairs, either $x \preceq y$ or $y \preceq x$.

In order to prove the next result, we need a more specific idea of a chain: we do not want to consider chains that are proper subsets of other chains with respect to the same partial ordering. The following well-known definition will help us do so:

Definition 3.31. Let $\langle P, \prec \rangle$ be a partially ordered set. We call a chain C in P a maximal chain if for every element $x \in P \setminus C$ there is an element $y \in C$ such that x is incomparable with y. That is,

$$\forall x \in P \, \exists y \in C (x \in P \setminus C \Rightarrow \neg (y \preceq x \vee x \preceq y)).$$

It is clear that a partially ordered set is linearly ordered if and only if it consists of exactly one maximal chain.

This new setup of partially ordered sets in combination with the introduced definitions of cofinal and well-founded subsets as well as maximal chains gives rise to a result similar to Hausdorff's Cofinality Theorem. The problem is taken from [HH99, p. 91 (6)].

Theorem 3.32. *Every partially ordered set $\langle P, \prec \rangle$ includes a well-founded subset Q that is cofinal in P with respect to $\prec$.*

It is known to the author that the following proof may not be correct.

Proof. This is a proof by construction.
First, we use AC to find the cardinality of the set system comprising all maximal chains in P. We denote it by Γ.
Now define $\{C_\gamma : \gamma \in \Gamma\} \subset \mathcal{P}(P)$ to be the set of all maximal chains in P. Note that $\cup_{\gamma \in \Gamma} C_\gamma = P$. We crucially note that every such chain is linearly ordered by $\prec$. Now we can apply Hausdorff's Cofinality Theorem (theorem 3.5) on C_γ: that is, there exists B_γ such that $B_\gamma \subset C_\gamma$ and B_γ is cofinal and well-ordered in C_γ. Now define

$$Q := \bigcup_{\gamma \in \Gamma} B_\gamma,$$

which is, by definition of each B_γ, cofinal in C_γ for every $\gamma \in \Gamma$. In addition, as B_γ is well-ordered in the chain C_γ, it follows that Q is well-founded in P. We can apply the following reasoning in order to prove the well-foundedness:

Take any non-empty subset $Y \subset Q$. In one case, Y is a subset of one B_i, which is well-ordered by Hausdorff's Cofinality Theorem. The result follows since the subset of a well-ordered set is well-ordered and since every well-ordered set is well-founded (this follows directly from the definition).

In the other case, Y intersects multiple B-sets, which are, by definition, well-ordered. Let $\Gamma_0 \subset \Gamma$ and denote such Y-intersecting B-sets by B_{γ_0} for $\gamma_0 \in \Gamma_0$. Now each $Y \cap B_{\gamma_0}$ has a least element, and so we can define a set, Z say, comprising all such least elements (i.e. the least element of each $Y \cap B_{\gamma_0}$). This constructed set is a subset of Y and contains all the minimal elements in Y, as required. Indeed, by construction, for every element in Z there is no smaller element in Y.

Hence we have proved that Q is well-founded, as required. □

Remark. *This proof merely applies Hausdorff's Cofinality Theorem multiple times, as can be visualised by applying it to a linearly ordered set: When the proof is applied to any linearly ordered set, $\langle \mathbb{Q}, < \rangle$ (the rationals and the usual order) for instance, then* $\Gamma = 1$*. Hence we can apply Hausdorff's Cofinality Theorem directly and obtain a cofinal well-ordered subset, which is therefore also well-founded.*

The problem with the proof above is that an arbitrary union of well-ordered sets is not necessarily well-ordered. A trivial example is given by the set of rationals endowed with the natural ordering. Clearly, each singleton in $\mathbb{Q}$ is well-ordered, its union, however, is not.

Using the notions of regular and singular cardinals, we can also consider the following:

Example 9. *Let $\langle P, \prec \rangle$ be a partially ordered set. Then $\mathrm{cf}(\langle P, \prec \rangle)$ may be a singular cardinal, as can be shown easily: Consider the partially ordered set $\langle P, \prec \rangle = \langle \omega_\omega, <_R \rangle$, where the relation is defined as $<_R := \{\langle a, a \rangle : a \in \omega_\omega\}$. It is worth noting explicitly that since ω_ω is a cardinal, it follows directly that $|\omega_\omega| = \aleph_\omega$. Following the definition of cofinality in partially ordered sets, we see immediately that ω_ω is the only cofinal subset, and hence $\mathrm{cf}(\langle \omega_\omega, <_R \rangle) = \omega_\omega$ (again, this follows since ω_ω is a cardinal). As ω is a limit ordinal, applying proposition 3.21 yields that $\mathrm{cf}(\mathrm{cf}(\langle \omega_\omega, <_R \rangle)) = \mathrm{cf}(\omega_\omega) = \omega$, and hence we have*

$$1 < \omega = \mathrm{cf}(\mathrm{cf}(\omega_\omega)) < \omega_\omega = \mathrm{cf}(\langle \omega_\omega, <_R \rangle),$$

which verifies the definition of singular cardinals.

Further, by introducing the following notion, we will be able to formulate the main problem of this section:

Definition 3.33. Let $\langle P, \prec \rangle$ be a partially ordered set. A set $A \subset P$ is called an <u>independent subset of P</u> if whenever $a, b \in A$ and $a \neq b$ then a and b are incomparable. Such subsets are also called <u>antichains</u>.

We can use this newly introduced definition in order to prove the following interesting theorem (this is a weakening of a problem that can be found in [HH99, p. 92 (6)]). We require another notion first: we say that a partially ordered set P has <u>(finite) width k</u> if there exists an antichain A in P and a natural number k for which $|A| = k$ and every antichain in P has at most cardinality k.

Theorem 3.34. *Let $\langle P, \prec \rangle$ be a partially ordered set. If $\mathrm{cf}(P, \prec)$ is singular then P is not of finite width.*

The proof of this theorem will conclude the section. However, we require a few lemmas and notions first in order to give the proof. We begin by showing the following lemma which is a natural analogue of proposition 3.7 but applied to partially ordered sets:

Lemma 3.35. *Let $\langle P, \prec \rangle$ be a partially ordered set and suppose that Q is a subset of P. If Q is cofinal in P, then $\mathrm{cf}(Q) = \mathrm{cf}(P)$.*

Proof. The proof mimics the reasoning in the justification of proposition 3.7 almost verbatim; interchange the words *well-ordered* and *well-founded* and use definition 3.28 as well as theorem 3.32 in order to complete the proof. □

We also require the following salient theorem due to Dilworth (1950) in our main proof.

Theorem 3.36 (Dilworth's theorem)**.** *Let $\langle P, \prec \rangle$ be a partially ordered set. If P has finite width k then P can be expressed as the disjoint union of k chains in P.*

I am grateful to Dr. Tressl for making me aware of this striking result. We only give the proof in the case that P is finite. It is based on Galvin's elegant approach (see [Gal94] for the original presentation). Dilworth's original proof in [Dil50] includes both the finite and infinite case.

Proof. We use induction on the cardinality of P and consider all widths simultaneously. If $|P| = 1$ there is nothing to prove, hence suppose $|P| = n \geq 2$ and suppose that the result holds for all partially ordered sets with cardinality less than n. Now, as P is finite, it contains a maximal element, m say. Consider the set

$$P' \coloneqq P \setminus \{m\}.$$

Suppose P' has width k. By the inductive hypothesis, we know that we can write

$$P' = \bigcup_{i=1}^{k} C_i$$

where each C_i is a chain and all such chains are pairwise disjoint. There are two cases to consider: either P is also the union of k chains or P has in fact width $k+1$. In both cases, we will have to provide a decomposition of P into chains.

Clearly, every antichain in P' with cardinality k contains exactly one element from each chain C_i. We would like to investigate one antichain in particular: define

$$a_i = \max(a \in C_i : a \text{ is a member of a } k\text{-element } P'\text{-antichain})$$

and consider

$$A = \{a_i : 1 \leq i \leq k\}.$$

This set A is an antichain.

Claim: A is an antichain in P'.
Proof: In view of a contradiction, suppose that A is not an antichain. Then there exist a_i and a_j in A for which $a_i \prec a_j$ (as before, $a_i \in C_i$ and $a_j \in C_j$). By definition, a_j is a member of a k-element antichain in P'. Note that $P' = C_1 \cup C_2 \cup \ldots \cup C_k$ and that every k-element antichain in P' contains exactly one element from each such chain. Hence, as a_j is an element of a k-element antichain, there must exist $r \in C_i$ with $r \succ a_i$ such that r and a_j are incomparable, i.e. both are elements of a k-element antichain. Note that r must be greater than a_i as otherwise $r \prec a_i \prec a_j$, which contradicts the required incomparability. But the existence of such an r contradicts the maximality of a_i. This yields the required contradiction. ■

The maximality is explained as follows: if we consider an element $b \in C_i$ for which $a_i \prec b$ then b is not an element of a k-element antichain in P'. Hence, A is the "largest" k-element P'-antichain with respect to the underlying ordering $\prec$.

Now we consider the two cases outlined above:

- If $A \cup \{m\}$ is an antichain, then we are done: we have found a P-antichain of cardinality $k+1$, and hence the chains $C_1, C_2, \ldots, C_k, \{m\}$ (where $\{m\}$ is the trivial chain) provide the required decomposition.

- If $A \cup \{m\}$ is not an antichain then we have $a_i \prec m$ for some i, by definition. We consider the set of elements in C_i that are smaller than or equal to a_i and denote it by C_{a_i}, i.e.
$$C_{a_i} \coloneqq \{x \in C_i : x \preceq a_i\}.$$
Note that $a_i \in C_{a_i}$. Now, the set $C_{a_i} \cup \{m\}$ is a chain by definition. Further, we show the following:

Claim: The set $P \setminus (C_{a_i} \cup \{m\})$ does not contain any k-element antichains.

Proof: The element a_i is defined to be the largest element of C_i that is part of a k-element antichain in P'. Hence, if we subtract C_{a_i} from P, the only element that could possibly form part of a k-element antichain in P is m (it is comparable to and, in particular, larger than a_i). Thus, after subtracting $\{m\}$, there is no element in $P \setminus (C_{a_i} \cup \{m\})$ that could possibly provide the C_i-component of any k-element antichain in P. Indeed, if there were such an element $a' \in C_i \setminus C_{a_i}$ then a' would contradict the maximality of a_i. ■

Note that the reasoning in the claim above is valid only since m is maximal in P. As $P \setminus (C_{a_i} \cup \{m\})$ does not contain any k-element antichains, it is of width $k-1$. Hence, by the inductive hypothesis, it can be written as the union of $k-1$ pairwise disjoint chains, $D_1, \ldots, D_{k-1}$ say, and thus a decomposition of P is

$$P = (C_{a_i} \cup \{m\}) \cup \bigcup_{i=1}^{k-1} D_i.$$

Therefore, P can be written as the union of k pairwise disjoint chains, as required.

In both cases, we have expressed P as a union of pairwise disjoint chains, and the number of chains equals the width of P. Hence the proof is complete. □

Dilworth's theorem proves a notion similar to compactness in topology; we can find a finite "subcover" whose union equals the underlying set. But Dilworth's theorem is even stronger: it tells us that the components that cover the partially ordered set are also disjoint, and that the minimum number of chains decomposing the partially ordered set is determined by its width. Both these properties will be immensely helpful in the following lemma.

However, and this is crucial, note that we require that the cardinality of the P-independent subsets is uniformly bounded. The following counterexamples show why finite but unbounded antichains do not suffice:

Remark (Counterexamples)**.** *We consider two counterexamples:*

- *I am grateful to Dr. Tressl for the following counterexample: consider an infinite well-ordered set W and the product partial order $\preceq$ on the set $W \times W$ which is defined as follows: for elements $\langle a, b\rangle$ and $\langle a', b'\rangle$ of $W \times W$ we define*

  $$\langle a, b\rangle \preceq \langle a', b'\rangle \Leftrightarrow a \leq a' \text{ and } b \leq b'$$

  and note that an element $\langle c, d\rangle$ is incomparable to $\langle a, b\rangle$ only if

  $$a < c \text{ and } b > d \qquad (*)$$

  or if

  $$a > c \text{ and } b < d.$$

  Now every independent subset is finite. To show this, assume not and fix $\langle a, b\rangle$. Without loss of generality, we may consider case $(*)$ (this follows from the symmetry in the definition of the product partial order above). In order to obtain an infinite antichain, there must exist sequences

  $$d < b < b_1 < b_2 < \ldots \text{ and } c > a > a_1 > a_2 > \ldots,$$

which is equivalent to

$$\langle a_i, b_i \rangle \text{ is incomparable to } \langle a_j, b_j \rangle \text{ for all } i \neq j.$$

Hence the set $\{\langle a, b\rangle, \langle a_1, b_1\rangle, \langle a_2, b_2\rangle, \ldots\}$ forms an antichain. However, the required a-sequence cannot exist as W is well-ordered. Hence there is no infinite antichain.

Crucially, $W \times W$ is not a finite union of disjoint chains. In fact, if $|W| = \kappa \geq \aleph_0$, then we cannot express W as the union of fewer than κ chains, as can be verified easily.

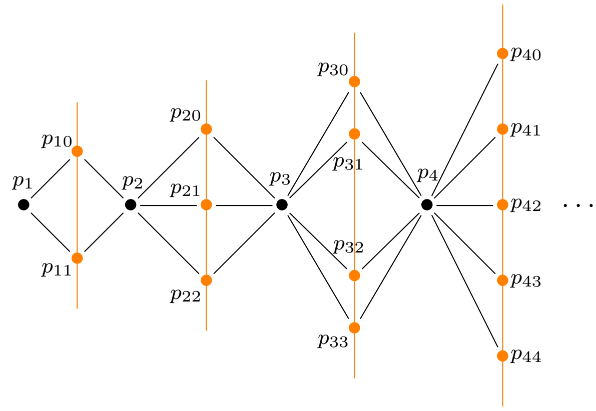
- Now consider a countably infinite well-ordered set $\langle P, \prec\rangle$ with

$$P = \{p_1, p_{10}, p_{11}, p_2, p_{20}, p_{21}, p_{22}, p_3, \ldots\}$$

and endowed with the partial ordering $\prec$. Here, $\prec$ is defined such that

  - p_1 is the minimal element;
  - $\langle p_i, p_{ij}\rangle \in \prec$; and
  - $\langle p_{ij}, p_{i+1}\rangle \in \prec$

for all $i \in \omega \setminus \{0\}$ and for all $j \leq i$. The Hasse diagram of the partially ordered set is given below:



It is clear that P has only finite independent subsets (indicated in orange above). Suppose that P can be expressed as the union of finitely many disjoint chains, k say. Then the subset $P_k = \{p_{k0}, p_{k1}, \ldots, p_{kk}\}$ is an antichain by definition with cardinality $k+1$. However, if P is a union of k disjoint chains then it is clear that P_k cannot be expressed as the union of k chains.

In both cases, the problematic stems from the unboundedness of the antichains. Even though the cardinality of the antichains never attain $\aleph_0$, the fact that their limit does renders Dilworth's theorem inapplicable.

The following lemma is required in the subsequent proposition:

Lemma 3.37. *Let $A_1, A_2, \ldots, A_n$ be well-ordered pairwise disjoint sets. Then*

$$\operatorname{cf}\left(\bigcup_{i=1}^{n} A_i\right) = \max_{1 \leq i \leq n} \operatorname{cf}(A_i).$$

Proof. We prove the result by induction on n.

If $n = 1$ then the result is trivial. Hence suppose that we have sets $A_1, \ldots, A_{k+1}$ as described above, and further assume that the result holds for all unions of k-many such well-ordered pairwise disjoint sets. We rewrite the union of sets as

$$\bigcup_{i=1}^{k+1} A_i = A_{k+1} \cup \bigcup_{i=1}^{k} A_i$$

and hence we see that

$$\mathrm{cf}\left(\bigcup_{i=1}^{k+1} A_i\right) = \mathrm{cf}\left(A_{k+1} \cup \bigcup_{i=1}^{k} A_i\right).$$

Recall that the sets A_i are disjoint, and hence any cofinal subset of $\bigcup_{i=1}^{k+1} A_i$ contains a subset $A'_{k+1} \subset A_{k+1}$ that is cofinal in A_{k+1} and disjoint from $\bigcup_{i=1}^{k} A_i$. Hence

$$\begin{aligned}
\mathrm{cf}\left(A_{k+1} \cup \bigcup_{i=1}^{k} A_i\right) &= \mathrm{cf}(A_{k+1}) + \mathrm{cf}\left(\bigcup_{i=1}^{k} A_i\right) \\
&= \mathrm{cf}(A_{k+1}) + \max_{1 \le i \le k} \mathrm{cf}(A_i) && \text{by the inductive hypothesis} \\
&= \max\left(\mathrm{cf}(A_{k+1}), \max_{1 \le i \le k} \mathrm{cf}(A_i)\right) && \text{by the Fundamental Theorem} \\
&= \max(\mathrm{cf}(A_{k+1}), \mathrm{cf}(A_1), \ldots, \mathrm{cf}(A_k))
\end{aligned}$$

which proves the claim. (Note that we have used cardinal arithmetic here, as the cofinality of the partially ordered set $\bigcup_{i=1}^{k+1} A_i$ is defined to be a cardinal.)

Hence the proof by induction is complete. □

One more result is needed:

Proposition 3.38. *Assume P is an infinite partially ordered set with finite width k. Suppose that $C_1, C_2, \ldots, C_k$ are pairwise disjoint chains such that $P = \bigcup_{i=1}^{k} C_i$. Then the following hold:*

- *There exist sets $C'_1, C'_2, \ldots, C'_k$ such that for each i the set C'_i is well-ordered and cofinal in C_i.*
- $\mathrm{cf}(P) = \max_{1 \le i \le n} \mathrm{cf}(C'_i)$

Proof. We ignore the trivial case and hence assume that $k \geq 2$. Thus, in particular, P is not linearly ordered.

Firstly, by Dilworth's theorem, we can guarantee the existence of pairwise disjoint chains $C_1, C_2, \ldots, C_k$ as needed. Now we use Hausdorff's Cofinality Theorem to find a well-ordered cofinal subset $C'_i \subset C_i$ for each i. It is easy to see that $\bigcup_{i=1}^{k} C'_i$ is cofinal in P (and indeed well-founded as each chain C'_i is well-ordered). This proves the first part.

As the case of finite chains is trivial (their cofinality is trivially 1), assume that every chain is infinite. By lemma 3.35 we hence know that

$$\mathrm{cf}(P) = \mathrm{cf}\left(\bigcup_{i=1}^{k} C'_i\right).$$

The second part of the result now follows from the previous proposition: we see that each C_i' is well-ordered and further all C_i' are pairwise disjoint. Hence

$$\begin{aligned}\mathrm{cf}(P) &= \mathrm{cf}\left(\bigcup_{i=1}^{k} C_i'\right)\\ &= \max_{1\leq i\leq k} \mathrm{cf}(C_i')\end{aligned}$$

which proves the claim. □

We are now ready to proceed with the main proof:

Proof of theorem 3.34: We aim to prove the contrapositive. Hence suppose P is infinite and assume that P is of finite width k. We need to show that $\mathrm{cf}(P)$ is not singular.

First, consider the case in which the only independent subsets of P are singletons (i.e. P is of width 1). Then all elements in P are comparable, ergo P is in fact linearly ordered. Hence, P contains a well-ordered cofinal subset (this follows from Hausdorff's Cofinality Theorem), which has an order type, α say. Therefore $\mathrm{cf}(P) = \mathrm{cf}(\alpha)$, where $\mathrm{cf}(\alpha)$ is a regular cardinal by theorem 3.25, as required.

Now consider the case in which P is of width k with $k \geq 2$. We apply Dilworth's theorem and hence obtain disjoint chains $C_1, \ldots, C_k$ whose union equals P. As before, we apply Hausdorff's Cofinality Theorem and obtain well-ordered chains $C_1', \ldots, C_k'$ where C_i' is cofinal in C_i for each i. Hence, the union $\bigcup_{i=1}^{k} C_i'$ is cofinal in P.

If we now use lemma 3.35 as well as the previous proposition then we obtain

$$\begin{aligned}\mathrm{cf}(P) &= \mathrm{cf}\left(\bigcup_{i=1}^{k} C_i'\right)\\ &= \max_{1\leq i\leq n} \mathrm{cf}(C_i').\end{aligned}$$

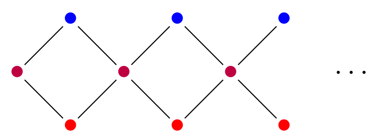
But theorem 3.25 tells us that, as each C_i' is in fact linearly ordered, we obtain that $\mathrm{cf}(C_i')$ is regular. Hence the proof is complete. □

The crucial step in the proof above is to realise that if P has only finite independent subsets, then we must be able to "cover" P with finitely many chains. Employing the useful properties of cardinal arithmetic then yields the result quite easily.

The following remark visualises this idea of "covering" the partially ordered set:

Example 10. *It is clear that there are infinite partially ordered sets with infinitely many maximal chains (where no two are disjoint) but only finite independent sets: consider the set $P = \{p_0, p_1, p_2, \ldots\}$ (i.e. with cardinality $\aleph_0$) with the partial order $\prec$ defined to be as follows:*

- *p_0 is the unique minimal element in P.*
- *$p_{3k} \prec p_{3k+1}$ and $p_{3k} \prec p_{3k+2}$ for all $k \in \omega$.*
- *$p_{3k} \succ p_{3k-1}$ and $p_{3k} \succ p_{3k-2}$ for all $k \in \omega \setminus \{0\}$.*
- *p_{3k+1} and p_{3k+2} are incomparable for all $k \in \omega$.*



This is obviously an infinite partially ordered set with infinitely many maximal chains (where all of which are infinite), but every independent subset has cardinality 2 *(comprising elements of the form* p_{3k+1} *and* p_{3k+2} *for some* $k \in \omega$*).*

It is now clear that, from the definition of the partial ordering on P*, two maximal chains suffice to cover* P*. Informally, define* $C_1 \subset P$ *to always cover the "blue" element (that is, elements of the form* p_{3k+1}*) and define* $C_2 \subset P$ *to cover the "red" element (that is, elements of the form* p_{3k+2}*) as well as the intersecting elements (purple). Now* $C_1 \cup C_2 = P$*, and* $C_1 \cap C_2$ *is empty (as per the Hasse diagram above). We can now apply the contrapositive of theorem 3.34 and obtain that the cofinality of* P *is regular (it is in fact* $\aleph_0$*, as can be verified very easily).*

It can be shown that the result we have just proved can in fact be strengthened; one can even show that if $\mathrm{cf}(P)$ is singular then P contains an infinite independent subset (this is not a consequence of our result, as P could have only finite but unbounded antichains, as seen in the remark above covering two counterexamples). The statement can be found in [HH99, p. 92].

3.4 Cardinal Arithmetic

In the course of this report, we have mentioned the Generalised Continuum Hypothesis only once. Its importance is undeniable, however, and in order to give some insight into its paramount role, this subsection will serve as a compilation of important results in this area.

As outlined previously, cardinal addition and multiplication are easy: for any two cardinals, their sum and product equal their maximum provided one of them is infinite. Cardinal exponentiation, however, is significantly harder, and determining powers of cardinals is highly non-trivial and depends to a great extent on the axioms we assume.

However, cardinal exponentiation is of utmost interest. Of course, cardinal exponentiation is the main ingredient of GCH, for example, and giving better bounds on operations involving cardinal exponentiation is hence crucial. We now present a few standard results in this area, which will allow us to obtain further inequalities on cardinal exponentiation.

In this section, all operations applied to well-ordered sets are the respective cardinal-operations (cardinal addition, multiplication, exponentiation), unless stated otherwise.

The reader is certainly familiar with the following result:

Theorem 3.39 (Cantor's theorem)**.** *Let* κ *be a cardinal. Then*

$$\kappa < 2^\kappa.$$

This is a fairly weak statement, as it only gives a bound on one cardinal at a time. Can we do better? The following standard theorem shows that we can indeed:

Theorem 3.40 (König's theorem)**.** *Let* α *be an ordinal. Consider two set systems of cardinals of the form* $\{\lambda_\xi : \xi < \alpha\}$ *and* $\{\kappa_\xi : \xi < \alpha\}$*. If* $\lambda_\xi < \kappa_\xi$ *for all* $\xi < \alpha$*, then*

$$\sum_{\xi<\alpha} \lambda_\xi < \prod_{\xi<\alpha} \kappa_\xi.$$

This result is in fact a generalisation of Cantor's theorem, as can be verified easily: if we put $\lambda_\xi = 1$ and $\kappa_\xi = 2$ for all $\xi < \alpha$, then we obtain that

$$\sum_{\xi<\alpha} \lambda_\xi = \sum_{\xi<\alpha} 1 = |\alpha| < \prod_{\xi<\alpha} 2 = 2^{|\alpha|},$$

which confirms what we would expect from Cantor's theorem.

Remark. *The following standard result will be used in the following proof; it will be stated for convenience: let A and B be sets. If there does not exist a surjection from A to B, then $|A| < |B|$.*

We will return to the proof of theorem 3.40 later; the following result and proof which can be found in [Kun80, p. 34] (comments and further explanations have been added accordingly) will be very useful:

Proposition 3.41 (König's lemma)**.** *If $\kappa \geq \aleph_0$ is a cardinal and $\mathrm{cf}(\kappa) \leq \lambda$, then $\kappa^\lambda > \kappa$.*

Proof. Consider any cofinal map f from λ into κ. Such a map exists by our assumption that the cofinality of κ is less than or equal to λ. Let G be a function from κ to ${}^\lambda\kappa$. Then the image of G comprises functions with domain λ and range κ. Define a function $h\colon \lambda \to \kappa$ by

$$\alpha \mapsto \min(\kappa \setminus \{(G(\mu))(\alpha) : \mu < f(\alpha)\}).$$

Claim: The set $\kappa \setminus \{(G(\mu))(\alpha) : \mu < f(\alpha)\}$ is non-empty.
Proof: In view of a contradiction, fix any $\alpha \in \lambda$ and suppose that

$$\kappa = \{(G(\mu))(\alpha) : \mu < f(\alpha)\}.$$

Define $F_\alpha(\mu) \coloneqq (G(\mu))(\alpha)$ and hence obtain the equality

$$\kappa = \{F_\alpha(\mu) : \mu < f(\alpha)\}$$

where $F_\alpha(\mu)\colon \kappa \to \kappa$. Further, note that $f(\alpha) \in \kappa$. But now κ is the image of the function F_α restricted to $f(\alpha)$. So we must have an injection from $f(\alpha)$ into κ. As $f(\alpha) < \kappa$, this contradicts the fact that κ is a cardinal. ■

Trivially, a least element always exists as κ is well-ordered.

Is h in the image of G? Assume (in view of a contradiction) that h is indeed in the image of G so that there exists $\beta \in \kappa$ for which $G(\beta) = h$. This is the same as saying that $(G(\beta))(\gamma) = h(\gamma)$ for all $\gamma \in \lambda$. Fix any $\gamma \in \lambda$. Then

$$(G(\beta))(\gamma) = h(\gamma) = \min(\kappa \setminus \{(G(\mu))(\gamma) : \mu < f(\gamma)\}).$$

Hence we have that

$$(G(\beta))(\gamma) \in \kappa \setminus \{(G(\mu))(\gamma) : \mu < f(\gamma)\}$$

and thus it follows that

$$(G(\beta))(\gamma) \in \{(G(\mu))(\gamma) : f(\gamma) \leq \mu\}.$$

This final deduction is allowable since $(G(\mu))(\delta) \in \kappa$ for all $\mu \in \kappa$ and for all $\delta \in \lambda$ by the definition of G. Further, note that the right hand side is non-empty.

Hence $f(\gamma) \leq \beta$. We have fixed an arbitrary $\gamma \in \lambda$, though, and hence it follows that $f(\gamma) \leq \beta$ for all $\gamma \in \lambda$. As β is an element of κ, it follows that f is not a cofinal map into κ. Contradiction.

Now our defined function h is not in the image of G. Thus G is not surjective. As G was chosen arbitrarily, we can deduce (using the previous remark) that $\kappa^\lambda > \kappa$, as required. □

Remark. *The crucial point to remark in the proof above is the diagonal argument that allows us to deduce that h cannot be in the image of G. We will use a similar reasoning later in the proof of König's theorem.*

The following example can be found in [BP77, p. 109]. In the paper, however, it is an application of König's theorem, which we shall examine shortly. Its statement can be derived from König's lemma directly.

Example 11. *Let $\kappa = \aleph_\omega$. Then clearly $\mathrm{cf}(\aleph_\omega) = \omega$, and hence applying König's lemma yields*

$$\aleph_\omega < \aleph_\omega^{\aleph_0}.$$

Aside. *The reader should note that both the terms "König's lemma" and "König's theorem" are not standard; in the literature, one can find various statements with names resembling our nomenclature (whilst a few such results are due to Julius König, the set theorist whose results are of interest to us, there are equally theorems proved by his son Dénes Kőnig, whose work was mainly focused on graph theory). In [WM14, p. 379], Wate-Mizuno gives further details on the historical background.*

Corollary 3.42. *If κ is an infinite cardinal, then $\kappa^{\mathrm{cf}(\kappa)} > \kappa$.*

Proof. This is immediate from König's lemma above; simply identify λ in the statement of König's lemma with $\mathrm{cf}(\kappa)$ and the result follows. □

This corollary is very useful: for any cardinal κ, we are now given an upper bound on the least cardinal λ for which κ^λ is strictly bigger than κ. However, we have not made any progress on determining κ^θ for $\kappa < \theta < \mathrm{cf}(\kappa)$. As we will see later, a seminal theorem by Easton will show that we cannot say much more than this.

Example 12. *We have seen examples of the corollary above: we know that $\aleph_0^{\aleph_0} = 2^{\aleph_0}$ (this can be shown using a simple argument involving cardinal arithmetic), which in turn is strictly greater than $\aleph_0$ by Cantor's theorem.*

Another corollary to König's lemma can be found by considering its contrapositive.

Proposition 3.43. *Assume κ and λ are cardinals. If $\kappa^\lambda \leq \kappa$ then κ is finite or $\lambda < \mathrm{cf}(\kappa)$.*

If we assume that κ is infinite from the outset, we obtain the following:

Corollary 3.44. *Assume κ and λ are cardinals. If $\kappa \geq \aleph_0$ and $\kappa^\lambda \leq \kappa$ then $\lambda < \mathrm{cf}(\kappa)$.*

In particular, and much more useful at that, is the following:

If $\kappa \geq \aleph_0$ and $\kappa^\lambda = \kappa$, then $\lambda < \mathrm{cf}(\kappa)$.

Remark. *There are two simple remarks to make about the previous two results:*

- *If $\kappa^\lambda = \kappa$ and κ is finite, then we must have $\kappa = 0$.*
- *If κ is infinite, then it is clear that we obtain the strict inequality $\kappa^\lambda < \kappa$ if and only if $\lambda = 0$.*

We will ignore these two trivial cases in the future.

Let κ and λ be cardinals. As we are particularly interested in exponentiation of infinite cardinals, we might wonder what the cofinality of κ^λ is. Further, we want to apply corollary 3.44, and hence assume that κ is infinite. As we have met all the hypotheses, we can now give a bound on the value of $\mathrm{cf}(\kappa^\lambda)$ using the previous corollary:

Consider corollary 3.44 and hence the statement

$$\mu^\theta = \mu$$

with an infinite cardinal μ and an arbitrary non-zero cardinal θ. Put $\mu = \kappa^\lambda$ and hence obtain

$$\left(\kappa^\lambda\right)^\theta = \kappa^\lambda.$$

Once we have established this equality (i.e. once we have determined for which values of θ this equality holds) we can appeal to corollary 3.44 in order to obtain the required bound on the cofinality of κ^λ. But this is easy: by the Fundamental Theorem of Cardinal Arithmetic, we have

$$\max(\lambda, \theta) = \lambda \Rightarrow \left(\kappa^\lambda\right)^\theta = \kappa^{(\lambda \times \theta)} = \kappa^\lambda$$

which provides us with the required lower bound.

Hence we can state the following:

Corollary 3.45. *Assume κ and λ are non-zero cardinals and suppose κ is infinite. If a cardinal θ satisfies $0 < \theta \leq \lambda$, then $\mathrm{cf}(\kappa^\lambda) > \theta$.*

Hence, in particular, we have that $\mathrm{cf}(\kappa^\lambda) > \lambda$.

This observation gives rise to a very useful result which we will call *König's inequality* (this choice of nomenclature can be found in [FH08, p. 191]):

Lemma 3.46 (König's inequality)**.** *Let κ be an infinite cardinal. Then $\mathrm{cf}(2^\kappa) > \kappa$.*

Proof. We can apply corollary 3.45 directly: we see that 2^κ is infinite, and then note that

$$(2^\kappa)^\kappa = 2^{(\kappa \times \kappa)} = 2^\kappa.$$

Therefore

$$\mathrm{cf}((2^\kappa)^\kappa) = \mathrm{cf}(2^\kappa)$$

and thus, by the second part of corollary 3.45, we obtain $\mathrm{cf}(2^\kappa) > \kappa$. □

König's inequality plays a major role in the aforementioned theorem by Easton published in 1970.

After this brief digression drawing our attention to König's lemma, we can now proceed with the main proof of the stronger theorem by König (this is a proof given by Devlin in [Dev17, p. 22]; further comments and explanations have been added accordingly):

Proof of König's theorem. We begin by considering the trivial cases first: if α is finite, then we can apply the Fundamental Theorem of Cardinal Arithmetic whenever there exists an infinite κ_ξ. If also all κ_ξ are finite, then the fact that addition and multiplication on finite cardinals behave just like their counterparts on the usual integers implies the inequality. In both cases the result is immediate.

For ease of notation, define

$$\Lambda = \bigcup_{\beta<\alpha} \{\beta\} \times \lambda_\beta$$

and

$$K = \{f : \mathrm{dom}(f) = \alpha \wedge \forall \beta < \alpha (f(\beta) \in \kappa_\beta)\}.$$

Note that

$$|\Lambda| = \sum_{\beta<\alpha} \lambda_\beta \qquad \text{and that} \qquad |K| = \prod_{\beta<\alpha} \kappa_\beta.$$

When we now consider a function

$$S : \Lambda \to K$$

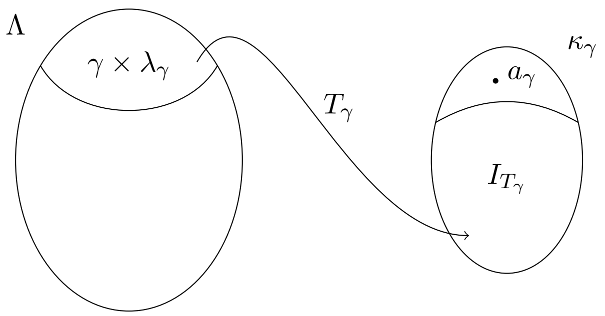
then it suffices to show that S cannot be surjective in order to prove the result.

To prove the required strictness, consider the following setup: for any function S as defined above, consider a function T_γ for $\gamma \in \alpha$ that is defined as follows:

$$T_\gamma \colon \Lambda \to \kappa_\gamma$$
$$\langle \beta, \xi \rangle \mapsto [S(\langle \beta, \xi \rangle)](\gamma)$$

(It is clear that $[S(\langle \beta, \xi \rangle)](\gamma)$ is an element of κ_γ as the function S returns a function, s say, for which $s(\gamma) \in \kappa_\gamma$ by definition of K.)

By our initial assumption, we know that $\lambda_\gamma < \kappa_\gamma$. Note that $|\{\gamma\} \times \lambda_\gamma| = \lambda_\gamma$. Thus $T_\gamma \restriction (\{\gamma\} \times \lambda_\gamma)$ cannot map surjectively to κ_γ. Hence the set $\kappa_\gamma \setminus \mathrm{img}(T_\gamma \restriction (\{\gamma\} \times \lambda_\gamma))$ is non-empty, and we can pick an element from it (its minimal element for instance) and denote it by a_γ.



$$I_{T_\gamma} \coloneqq \mathrm{img}(T_\gamma \restriction (\gamma \times \lambda_\gamma))$$

If we now define a function $g \colon \alpha \to K$ by $\beta \mapsto a_\beta$, then $g \in K$ (its domain is α and for any $\beta \in \alpha$, $g(\beta) = a_\beta \in \kappa_\beta$, by definition).

However, $g \notin \mathrm{img}(S)$. This follows immediately as we have chosen the values of g based on the limitations of the values of S.

(Technically, we have chosen a_γ for each γ based on the fact that a_γ is not an element of the image of $T_\gamma \restriction (\{\gamma\} \times \lambda_\gamma)$. The fact that T_γ is defined using S proves the result).

As there is no surjection from Λ to K, we can conclude that

$$|\Lambda| = \sum_{\beta<\alpha} \lambda_\beta < \prod_{\beta<\alpha} \kappa_\beta = |K|,$$

by the remark given above König's lemma.

Therefore S is not surjective and, similar to König's lemma, the result follows. $\square$

Note that the diagonal argument employed in the previous proof is inherently similar to our reasoning in the proof of König's lemma. Further, it is evident that this result can be interpreted as a very natural analogue to addition and multiplication of integers.

Remark. *In the previous proof, we used the definitions of cardinality of sums and products of cardinals, respectively. These seem abstract at first sight, can be visualised informally as follows:*

- *For the sum, note that this is a natural extension of ordinary summation over cardinals. We introduce the first element of the ordered pair so as to avoid conflating identical elements.*
- *For the product, the reader might want to think of α-many boxes which are numbered through by $\xi < \alpha$; each box ξ contains κ_ξ-many elements. In order to determine the product of all κ_ξ we "count" the number of possible ways of picking exactly one element from each box. This number gives the product of all κ_ξ. Indeed, this interpretation coincides with our combinatorial understanding of finite integer multiplication and finite choice (one might want to think of classic examples of probability theory including drawing balls from urns).*

We may now revisit example 11:

Example 13. *Let $i = \omega$ and put $\kappa_i = \aleph_i$ for all $i \in \omega$. Then $\sum_{i\in\omega} \kappa_i = \sum_{i\in\omega} \aleph_i = \aleph_\omega$. If we now set $\lambda_i = \aleph_\omega$, then clearly König's theorem is applicable, and hence we obtain the inequality*

$$\aleph_\omega = \sum_{i\in\omega} \aleph_i < \prod_{i\in\omega} \aleph_\omega = \aleph_\omega^{\aleph_0},$$

which agrees with our result from example 11.

Note that we may go one step further: we can now derive König's lemma from König's theorem (this is exercise [Kun80, p. 45 (18) (b)]):

We may use the notation as employed above and set $\alpha = \lambda$, $\kappa_\xi = \kappa$ for all $\xi \in \lambda$, and consider the κ-unbounded set of cardinals $\{\theta_\xi : \xi \in \lambda\}$ (this exists as $\mathrm{cf}(\kappa) \leq \lambda$). Note that $\theta_\xi < \kappa$ for every $\xi \in \lambda$. Hence we can apply König's theorem and obtain

$$\prod_{\xi\in\lambda} \kappa_\xi = \prod_{\xi\in\lambda} \kappa = \kappa^\lambda > \sum_{\xi\in\lambda} \theta_\xi \geq \sup_{\xi\in\lambda} \theta_\xi = \bigcup_{\xi\in\alpha} \theta_\xi = \kappa.$$

where the final equality follows from the definition of cofinality. Also, here we have used the cardinal arithmetic properties derived at the beginning of section 3.2.

3.4.1 Aside: Easton's Theorem

The following theorem is included for completeness only; its importance in the field is crucial. The original statement and proof can be found in [Eas70]; here we give a slightly more accessible version due to Hajnal and Hamburger, which can be found in [HH99, p. 243]:

Theorem 3.47 (Easton's theorem)**.** *Assume $\mathcal{M}$ is a countable model of* ZFC *with universe M in which* GCH *holds and suppose that G is a unary function in $\mathcal{M}$. Assume the following two conditions hold:*

(1) G is an increasing function on ordinals in M

(2) $\mathrm{cf}(\aleph_{G(\alpha)}) > \aleph_\alpha$ for all $\alpha \in M$

Then there exists an extension $\mathcal{N}$ of $\mathcal{M}$ in which the cardinals and cofinalities in $\mathcal{N}$ coincide with those in $\mathcal{M}$. Further, such an extension satisfies

$$2^{\aleph_\alpha} = \aleph_{G(\alpha)}$$

for all regular cardinals $\aleph_\alpha$ for which $\alpha \in M$.

The proof requires advanced techniques such as forcing, hence it is omitted. An alternative proof to the source mentioned above can be found in [Kun80, p. 264].

The statement of this result is very powerful. In essence, if ZFC is consistent and we are given a particular model, we can then choose any function G, subject to the aforementioned conditions, and find a model of ZFC in which the cardinality of the power set of any regular cardinal $\aleph_\alpha$ equals exactly $\aleph_{G(\alpha)}$.

A function satisfying the two conditions stated above is called an Easton function.

Remark. *It is easy to derive the necessity of conditions (1) and (2) in the statement of Easton's theorem above: condition (1) follows from the fact that the continuum function $\alpha \mapsto 2^\alpha$ is strictly increasing on* **ON**. *Hence so is the function $\alpha \mapsto 2^{\aleph_\alpha}$. Therefore, any function G we consider must satisfy this condition. The second condition stems from the limitations determined in König's theorem; it is, in fact, a restatement of König's inequality. If either were violated, $\mathcal{M}$ would not be a model for* ZFC.

As pointed out in Easton's original paper (see [Eas70, pp. 140, 175] for interest), it is worth mentioning that in the constructed extension the cardinality of power sets of singular cardinals equals exactly the least cardinal permissible by König's theorem.

We can consider an easy application of Easton's theorem:

Example 14. *With the assumption as above, assume* ZFC *is consistent with a given model $\mathcal{M}$. Define $G\colon M \to M$ by $G(\alpha) = \alpha + 2$. Hence, by Easton's theorem, there exists an extension $\mathcal{N}$ for $\mathcal{M}$ in which $2^{\aleph_\alpha} = \aleph_{\alpha+2}$ for all ordinals α in M for which $\aleph_\alpha$ is regular.*

Example 15. *Although it is increasing, we cannot use the function $G(\alpha) = \omega \times \alpha$; condition (2) is violated as, for instance with $\alpha = 2$, we obtain $\mathrm{cf}(\aleph_{\omega\times 2}) = \mathrm{cf}(\omega \times 2) = \omega \not> \aleph_2$. Indeed, such a choice of G contradicts König's inequality.*

Similarly, if $G(\alpha) = \alpha$, i.e. a fixed point, then we obtain $2^{\aleph_\alpha} = \aleph_\alpha$, which contradicts Cantor's theorem.

Finally, as Prikry and Baumgartner point out (see [BP77, p. 109]), although one of the remarkable properties of Easton's theorem is the fact that it can treat many regular cardinals at once, it isn't limited to this universal case. For example, if we assume ZF to be consistent, then so is ZF + "$2^{\aleph_4} = \aleph_6$". Further, the Easton function need of course not be linear. Hence one can apply Easton's theorem and show that if ZF is consistent then so is

$$\mathrm{ZF} + \text{“}2^{\aleph_2} = \aleph_4\text{”} + \text{“}2^{\aleph_3} = \aleph_9\text{”} + \dots$$

for example.

Again, we cannot have $2^{\aleph_4} = \aleph_\omega$ for example, as this would violate condition (2): $\mathrm{cf}(\aleph_\omega) = \omega \not> \aleph_4$.

Remark. *Unfortunately, Easton's theorem does not provide any information on singular cardinals. The theory behind the continuum function of singular cardinals is called the Singular Cardinal Problem, and culminates in the so-called Singular Cardinal Hypothesis (*SCH *for short). An example as given in [BP77, p. 110] is the following: if for all $n \in \omega$ we have $2^{\aleph_n} = \aleph_{n+1}$, does that imply $2^{\aleph_\omega} = \aleph_{\omega+1}$? The answer is no as Magidor showed in [Mag77] (the assumption of the consistency of the existence of certain large cardinals was employed). We will briefly return to* SCH *later on.*

Hence it is possible for a singular cardinal with countable cofinality to be the least cardinal for which GCH *fails (provided we assume the existence of certain large cardinals). However, as we shall see later on, Silver's theorem found in 1974 gives a surprising result: if κ is a singular cardinal with uncountable cofinality and* GCH *holds for all cardinals below κ, then* GCH *also holds for κ. Hence a singular cardinal with uncountable cofinality cannot be the least cardinal failing* GCH.

We will return to further results concerning the Generalised Continuum Hypothesis later on (cf. Silver's theorem). For now, however, we will turn towards one of the main sections of this report: the next section covers the vital notion of *stationary sets*.

4 Clubs, ξ-Large Sets, and Stationary Sets

This sections introduces stationary sets, a crucial ingredient of infinitary combinatorics. In the following sections, we will examine the definition of stationary sets as well as applications thereof.

4.1 Filters and Ultrafilters

Filters, as presented in this section, will be particularly useful later once we have introduced clubs and stationary sets. We begin by stating the definition.

Definition 4.1. Let X be a set. Then $\mathcal{F} \subset \mathcal{P}(X)$ is a filter on X if $\mathcal{F} \neq \emptyset$ and if the following three conditions hold:

(i): $\emptyset \notin \mathcal{F}$

(ii): If $A \in \mathcal{F}$ and $B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$.

(iii): If $A \in \mathcal{F}$ and $A \subset B$ then $B \in \mathcal{F}$.

That is, $\mathcal{F}$ is closed under finite intersections and any superset of an element of $\mathcal{F}$ is also a member of $\mathcal{F}$.

Remark. *The set X from definition 4.1 is called the ground set. In some instances it is also called the underlying set. These notions are synonymous.*

We visualise this concept using the following examples:

Example 16. *For any set X, a filter is given by $\mathcal{F} = \{X\}$. We call this the trivial filter.*

Example 17. *Consider any infinite set X and consider the set of all subsets of X whose complements are finite. Such sets are called cofinite sets. That is, consider*

$$\mathcal{A} = \{Y \subset X : |Y^{\mathsf{c}}| < \omega\}.$$

This forms a filter as can be verified easily:

- *As X is infinite, the empty set is not in $\mathcal{A}$.*
- *Let $B, C \in \mathcal{A}$. Then B^{c} and C^{c} are finite sets. Now $(B \cap C)^{\mathsf{c}} = B^{\mathsf{c}} \cup C^{\mathsf{c}}$ (by De Morgan's laws) which is clearly finite. Hence $B \cap C \in \mathcal{A}$.*
- *Let $B \in \mathcal{A}$ and assume $B \subset C$. Now B^{c} is finite, and as $C^{\mathsf{c}} \subset B^{\mathsf{c}}$, we can deduce that $C \subset \mathcal{A}$.*

Remark. *The existence of the trivial filter shows that every set admits a filter. Further, it is worth mentioning that $\mathcal{F}$ is a filter on X only if $X \in \mathcal{F}$: assume $\mathcal{F}$ is a filter on X. For any $A \in \mathcal{F}$ we have that $A \subset X$. Condition 2 yields the result.*

Example 18. *Let D be a non-empty subset of a set X. Then*

$$\mathcal{F}_D = \{Y \subset X : D \subset Y\}$$

is a filter on X. This can easily be shown by verifying the definition:

- *Assume A, $B \in \mathcal{F}_D$. Then $D \subset A$ and $D \subset B$, and hence $D \subset A \cap B$, which implies that $A \cap B \in \mathcal{F}_D$. This proves closure under intersections.*

- *Assume* $A \in \mathcal{F}_D$ *and* $A \subset B$*, where* $B \in \mathcal{P}(X)$*. So* $D \subset A$*, and* $A \subset B$ *yields that* $D \subset B$*, and so* $B \in \mathcal{F}_D$*, as required.*

We call such a filter a principal filter.

It is clear that a principal filter is uniquely determined by its generating set D.

Example 19. *Take any set of propositional variables* $\mathcal{L}$ *and consider the set of all sentences constructed from these variables, which is denoted by* $S\mathcal{L}$*. By definition of* $S\mathcal{L}$*, it is the union of sets* $S_n\mathcal{L}$ *over all* $n \in \omega$*. Now fix any* $k \in \omega$ *and define*

$$\mathcal{F}_{S_k\mathcal{L}} = \{A \in \mathcal{P}(S\mathcal{L}) : S_k\mathcal{L} \subset A\}.$$

This constitutes a filter; it is the principal filter generated by the set $S_k\mathcal{L}$*.*

When considering this set and setup in an intuitive way, this definition indeed makes sense: if $A, B \supset S_k\mathcal{L}$*, then every sentence in* A *and every sentence in* B *is of complexity at least* k*, hence so is the intersection of* A *and* B*. Similarly, any superset of* A *is comprised of sentences of complexity at least* k*.*

Remark. *In the literature, elements of a given filter are considered "large" with respect to the underlying set. Principal filters as defined here give an intuitive understanding of this choice of terminology.*

Principal filters are especially useful as we now have a tool at hand to generate a non-trivial filter on any set by constructing it using only one subset. That is, we do not need any further knowledge about the respective set. On the flip side, such principal filters are in some sense trivial.

Remark. *A couple of counterexamples follow:*

- *Take any set* X *with more than one element and consider* $\mathcal{P}(X) \setminus \{\emptyset\}$*. Condition 2 holds by definition of the power set. However, Condition 1 is violated: We can find elements* $\{a\}$ *and* $\{b\} \in \mathcal{P}(X)$ *such that* $\{a\} \cap \{b\} = \emptyset$*, which is not an element of* $\mathcal{P}(X) \setminus \{\emptyset\}$*.*
- *Take the real line* $\mathbb{R}$ *as the underlying set and consider the set system*

$$\mathcal{E} = \{(-\infty, a] : a \in \mathbb{R}\},$$

which is clearly a subset of $\mathcal{P}(\mathbb{R})$*. Then* $\mathcal{E}$ *is closed under intersections (i.e. it satisfies Condition 1), but Condition 2 is violated as* $\mathbb{R} \notin \mathcal{E}$*.*

We can strengthen the idea of a filter as follows:

Definition 4.2. A filter $\mathcal{F}$ on a set X is called an ultrafilter on X if

$$A \in \mathcal{P}(X) \Rightarrow A \in \mathcal{F} \vee X \setminus A \in \mathcal{F}.$$

That is, for any element in the power set of the underlying set, either the element itself or its complement with respect to the underlying set is a member of the filter.

Remark. *We need not consider the case in which both* A *and* $X \setminus A$ *are an element of* $\mathcal{F}$*; in such a case observe that* $A \cap (X \setminus A) = \emptyset$*, which contradicts the fact that* $\mathcal{F}$ *was assumed to be a filter.*

An ultrafilter is therefore an extension of the original idea of a filter; by adding the additional constraint we have obtained a connection between the underlying set X and its power set $\mathcal{P}(X)$. Further, as mentioned earlier, elements of filters can be perceived as somewhat large. This interpretation is helpful in regard to ultrafilters, too: here, whenever we consider a subset, either the subset is large or its complement is. Informally, one is larger than the other.

(Technically, we would be required to put the term "large" in quotation marks as we have never defined size in a sense different from cardinality. The description above is simply an attempt to visualise the idea behind the notion of filters. When we consider ideals later on, this visual description and its importance will become obvious.)

One should note that we need not include the restriction $A \neq \emptyset$, as $X \in \mathcal{F}$ for any filter $\mathcal{F}$ on X by the first remark of this section.

This notion gives rise to the following result:

Theorem 4.3. *Let $D \subset X$. The principal filter $\mathcal{F}_D$ is an ultrafilter on X if and only if D is a singleton.*

Proof.

($\Rightarrow$): Assume D is not a singleton and, for contradiction, assume that $\mathcal{F}_D$ is an ultrafilter. Let $e \in D$ and consider the proper subset $E = \{e\} \subset D$. This subset is proper since D is not a singleton. Clearly, $E \in \mathcal{P}(X)$, and as $\mathcal{F}_D$ is an ultrafilter, either $E \in \mathcal{F}_D$ or $X \setminus E \in \mathcal{F}_D$. But $D \not\subset X \setminus E$, since $e \notin X \setminus E$. Hence $E \in \mathcal{F}_D$, that is, D must be a subset of E. But now we have that $D \subset E$ and E is a proper subset of D, a contradiction.

($\Leftarrow$): Assume $D = \{d\}$. Consider some $A \in \mathcal{P}(X) \setminus \{\emptyset\}$. If $d \in A$, then $D \subset A$, and so $A \in \mathcal{F}_D$. If $d \notin A$, then d must be an element of $X \setminus A$, and so $X \setminus A \in \mathcal{F}_D$. As A was chosen arbitrarily, $\mathcal{F}_D$ is an ultrafilter. □

Similar to the idea of principal filters and their construction requiring only a single subset, theorem 4.3 gives us a method to find ultrafilters on any set.

Remark. *There are several different definitions for principal filters; in some texts, principal filters are filters generated by a singleton, which is a special case of our definition.*

For completion, we include the following definition and theorem:

Definition 4.4. A set system $\mathcal{F}$ has the <u>Finite Intersection Property (FIP)</u> if

$$\forall \mathcal{G} \subset \mathcal{F}\,(\,\mathcal{G} \text{ is finite} \Rightarrow \bigcap \mathcal{G} \neq \emptyset\,).$$

It follows directly from definition 4.1 Condition 1 that every filter satisfies the FIP.

We state the following result without proof:

Theorem 4.5. *Let X be a non-empty set and assume that $\mathcal{G} \subset \mathcal{P}(X)$ has the FIP. Then there exists $\mathcal{U} \subset \mathcal{P}(X)$ such that $\mathcal{G} \subset \mathcal{U}$ and $\mathcal{U}$ is an ultrafilter on X.*

For the proof, see [HH99, p. 74]; as ultrafilters generated using the theorem will not be used in this report, it is omitted.

The same reference can be consulted for the following result: it can be shown that every ultrafilter is either generated by a singleton (as in theorem 4.3) or has been constructed using theorem 4.5.

This results concludes this short introduction on filters and ultrafilters. The next section will introduce ideals and, eventually, we will be considering the notions of clubs and stationary sets.

4.2 Clubs and Ideals

The first part of this subsection will cover *ideals*, the dual notion of filters. The theme of duality between them will be recurrent throughout.

Definition 4.6. Let X be a non-empty set and assume $\mathcal{A} \subset \mathcal{P}(X)$. Then we define the dual of $\mathcal{A}$ (which is denoted by $\operatorname{co}(\mathcal{A})$) as

$$\operatorname{co}(\mathcal{A}) := \{X \setminus A : A \in \mathcal{A}\}.$$

That is, the dual of $\mathcal{A}$ constitutes the set of the complements of all elements of $\mathcal{A}$ with respect to the underlying set X.

We immediately see the following:

Lemma 4.7. *The dual of the dual of $\mathcal{A}$ is $\mathcal{A}$ itself.*

Proof. $\operatorname{co}(\operatorname{co}(\mathcal{A})) = \{X \setminus A : A \in \operatorname{co}(\mathcal{A})\} = \{X \setminus A : X \setminus A \in \mathcal{A}\} = \mathcal{A}$. □

Remark. *Recall that if $A \subset X$ we define A^{c} to be the complement of A in X, i.e. $A^{\mathsf{c}} = X \setminus A$. By the definition of the dual of any collection $\mathcal{A}$ of subsets of X, we can write the following:*

$$\operatorname{co}(\mathcal{A}) = \bigsqcup_{A \in \mathcal{A}} \{A^{\mathsf{c}}\}$$

This union is clearly disjoint. Now assume that some set $\mathcal{B}$ is a subset of $\operatorname{co}(\mathcal{A})$*. Then, by the previous remark,*

$$\mathcal{B} \subset \operatorname{co}(\mathcal{A}) = \bigsqcup_{A \in \mathcal{A}} \{A^{\mathsf{c}}\}$$

and hence, when we apply the dual operation to both sides, we obtain

$$\operatorname{co}(\mathcal{B}) \subset \mathcal{A}.$$

Hence, under these circumstances, the dual operation differs from the complement operation on sets: whenever $A \subset B^{\mathsf{c}}$, then $B \subset A^{\mathsf{c}}$. This property fails for the dual operation. Similarly, it is clear that $|\operatorname{co}(\mathcal{A})| = |\mathcal{A}|$ for any $\mathcal{A} \subset \mathcal{P}(X)$ (the map $A \mapsto X \setminus A$ is a bijection). This is clearly not true in general for complements of sets.

Definition 4.8. Let X be a set. A set system $\emptyset \neq \mathcal{I} \subset \mathcal{P}(X)$ is called an ideal on X if $\mathcal{I}$ satisfies the following three conditions:

(i) $X \notin \mathcal{I}$

(ii) If A and $B \in \mathcal{I}$ then $A \cup B \in \mathcal{I}$

(iii) If $A \in \mathcal{I}$ and $B \subset A$ then $B \in \mathcal{I}$.

That is, $\mathcal{I}$ is closed under finite unions and any subset of an element of $\mathcal{I}$ is also a member of $\mathcal{I}$.

The following example is mentioned in [Kun80, p. 76]:

Example 20. *If A is an infinite set, then $\mathcal{I}$, the set of all finite subsets of A, is an ideal on A, as can be verified easily: A is infinite and hence not an element of $\mathcal{I}$, the cardinality of the union of two finite sets is finite, and, thirdly, every subset of a finite set is finite.*

Remark. *As for filters (whose elements, as we recall, can be considered as somehow "large"), we can interpret elements of ideals as "small" in the given set; this notion is very much in line with the previous example in which we can interpret "small" elements as "elements with small cardinality".*

Definition 4.8 above bears a striking (and by no means coincidental) resemblance to the notion of filters which we have already encountered. In fact, both definitions are tightly interlaced by duality. The following result illustrates this connection:

Proposition 4.9. *Let X be a set and let $\mathcal{F}$ be a filter on X. Then the dual of $\mathcal{F}$ is an ideal on X.*

Proof. We verify the definition:

- By definition of duals and of $\mathcal{F}$, $\emptyset \notin \mathcal{F}$ and so $X \notin \operatorname{co}(\mathcal{F})$.
- Consider $A, B \in \operatorname{co}(\mathcal{F})$. Then, by definition, $X \setminus A$ and $X \setminus B$ are elements of $\mathcal{F}$. As $\mathcal{F}$ is closed under intersection, $(X \setminus A) \cap (X \setminus B)$ is in $\mathcal{F}$. We can rewrite this as $X \setminus (A \cup B)$ by de Morgan's law. Again applying the definition of the dual, it follows that $A \cup B \in \operatorname{co}(\mathcal{F})$, which proves closure under taking unions.
- Let $A \in \operatorname{co}(\mathcal{F})$ and let $B \subset A$. Then $X \setminus A \in \mathcal{F}$, and so $X \setminus B \supset X \setminus A$. As filters are closed under taking supersets, $X \setminus B \in \mathcal{F}$, and so $B \in \operatorname{co}(\mathcal{F})$, as required.

As we have verified the definition, the proof is complete. □

Are there more analogues between filters and ideals? That is, is it possible to translate further results we have derived about filters into the world of ideals? Using the duality between the notions yields plenty more results of little mathematical sophistication that link filters to ideals. One main definition we have used in previous sections is that of principal filters. Indeed, it is possible to define principal ideals in a very similar fashion.

Definition 4.10. Let D be a proper subset of X. Then

$$\mathcal{I}_D = \{Y \subset X : Y \subset D\}$$

is called the principal ideal on X generated by D.

Of course, we need to verify that $\mathcal{I}_D$ is in fact an ideal.

Proposition 4.11. *Let D be a proper subset of X. Then $\mathcal{I}_D$ as defined above constitutes an ideal on X.*

Proof. We simply verify the definition:

- As we assumed $D \neq X$, it immediately follows that $X \notin \mathcal{I}_D$.
- Suppose A and B are elements of $\mathcal{I}_D$. Then $A \subset D$ and $B \subset D$, and hence $A \cup B \subset B$, which implies that $A \cup B$ is in $\mathcal{I}_D$, as required.
- Now suppose that $A \in \mathcal{I}_D$ and B is a subset of A. Then A is a subset of D and B is a subset of A imply that B is a subset of D, which yields closure under taking subsets.

As we have verified the definition, we can conclude that $\mathcal{I}_D$ is indeed an ideal on X. $\square$

None of the results above should be surprising: De Morgan's Laws allow us to interchange intersections and unions as well as sub- and supersets after applying the complement operation. We can therefore consider ideals as a type of imperfect complementing structure to filters. What this means practically can be described best when considering the following statement: let D be a proper non-empty subset of X. Then $\mathcal{I}_D \cup \mathcal{F}_D = X$ and $\mathcal{I}_D \cap \mathcal{F}_D = D$, as can be verified very easily.

Definition 4.12. Let ξ be an ordinal. We define the order topology with respect to $\in$ as follows: the subsets of the form $\{\eta \in \xi : \eta < \alpha\}$ and $\{\eta \in \xi : \alpha < \eta\}$ (for any $\alpha \in \xi$) as well as finite intersections and arbitrary unions thereof are considered *open*.

In its more general form, the order topology does apply to all ordered sets (and classes) and not just ordinals (since this piece exclusively considers set theory, the definition is given in a more specific fashion, i.e. on the class of ordinals only). A straightforward and well-known example of this topology can be found when applied to the real line: if we take the real numbers $\mathbb{R}$ and endow them with the order topology, we obtain what is known as the *usual* or *Euclidean topology* on $\mathbb{R}$. The basis of the topology is the collection of all open intervals in $\mathbb{R}$.

In the literature, we can find a number of alternative definitions for the order topology on linearly ordered sets/classes. Some of these are equivalent. We show this exemplarily using the following definition below:

Definition 4.13. Let ξ be an ordinal and consider $A \subset \xi$. We call A a closed set in ξ if whenever $\eta < \xi$ and B is a nonempty subset of $A \cap \eta$, then $\sup(B) \in A$.

Remark. *We recall that as B is a set of ordinals, it is clear that $\bigcup B := \sup(B)$, where $\sup(B)$ is, again, an ordinal. For more details on this, see proposition 2.7.*

Proposition 4.14. *Definition 4.12 and definition 4.13 give the same topology.*

Before we give the proof, we provide some examples:

Example 21. *For any cardinal κ, the set $\mathrm{Lim}(\kappa)$ comprising all limit ordinals in κ is a club. We showed unboundedness in example 4. For closedness, consider an ordinal $\alpha < \kappa$ and any non-empty subset $A \subset \alpha \cap \mathrm{Lim}(\kappa)$. Now, any element of A is a limit ordinal, and hence $\sup(A)$ is a limit ordinal, too, by definition of $\sup$. Thus $\mathrm{Lim}(\kappa)$ is closed, as required.*

It is clear that every infinite subset of ω is cofinal and that every finite subset A of ω is closed ($\sup(B) = \max(B)$ for any non-empty subset of A).

Example 22. *It is clear that any subset in ω is closed: assume A is a subset of ω. Take any $k \in \omega$ and consider any $B \subset A \cap k$. Then $\sup(B) = \max(B)$, which is an element of A since B is a subset of A. Thus A is closed in ω, as required.*

It is easy to find sets that are not closed:

Example 23. *Consider ω_1 and define the set of successor ordinals in ω_1,*

$$\mathrm{Succ}(\omega_1) = \{\alpha + 1 : \alpha \in \omega_1\}.$$

Pick any limit ordinal $\xi \in \mathrm{Lim}(\omega_1)$. Then $\xi \cap \mathrm{Succ}(\omega_1)$ is non-empty, but it is obvious that $\sup(\xi \cap \mathrm{Succ}(\omega_1)) = \xi$, which is a limit ordinal and hence not a member of $\mathrm{Succ}(\omega_1)$, as required.

We now proceed with the proof showing the equivalence of our definitions of closedness.

Proof of proposition 4.14. Let ξ be an ordinal and suppose that A is a subset of ξ.

($\Rightarrow$): Assume A is closed in ξ with respect to definition 4.12. Then $\xi \setminus A =: A^{\mathsf{c}}$ is open w.r.t. definition 4.12 and can thus be expressed as the arbitrary union of open sets:

$$A^{\mathsf{c}} = \{\alpha \in \xi : \alpha < \beta\} \sqcup \bigsqcup_{\epsilon \in \kappa} \{\alpha \in \xi : \gamma_\epsilon < \alpha < \delta_\epsilon\}, \tag{$*$}$$

where $\beta, \gamma_\epsilon, \delta_\epsilon \in \mathbf{ON}$ and $\kappa \in \text{Card}$, where we can assume the union is disjoint.

Now fix any $\eta \in \xi$ and consider any non-empty subset $B \subset A \cap \eta$. We can consider two cases:

- If $\max(B)$ exists, then $\max(B) = \sup(B)$ and hence $\sup(B) \in A$, as required.
- If $\max(B)$ does not exist, then $\sup(B)$ is a limit ordinal. Suppose (for contradiction) $\sup(B) \in A^{\mathsf{c}}$. Then there exist $\gamma_{\epsilon'}$ and $\delta_{\epsilon'}$ such that $\sup(B) \in \{\alpha \in \xi : \gamma_{\epsilon'} < \alpha < \delta_{\epsilon'}\}$. As $\sup(B)$ is a limit ordinal and, by definition, the least ordinal greater than the elements of B, the open set $\{\alpha \in \xi : \gamma_{\epsilon'} < \alpha < \delta_{\epsilon'}\}$ must intersect B. Contradiction. Hence $\sup(B) \notin A^{\mathsf{c}}$ and thus $\sup(B) \in A$, as required.

Hence we have shown that closed sets as defined in definition 4.12 are also closed w.r.t definition 4.13.

($\Leftarrow$): Assume the set A is closed in ξ with respect to definition 4.13. Fix any $\gamma \in \xi$ and consider any non-empty subset $B \subset A \cap \gamma$. By the definition, $\sup(B)$ is an element of A. As $\sup(B)$ is, by definition, the least ordinal greater than or equal to the elements of B, and as B was chosen arbitrarily, it follows that A includes its boundary. This is equivalent to closedness with respect to the order topology as defined in definition 4.12.

As we have shown both implications to be true, the proof is complete. □

Due to the equivalence proven above, we will use the following conventions:

- We write $[\alpha, \beta)$ for the set $\{\gamma \in \mathbf{ON} : \alpha \leq \gamma < \beta\}$ as well as for all permutations that include or exclude the boundary.
- Further, we will occasionally call such sets <u>intervals</u>.

Remark. *In the literature, there are definitions of closed subsets of ordinals that are only defined for limit ordinals (cf. [Kun80, p. 77], among others). As we will mainly work with limit ordinals, we will not need pay much attention to successor ordinals, hence we will implicitly follow this convention.*

The following definitions will be crucial when working with stationary sets, one of the main structures of this section:

Definition 4.15. Let X be a set and let κ be an infinite cardinal. A filter $\mathcal{F}$ on X is called <u>κ-complete</u> if for any subset $\mathcal{F}'$ of $\mathcal{F}$ with $|\mathcal{F}'| < \kappa$ the intersection $\bigcap \mathcal{F}'$ is an element of the filter $\mathcal{F}$.

Remark. *Note that κ-completeness of filters implies closedness under intersection of any set system with cardinality less than κ, not equalling κ. This strict inequality is crucial as we shall see later.*

By duality, we can give a very similar definition for ideals:

Definition 4.16. Let X be a set and let κ be an infinite cardinal. An ideal $\mathcal{I}$ on X is called κ-complete if for any subset $\mathcal{I}'$ of $\mathcal{I}$ with $|\mathcal{I}'| < \kappa$ the union $\bigcup \mathcal{I}'$ is an element of the ideal $\mathcal{I}$.

The reader is certainly already familiar with specific examples of this special notion of completeness: when considering σ-fields (also known as σ-algebras), for instance, we see that they are closed under countable unions or, using our newly introduced definition, that every σ-field is ω_1-complete. The same can be said about σ-rings.

It is common to call ω_1-complete set systems σ-complete. When we consider ideals, we may even say σ-ideal for short. We will adopt this nomenclature from now on.

Again, using the connection between ideals and filters, we can show the following:

Proposition 4.17. *Let X be a set and κ an infinite cardinal. Then $\mathcal{F}$ is a κ-complete filter on X if and only if* $\mathrm{co}(\mathcal{F})$ *is a κ-complete ideal on X.*

Proof.
($\Rightarrow$): Assume $\mathcal{F}$ is a κ-complete ideal. From proposition 4.9 we know that $\mathrm{co}(\mathcal{F})$ is an ideal. Hence it suffices to show that the dual of $\mathcal{F}$ is κ-complete. We do this by verifying the definition:

Suppose $\mathcal{I}'$ is a subset of $\mathrm{co}(\mathcal{F})$ and assume that $|\mathcal{I}'| < \kappa$. Then $\mathrm{co}(\mathcal{I}')$ is an element of $\mathcal{F}$ by the definition of the dual. Moreover, $\bigcup \mathrm{co}(\mathcal{I}') \in \mathcal{F}$ as $\mathcal{F}$ is κ-complete. We can rewrite the union of elements of the dual of $\mathcal{I}'$ as follows:

$$\bigcup \mathrm{co}(\mathcal{I}') = \bigcup_{I \in \mathcal{I}'} I^{\mathsf{c}} = \Big(\bigcap_{I \in \mathcal{I}'} I \Big)^{\mathsf{c}} = \Big(\bigcap \mathcal{I}' \Big)^{\mathsf{c}}$$

Note that the equation $\bigcup_{I \in \mathcal{I}'} I^{\mathsf{c}} = (\bigcap_{I \in \mathcal{I}'} I)^{\mathsf{c}}$ holds by De Morgan's laws.

Thus $(\bigcap \mathcal{I}')^{\mathsf{c}} \in \mathcal{F}$. As $(\bigcap \mathcal{I}')^{\mathsf{c}} = X \setminus (\bigcap \mathcal{I}')$, it follows directly that $\bigcap \mathcal{I}' \in \mathrm{co}(\mathcal{F})$, as required.

($\Leftarrow$): This direction is very similar. The proof is omitted. □

We can now draw our attention to one of the main definitions of this section:

Definition 4.18. Let ξ be an ordinal. If $A \subset \xi$ is both closed and cofinal in ξ, then we say that A is a club in ξ.

This is simply a task of relabelling; the word *club* is merely a contraction of the terms *closed* and *unbounded*. Such sets may also be called ξ-clubs, for short.

Theorem 4.19. *Let ξ be an ordinal with uncountable cofinality. Consider a set system $\{C_\gamma : \gamma < \mu\}$ of ξ-clubs. If $\mu < \mathrm{cf}(\xi)$ then the intersection $\bigcap \{C_\gamma : \gamma < \mu\}$ is again a ξ-club.*

In the proof below we use ideas presented by Hajnal and Hamburger in [HH99, p. 147]. Further explanations and details have been added.

Proof. Let ξ be a limit ordinal with uncountable cofinality and assume $\{C_\gamma : \gamma < \mu\}$ with $\mu < \mathrm{cf}(\xi)$ is a system of ξ-clubs. By our proven equivalence of closedness in the order topology, it is clear that $\bigcap_{\gamma < \mu} C_\gamma$ is closed.

To prove unboundedness, we go about as follows: given an arbitrary ordinal $\alpha \in \xi$ we define recursively a countable increasing sequence of ordinals whose supremum is greater than or equal to α. The main focus will lie on the definition of our sequence. Once we are done defining the sequence, we remain to show that the supremum of our sequence is a member of the intersection of our clubs. This will follow quite easily.

Suppose $\alpha \in \xi$. We define our countable sequence of ordinals $\{\alpha_n : n \in \omega\}$ (where $\alpha_n \in \xi$ for each $n \in \omega$) as follows:

- Set $\alpha_0 \coloneqq \alpha$.
- Let $n \in \omega \setminus \{0\}$. We assume that we have defined α_n so that $\alpha_n < \xi$.
  Now we pick a non-empty A_n such that

  $$A_n \subset \xi \setminus \alpha_n.$$

  (Note that this is always possible as $\alpha_n < \xi$ and hence $\xi \setminus \alpha_n \neq \emptyset$.)
  We require the following two conditions to be satisfied by A_n:

  (1) We must have $|A_n| < \mathrm{cf}(\xi)$.
  (2) Further, $C_\gamma \cap A_n$ must be non-empty for each $\gamma \in \mu$.

  Such a set A_n exists. We show the existence in reverse order: for (2), note that C_γ is a ξ-club for each $\gamma \in \mu$. Hence, in particular, it is ξ-unbounded. By our assumption, α_n is strictly less than ξ, and hence bounded in ξ, and thus

  $$C_\gamma \setminus \alpha_n \neq \emptyset \qquad \text{for every } \gamma \in \mu.$$

  As A_n is a subset of $\xi \setminus \alpha_n$, we can use a choice function and define A_n to contain at least one element from each C_γ. This verifies (2). As $\mu < \mathrm{cf}(\xi)$, we see that, if we define A_n as described, then $|A_n| \leq \mu < \mathrm{cf}(\xi)$, which verifies (1).
  Now define

  $$\alpha_{n+1} \coloneqq \sup(A_n) + 1.$$

  We see that $\alpha_{n+1} < \xi$ as $|A_n| < \mathrm{cf}(\xi)$. Informally, this is true since α_{n+1} will never attain ξ as A_n is not large enough (i.e. bounded in ξ).
- Now we consider the sequence $\{\alpha_n : n \in \omega\}$; it is completely defined by the rules given above.

Clearly, this is an increasing sequence (this follows since $\alpha_{n+1} \in \xi \setminus \alpha_n$).

To finish the proof, we define

$$\tau = \sup\{\alpha_n : n \in \omega\}.$$

In order to progress, we need to recall our initial assumptions: we supposed that $\mathrm{cf}(\xi)$ is uncountable. Therefore the above α_n-sequence cannot be cofinal in ξ, and hence $\tau < \xi$. Further, again by definition of our sequence, τ is bounded below by α. Thus

$$\alpha \leq \tau < \xi.$$

Recall that we are trying to prove cofinality of the set $\bigcap_{\gamma\in\mu} C_\gamma$ in ξ. We were given an arbitrary α and have now found an element $\tau \in \xi$ (which depends on α) greater than or equal to α, which will verify cofinality provided we can show that $\tau \in \bigcap_{\gamma\in\mu} C_\gamma$.

But this is easy: note that for any C_γ and for any $n \in \omega$, the set

$$C_\gamma \cap (\alpha_{n+1} \setminus \alpha_n)$$

is non-empty. This is due to the fact that we defined A_n to be a subset of $\xi \setminus \alpha_n$ (i.e. one could say A_n is "greater than α_n") and we required A_n to intersect every C_γ (this is condition (2) above). Further, α_{n+1} is defined to be greater than $\sup(A_n)$. Hence, in particular,

$$\alpha_n < \beta < \alpha_{n+1} \qquad \text{for every } \beta \in A_n$$

and thus

$$A_n \subset \alpha_{n+1} \setminus \alpha_n.$$

Hence the intersection of C_γ and $\alpha_{n+1} \setminus \alpha_n$ is indeed non-empty for every $\gamma \in \mu$ and for every $n \in \omega$. In particular, as $\tau = \bigcup_{n \in \omega} \alpha_n$, we hence have

$$C_\gamma \cap \tau \neq \emptyset.$$

As τ is a limit ordinal and C_γ is unbounded in $\xi > \tau$, we must have

$$\sup(C_\gamma \cap \tau) = \tau.$$

As C_γ is a club (and hence closed) for every $\gamma \in \mu$, it follows from the definition that $\sup(C_\gamma \cap \tau) \in C_\gamma$. Thus $\tau \in C_\gamma$ for every $\gamma \in \mu$, and thus

$$\tau \in \bigcap_{\gamma \in \mu} C_\gamma,$$

as required. □

Remark. *One might wonder why we are required to impose the assumption of uncountable cofinality upon ξ. Obviously, it was indispensable in the proof when we derived that $\tau < \xi$ (we would not be able to deduce this fact if $\mathrm{cf}(\xi)$ were countable). Secondly, a counterexample is readily found: consider the cofinal sequences $(0, 2, 4, 6, \ldots)$ and $(1, 3, 5, \ldots)$, both in ω. They are clearly disjoint, but as they are unbounded they are also closed (by example 22).*

For an ordinal ξ meeting the requirements of the theorem above, we are required to consider a set system of clubs with cardinality less than $\mathrm{cf}(\xi)$ in order to guarantee closedness, as the following counterexample verifies (the particular case $\xi = \omega_1$ can be found in Prikry and Baumgartner's paper [BP77, p. 110]): consider a cardinal κ with uncountable cofinality. Define sets $C_\alpha = [\alpha, \kappa)$ for each $\alpha < \kappa$. Note that $|\{C_\alpha : \alpha < \kappa\}| = \kappa$. Clearly, each C_α is a club, but the intersection $\bigcap_{\alpha<\kappa} C_\alpha$ is empty.

The techniques used in the proof of theorem 4.19 can be found in various other instances: Baumgartner and Prikry give a very elegant proof of unboundedness for the special case $\kappa = \omega_1$. Details can be found in [BP77, p. 110].

4.3 Stationary Sets

We now proceed with the introduction of stationary sets. Stationary sets play a major role in combinatorial set theory due to their applicability. Results such as *Fodor's theorem*, which we shall investigate later, illustrate the theory's importance. Moreover, Fodor's theorem will be very useful as it simplifies several proofs of problems in infinitary combinatorics. Last but not least, by providing a purely combinatorial proof of *Silver's theorem* with its link to the Generalised Continuum Hypothesis, we provide an example of the theory's versatility.

Definition 4.20. Let ξ be an ordinal. Then we define the set system

$$\mathcal{C}(\xi) := \{A \subset \xi : \exists B \subset A (B \text{ is a club in } \xi)\}$$

to be the collection of so called ξ-large sets.

Let ξ be an ordinal. Then $\mathcal{C}(\xi)$ denotes the collection of all sets that contain a club in ξ. It is clear that if A is a club in ξ, then any superset of A that is itself a subset of ξ (and hence A itself) is a member of $\mathcal{C}(\xi)$. This illustration justifies the terminology: any set in $\mathcal{C}(\xi)$ is unbounded (which could be interpreted as "somewhat large in ξ") and is closed in ξ (which could be interpreted as "has a similar limit behaviour to ξ"). This is kept rather vague on purpose; we will examine the behaviour of clubs in detail in due course and refer to these informal descriptions en route.

In order to begin our investigation, we can pose the following question: what can we say about $\mathcal{C}(\xi)$ in general?

Example 24. *For any ordinal ξ, it is clear that ξ is a club in ξ, and hence we see immediately that $\xi \in \mathcal{C}(\xi)$. Thus, $\mathcal{C}(\xi)$ is non-empty.*

The next result follows almost immediately from theorem 4.19:

Proposition 4.21. *Let ξ be an ordinal. If ξ is a limit ordinal and $\operatorname{cf}(\xi)$ is uncountable, then $\mathcal{C}(\xi)$ is a $\operatorname{cf}(\xi)$-complete filter.*

Proof. It suffices to verify the definition of filter and $\operatorname{cf}(\xi)$-complete, respectively.

- It is clear that the empty set does not contain a ξ-club, hence $\emptyset \notin \mathcal{C}(\xi)$.
- We can apply theorem 4.19 and obtain the required $\operatorname{cf}(\xi)$-completeness.
- Assume $A \in \mathcal{C}(\xi)$ and consider $B \supset A$. As A contains a ξ-club, so does its superset B, and hence $B \in \mathcal{C}(\xi)$.

Hence the proof is complete. □

Remark. *We could easily drop the assumption that ξ is a limit ordinal as $\operatorname{cf}(\xi)$ is uncountable only if ξ is a limit ordinal. This was one of the results in section 3.1.*

We now have all the tools we need in order to present the main definition of the section:

Definition 4.22. Let ξ be an ordinal. We call a set $A \subset \xi$ a ξ-stationary set if A intersects every ξ-club.

Using our previous definitions of $\mathcal{C}(\xi)$, we can show the following equivalence:

Lemma 4.23. *A subset $A \subset \xi$ is ξ-stationary if and only if $\xi \setminus A$ is not ξ-large.*

The proof is straightforward:

Proof.
($\Rightarrow$): If A is stationary and $\xi \setminus A \in \mathcal{C}(\xi)$, then, by definition, there exists a subset $C \subset \xi \setminus A$ that is a club. But as $C \subset \xi \setminus A$, it cannot intersect A. Hence A does not intersect all ξ-clubs. Contradiction.
($\Leftarrow$): Immediate. □

Proposition 4.24. *The intersection of a stationary set with a club is again stationary.*

The proof is simple.

Proof. Let ξ be an ordinal. Let A be a stationary subset of ξ and consider any ξ-club C. We need to show that for any ξ-club C' we have $C' \cap (C \cap A)$ is non-empty. But this is easy: we write

$$C' \cap (C \cap A) = (C' \cap C) \cap A,$$

and now apply theorem 4.19 to $C' \cap C$. Hence it is clear that $C \cap C'$ is a club, and thus the result follows from the definition of the stationary set A. □

Example 25. *From the definition, it is clear that every cofinite subset of ω is stationary; clearly, any infinite subset of ω is also unbounded. We have shown before that every (cofinal) subset of ω is closed, hence the result follows.*

This example is, in fact, no coincidence as the following result shows (its statement without proof can be found in [BP77, p. 111]).

Proposition 4.25. *Every stationary set is unbounded.*

Proof. Let ξ be an ordinal and assume a subset $A \subset \xi$ is ξ-stationary. If A is bounded, then there exists $\beta \in \xi$ such that $\alpha < \beta$ for all $\alpha \in A$. Now consider the interval $[\beta, \xi)$. This is clearly a subset of ξ and both closed and unbounded in ξ; hence it is a club in ξ. But by its construction, it is disjoint from A. Hence A does not intersect every ξ-club. Contradiction. □

As Baumgartner and Prikry suggest in [BP77, p. 111], we can interpret κ-clubs as very large sets. In contrast, κ-stationary sets may be considered as fairly large.

The following result is stated without proof in [HH99, p. 148]. We give the proof below.

Proposition 4.26. *Let ξ be an ordinal and assume that $\mathrm{cf}(\xi) \in \{1, \omega\}$. Then a subset $A \subset \xi$ is ξ-stationary if and only if $\xi \setminus A$ is not cofinal in ξ.*

Proof.

($\Rightarrow$): Suppose that $A \subset \xi$ is ξ-stationary.

If $\mathrm{cf}(\xi) = 1$, then ξ is a successor ordinal. Denote its last element by β. Since A is ξ-stationary, it intersects all clubs (by definition) and is unbounded (by proposition 4.25), i.e. $\beta \in A$. Hence it follows directly that $\xi \setminus A$ does not contain β. As β is the last element of ξ, we obtain that $\xi \setminus A$ is not cofinal.

If $\mathrm{cf}(\xi) = \omega$, then for contradiction assume $\xi \setminus A$ is unbounded. As A is ξ-stationary, it is also unbounded (by proposition 4.25). Hence assume that

$$\Theta := \{\theta_n : n \in \omega\}$$

is a cofinal sequence in A (existence follows from the definition of cofinality). We show that there is a club in $\xi \setminus A$, which contradicts the fact that A is ξ-stationary. We define the following sequence inductively:

- Set
$$\lambda_0 := \min((\theta_{s+1} \setminus \theta_s) \cap (\xi \setminus A)),$$
where $s \in \omega$ is the least cardinal for which the set $(\theta_{s+1} \setminus \theta_s) \cap (\xi \setminus A)$ is non-empty. (Such an s exists as $\xi \setminus A$ is unbounded.)

- Now assume λ_n exists and define

  $$\lambda_{n+1} \coloneqq \min((\theta_{t+1} \setminus \theta_t) \cap (\xi \setminus (A \cup \lambda_n))),$$

  where, as in the case above, $t \in \omega$ is defined to be the least cardinal such that the set $(\theta_{t+1} \setminus \theta_t) \cap (\xi \setminus (A \cup \lambda_n))$ is non-empty. (For the same reason as above, such a t must exist.)

- Now consider

  $$\Lambda \coloneqq \{\lambda_n : n \in \omega\}.$$

  It is clear that Λ is both strictly increasing and unbounded.

However, it follows easily that Λ is also closed: as $\mathrm{cf}(\xi) = \omega$ and Θ is increasing, any subset $C \subset \Lambda \cap \eta$ for any $\eta < \xi$ must be finite since Λ is also unbounded. Hence $\sup(C) = \max(C) \in \Lambda$, which proves closedness of Λ. But now Λ is closed, unbounded, and disjoint from the ξ-stationary set A, which is a contradiction.

($\Leftarrow$): Assume that $\mathrm{cf}(\xi) = 1$ and suppose $\xi \setminus A$ is bounded in ξ. Then clearly A is unbounded in ξ, and hence it includes the last element of ξ. Thus A intersects every unbounded subset of ξ and thus every club. Clearly, the same reasoning holds for the case $\mathrm{cf}(\xi) = \omega$, too: here, note that if we assume A to be non-stationary, then the bounded set $\xi \setminus A$ contains a club, which yields an immediate contradiction. □

Remark. *It is clear that, if ξ is an ordinal, then for any $\eta < \xi$, the set $\xi \setminus \eta$ contains a club (the closed interval $[\eta, \xi)$, for example). This trivial example can be found in [HH99, p. 146]. However, the previous proposition visualises a quirk of ordinals with countable cofinality: if ξ is of countable cofinality, then we can find disjoint ξ-clubs A_1 and A_2, as proven above. Hence the notion of clubs (and hence of stationary sets) is not meaningful for sets with countable cofinality.*

We can generalise the previous proof as follows (the result, without proof, can be found in [Kun80, p. 77]):

Proposition 4.27. *Let ξ be a limit ordinal. If $\mathrm{cf}(\xi) = \omega$, then any ξ-unbounded sequence of cardinality ω is closed in ξ.*

Proof. Assume $(\gamma_\beta)_{\beta \in \omega}$ is a ξ-unbounded sequence and suppose w.l.o.g. that the sequence is strictly increasing. Define $A \coloneqq \{\gamma_\beta : \beta < \omega\}$. Take any $\eta < \xi$ and consider any non-empty subset B of $A \cap \eta$. Note that $|A \cap \eta| < \omega$, and thus B is finite. Therefore $\sup(B) = \max(B)$, and as B is, in particular, a subset of A, we have that $\sup(B) \in A$, as required. □

Remark. *In theorem 3.19, we showed that the cofinality of any ordinal is in fact a cardinal. Why do we not write $\aleph_0$ then instead of ω whenever we refer to the cofinality of a set with countable cofinality, for example? Although it sounds tempting to do so, this is a slippery slope. The following example serves as an warning for why this is the case:*

Consider the ordinal ω_ω. Clearly $\mathrm{cf}(\omega_\omega) = \omega$. Note that $A \coloneqq \{2, 4, 6, \ldots, \omega_2, \omega_4, \omega_6, \ldots\}$ is a countable subset of ω_ω. Further, A is ω_ω-unbounded. But consider the following: if we take any infinite subset of $\{2, 4, 6, 8, \ldots\}$, then its limit is ω, which is not an element of A. Hence A is not closed, and we appear to have given a counterexample to the previous proposition.

Careful: by conflating the notation (and hence the concepts of cardinal and ordinal numbers), we put ourselves in treacherous waters. In this case we have considered the *cardinality of A rather than the order type, which must be considered when we aim to apply proposition 4.27. Indeed, although A is clearly infinitely countable (and hence has cardinality $\aleph_0$), it is not of order type ω, but of order type $\omega \times 2$. Therefore proposition 4.27 is not applicable in the first place.*

The remaining results in this section build up to Fodor's theorem, an interesting theorem whose proof puts our acquired knowledge about stationary sets into practice.

We will begin by giving details regarding notation: let ξ be an ordinal.

- The set of all stationary sets of ξ is denoted by $\mathrm{Stat}(\xi)$.
- The dual of the set of ξ-large sets, $\mathrm{co}(\mathcal{C}(\xi))$, will be denoted by $\mathrm{NS}(\xi)$. (If NS is remembered as "Not Stationary", this is a very natural shorthand. We shall refer to such sets by the term ξ-nonstationary sets.)

Finally, after introducing the notions above, we are now ready to exploit their versatility. The following seminal theorem is due to Fodor (1956), we will introduce it here and give the proof later – a few notions need to be learned first.

Definition 4.28. Assume f is a function whose domain and range are subsets of **ON**. Then we call f a regressive function if for all $\alpha \in \mathrm{dom}(f)$ we have $f(\alpha) < \alpha$.

For obvious reasons, regressive functions are also called pressing-down functions.

In the literature, sometimes the condition that if $0 \in \mathrm{dom}(f)$ then $f(0) = 0$ is included in the definition for completeness.

Theorem 4.29 (Fodor's theorem)**.** *Suppose κ is an uncountable regular cardinal and assume that $S \in \mathrm{Stat}(\kappa)$. If $f\colon S \to \kappa$ is a regressive function, then there exists $\alpha \in \kappa$ such that $f^{-1}(\{\alpha\}) \in \mathrm{Stat}(\kappa)$.*

Intuitively, note that this is the same as saying there exists a stationary set in κ on which the regressive function f is constant (in fact it has value α). Additionally, note that $f^{-1}(\{\alpha\}) \subset S$, which says that S contains such a suitable stationary set.

We give a proof following Kunen's approach (cf [Kun80, p. 80]). An alternative proof can be found in [HH99, p. 151].

The following notion will be required:

Definition 4.30. Let ξ be an ordinal. If $(A_\alpha)_{\alpha<\xi}$ is a sequence of subsets of ξ, then we define the diagonal intersection of $(A_\alpha)_{\alpha<\xi}$ by

$$\begin{aligned}\Delta_{\alpha<\xi} A_\alpha &\coloneqq \left\{ \alpha < \xi : \alpha \in \bigcap_{\beta<\alpha} A_\beta \right\} \\ &= \{\alpha < \xi : \forall \beta < \alpha \; (\alpha \in A_\beta)\}.\end{aligned}$$

Intuitively, an ordinal α is in the diagonal intersection if and only if it is a member of the first α-many elements of the sequence.

For completeness, Hajnal and Hamburger also require each element of the diagonal intersection to be non-zero.

Remark. *Note that we require a sequence of subsets of ξ, not just a set. The order of the elements as determined by the sequence defines the diagonal intersection, as is obvious from the definition.*

Lemma 4.31. *Fix a regular uncountable cardinal κ and let $(A_\alpha)_{\alpha<\kappa}$ be a sequence of κ-clubs. Then $\Delta_{\alpha<\kappa}A_\alpha$ is also a κ-club.*

In the following proof, we will confine ourselves to showing closedness only; that proof is due to Hajnal and Hamburger in [HH99, p. 149]. We extend the proof by adding further comments and explanations. The reasoning proving unboundedness is very similar to the proof of theorem 4.19. We omit the details.

Proof. We show that the diagonal intersection is closed. We write

$$\Delta A := \Delta_{\alpha<\kappa}A_\alpha$$

for brevity.

We verify the definition of closed sets: fix an ordinal $\eta < \kappa$ and consider a nonempty subset $B \subset \eta \cap \Delta A$. We need to show that $\sup(B)$ is an element of ΔA. As before, if $\sup(B) = \max(B)$ then we are trivially done. Hence assume $\sup(B)$ is a limit ordinal, and write

$$\sup(B) = \xi.$$

Fix any ordinal $\alpha < \xi$. As ξ is a limit ordinal, it cannot be attained by the successor operation, and hence $\alpha + 1 < \xi$. Note that we may rewrite

$$\sup(B) = \sup(B \setminus (\alpha + 1))$$

as $B \setminus (\alpha + 1)$ is non-empty and since $\sup(B) > \alpha + 1$. We now use the fact that each element of ΔA is closed: by definition,

$$B \setminus (\alpha + 1) \subset A_\alpha \cap \eta.$$

We verify this explicitly:

<u>Claim:</u> The set $B \setminus (\alpha + 1)$ is a subset of $A_\alpha \cap \eta$.
<u>Proof:</u> As $\alpha + 1 < \sup(B)$ and $B \subset \Delta A \cap \eta$, it suffices to show that $B \setminus (\alpha + 1) \subset A_\alpha$. Assume $\gamma \in B \setminus (\alpha + 1)$. By definition, $\alpha + 1 \leq \gamma < \eta$ and γ is an element of the first γ-many sets A_α (this holds since $\gamma \in \Delta A \cap \eta$ and hence follows from the definition of the diagonal intersection). Hence, in particular, $\gamma < \alpha$, and hence $\gamma \in A_\alpha$, as required. ■

However, by assumption, A_α is closed. Hence, in particular,

$$\sup(B \setminus (\alpha + 1)) = \sup(B) \in A_\alpha.$$

As α was chosen arbitrarily, we may deduce that

$$\sup(B) \in \bigcap_{\alpha<\kappa} A_\alpha,$$

and so it is easily seen that

$$\sup(B) \in \Delta A$$

as required (clearly, if $\sup(B) = \xi$ and $\xi \leq \eta < \kappa$ is an element of all sets A_α, then, in particular, it a member of the first ξ-many). This proves the required closedness.

As mentioned above, unboundedness follows from a reasoning almost identical to the proof of theorem 4.19. Details are omitted, proofs can be found in [HH99, pp. 149-150] and [Kun80, p. 80]. □

We are now ready to proceed with the proof of Fodor's theorem:

Proof of Fodor's Theorem. We aim for a contradiction and hence assume the statement is false. Hence, with f, S and κ given as described, every set $f^{-1}(\{\alpha\})$ is a member of $\mathrm{NS}(\kappa)$. As we would like to use the previous lemma, define a sequence $(D_\alpha)_{\alpha<\kappa}$ of κ-clubs such that

$$D_\alpha \cap f^{-1}(\{\alpha\}) = \emptyset$$

for every $\alpha < \kappa$ (i.e. no element of D_α is mapped to α by f). Now consider the diagonal intersection of the sets D_α. By the previous lemma, $\Delta_{\alpha<\kappa} D_\alpha$ is a club. As S is stationary we have $\Delta_{\alpha<\kappa} D_\alpha \cap S \neq \emptyset$. But note that if an ordinal $\delta \in \Delta_{\alpha<\kappa} D_\alpha \cap S$, then δ is by definition of the diagonal intersection an element of D_γ for all $\gamma < \delta$. By definition of D_γ, we have $f(\delta) \neq \gamma$ for all $\gamma < \delta$. But now $f(\delta) \not< \delta$, which contradicts the fact that f is regressive. As we have arrived at a contradiction, the proof is complete. □

The following corollary will be very useful in the proof of Silver's theorem given in the next section. As we shall only use the result in the proof of Silver's theorem (which, in our case, concerns ω_1 only), we shall give a less general result than before.

First, we extend the definition of regressive functions: if f and g are functions whose domain and range are subsets of **ON** and whose domains coincide, then we say that f is g-regressive if $f(\alpha) < g(\alpha)$ for all $\alpha \in \mathrm{dom}(f)$.

Both the original result and the proof can be found in [BP77, pp. 111-2].

Corollary 4.32. *Suppose $S \subset \omega_1$ is stationary. Let g be the function on S defined by $g(\alpha) = \omega_\alpha$. If there is a function f on S that is g-regressive, then there exists an ordinal $\gamma < \omega_1$ as well as a stationary set $S' \subset S$ such that $f(\alpha) < g(\gamma) = \omega_\gamma$ for all $\alpha \in S'$.*

Proof. The proof will follow quite easily from results we have proven already. Consider $\mathrm{Lim}(\omega_1)$. This is a club by example 21. Given S and f as above, fix any $\alpha \in S \cap \mathrm{Lim}(\omega_1)$. Then there exists some $\beta < \alpha$ for which $f(\alpha) < \omega_\beta$. This holds as, by assumption, $f(\alpha) < \omega_\alpha$ and since α is a limit ordinal we have $\omega_\alpha = \sup(\omega_\beta : \beta < \alpha)$. In particular, there exists a least such $\beta < \alpha$ for which $f(\alpha) < \omega_\beta$; we denote this ordinal by $h(\alpha)$.

By definition, note that h is regressive (we defined h such that $h(\alpha) = \min(\beta < \alpha : f(\alpha) < \omega_\beta)$, which always exists as f is regressive and as α is a limit ordinal). Further, h maps to ω_1, which is regular and uncountable. By proposition 4.24, we also see that $S \cap \mathrm{Lim}(\omega_1)$ is in fact stationary, too. But now $\mathrm{dom}(h) = S \cap \mathrm{Lim}(\omega_1)$ is stationary. Hence Fodor's theorem is applicable, and we obtain a stationary set $S' \subset S \cap \mathrm{Lim}(\omega_1)$ as well as an ordinal γ such that $h^{-1}(\{\gamma\}) = S'$, that is, h is constant on S' with value γ. Now, if $\delta \in S'$, then

$$\begin{aligned} \gamma &= h(\delta) \\ &= \min(\beta < \delta : f(\delta) < \omega_\beta) \end{aligned}$$

and hence, by definition of h,

$$f(\delta) < \omega_\gamma.$$

Hence the result is proven. □

4.4 Silver's Theorem

After introducing the notion of stationary sets, we would like to use these newly introduced (and immensely powerful) notions and put them into practice; we will use them in order to prove a vital theorem due to Jack Silver that allows us to draw strong conclusions of the continuum function.

In this section, we closely follow Baumgartner and Prikry's presentation from [BP77]; all of the proofs presented here are taken from aforementioned source.

Let β be an ordinal. For ease of notation let $\mathrm{CH}(\beta)$ be the statement

$$\forall \alpha < \beta (2^{\aleph_\alpha} = \aleph_{\alpha+1}).$$

Theorem 4.33 (Silver's Theorem on ω_1)**.**

$$\mathrm{CH}(\omega_1) \Rightarrow \mathrm{CH}(\omega_1 + 1)$$

This result is indeed very surprising! Before this result was found by Silver, set theorists were convinced that a singular cardinal is most likely to be the first one for which GCH fails. Silver's theorem has proved this assumption wrong.

In order to progress to the proof, we require several lemmata:

Lemma 4.34. *Assume that ω_1 is a countable union of sets, which can be written as*

$$\omega_1 = \{A_n : n \in \omega\}.$$

Then there exists $n \in \omega$ such that A_n is stationary.

We consider the proof from [BP77, p. 111]; the argument is straightfroward.

Proof. In view of a contradiction, assume none of the sets A_n is stationary. Thus, there exists a club C_n for each A_n such that $C_n \cap A_n = \emptyset$. But by theorem 4.19 (which is applicable as ω_1 has uncountable cofinality), we see that the clubs C_n are closed under taking intersections, and hence $\bigcap_{n\in\omega} C_n$ is also an ω_1-club, and in particular non-empty. Now, by definition, $\bigcap_{n\in\omega} C_n$ contains all those elements from ω_1 that are disjoint from all A_n, and, at the same time, $\bigcap_{n\in\omega} C_n$ is non-empty. The fact that the union of the sets A_n equals ω_1 yields the required contradiction. □

We are now ready to give the proof of Silver's theorem:

Proof of Silver's Theorem on ω_1. The proof will be of a very direct nature. We will show that

$$|\mathcal{P}(\omega_{\omega_1})| = \aleph_{\omega_1+1}.$$

The proof comes in two parts. Firstly, we use the assumptions to list the subsets of ω_α indexed by ordinals less than $\omega_{\alpha+1}$. We then define functions f_A for subsets $A \subset \omega_{\omega_1}$ which return the index ξ for which $A \cap \omega_\alpha$ (informally, A restricted to ω_α) appears in the sequence of subsets of ω_α. We then define a relation on the subsets ω_{ω_1} that will allow us to obtain stationary sets defined by functions f_A and f_B.

In the latter half, we argue by contradiction and, using the previously derived results and the theory of stationary sets, show the existence of stationary sets which we have already shown not to be unbounded. This will yield the required contradiction.

We begin by considering $\mathcal{P}(\aleph_{\omega_1})$ in detail. Using the fact that $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for all $\alpha < \omega_1$, we consider the set

$$E_\alpha = \left\{ A^\alpha_\xi : \xi < \omega_{\alpha+1} \right\}$$

as the set of all subsets of ω_α without repititions. Technically, this is of course not different from $\mathcal{P}(\omega_\alpha)$, but by assumption, we are able to enumerate the elements using the ordinal $\omega_{\alpha+1}$. This will be crucial later on.

Fix any subset $A \subset \omega_{\omega_1}$. We define f on ω_1 such that

$$f_A(\alpha) = \xi \qquad \text{if and only if} \qquad A \cap \omega_\alpha = A^\alpha_\xi.$$

<u>Claim:</u> Let A, B be subsets of ω_{ω_1} and suppose $A \neq B$. Then the set $\{\alpha : f_A(\alpha) = f_B(\alpha)\}$ is bounded.

<u>Proof:</u> As $A \neq B$, there exists an $\alpha < \omega_1$ such that $A \cap \omega_\alpha \neq B \cap \omega_\alpha$ (if there were no such α then $A \cap \omega_{\omega_1} = B \cap \omega_{\omega_1}$, which in turn implies $A = B$ as both A and B are subsets of ω_{ω_1}; this contradicts the assumption $A \neq B$). Now assume $\beta \geq \alpha$. As $\omega_\alpha \subset \omega_\beta$, we have $A \cap \omega_\beta \neq B \cap \omega_\beta$ which implies, by definition, that $f_A(\beta) \neq f_B(\beta)$, as required. ■

We now introduce a relation on the subsets of ω_{ω_1} which will enable us to exploit the properties of f we have just derived. For subsets $A, B \subset \omega_{\omega_1}$, define

$$A \lessdot B$$

if and only if the set $\{\alpha : f_A(\alpha) < f_B(\alpha)\}$ is stationary. Note that if A and B are as above, then we always have either $A \lessdot B$ or $B \lessdot A$: as f is defined on ω_1 we may write

$$\omega_1 = \{\alpha : f_A(\alpha) < f_B(\alpha)\} \cup \{\alpha : f_A(\alpha) > f_B(\alpha)\} \cup \{\alpha : f_A(\alpha) = f_B(\alpha)\},$$

which is clearly a partition of ω_1 (i.e. in particular, the three sets on the right hand side are disjoint). Further, lemma 4.34 tells us that at least one of the sets must be a stationary sets. Lastly, the above claim shows that $\{\alpha : f_A(\alpha) = f_B(\alpha)\}$ is bounded and hence, by the contrapositive of proposition 4.25, not stationary.

In this second half of the proof, we proceed by assuming

$$|\mathcal{P}(\omega_{\omega_1})| > \aleph_{\omega_1+1}. \tag{$*$}$$

We aim for a contradiction.

We introduce the following notation: if $A \subset \omega_{\omega_1}$, then we write

$$R^{-1}_{\lessdot}(B) := \{A \subset \omega_{\omega_1} : A \lessdot B\}.$$

Now, consider the following claim (which makes fundamental use of our relation $\lessdot$):

<u>Claim:</u> Assuming ($*$) above, there exists an ω_{ω_1}-subset $B \in \mathcal{P}(\omega_{\omega_1})$ such that $R^{-1}_{\lessdot}(B)$ has cardinality at least $\aleph_{\omega_1+1}$.

<u>Proof:</u> Note that the set $R^{-1}_{\lessdot}(B)$ is really a set of subsets of ω_{ω_1} (i.e. those subsets A for which $A \lessdot B$). Hence consider $X \subset \mathcal{P}(\omega_{\omega_1})$. Further, assume $|X| = \aleph_{\omega_1+1}$. Clearly, if there exists a ω_{ω_1}-subset $B \in X$ such that the claim holds, we are trivially done. Thus, assume there is no such $B \in X$.

Consider the union

$$\begin{aligned}Y &:= \bigcup\{R_{\lessdot}^{-1}(B) : B \in X\}\\ &= \bigcup\{\{A : A \lessdot B\} : B \in X\}\\ &= \{A : A \lessdot B \text{ for some } B \in X\}.\end{aligned}$$

That is, Y is the set of all those $A \in \mathcal{P}(\omega_{\omega_1})$ for which $A \lessdot B$ for some $B \in X$. As $|\omega_{\omega_1}| = \aleph_{\omega_1}$, each subset of ω_{ω_1} has at most cardinality $\aleph_{\omega_1}$. Hence it is clear that

$$|Y| \leq \max(\aleph_{\omega_1}, \aleph_{\omega_1+1}) = \aleph_{\omega_1+1}$$

(there are at most $\aleph_{\omega_1+1}$ elements of cardinality at most $\aleph_{\omega_1}$ each). But by our initial assumption, we see that $|\mathcal{P}(\omega_{\omega_1})| > \aleph_{\omega_1+1}$, and hence there exists $B' \subset \omega_{\omega_1}$ such that $B' \notin Y$. Pick any such B' and any $A \in X$. Then it follows that $B' \notin R_{\lessdot}^{-1}(A)$. Indeed, for contradiction assume that $B' \notin Y$ and $B' \in R_{\lessdot}^{-1}(A)$. By definition, we then have $B' \lessdot A$. However, by definition again, $B' \in Y$, and a contradiction arises.

By the reasoning above, we see that $B' \not\lessdot A$ implies that $A \lessdot B'$ must hold. But A was chosen arbitrarily. Hence $A \lessdot B'$ for all $A \in X$, and thus B' is a subset of ω_{ω_1} satisfying the properties (as $\aleph_{\omega_1+1} = |X| \leq |\{A : A \lessdot B\}|$), as required. ■

We now fix a subset $B \subset \omega_{\omega_1}$ which satisfies the above claim of

$$|\{A \subset \omega_{\omega_1} : A \lessdot B\}| \geq \aleph_{\omega_1+2}.$$

Consider an ordinal $\alpha < \omega_1$. Now, by definition of E_α, we have $f_B(\alpha) < \omega_{\alpha+1}$. Hence, it follows that $|\{\beta : \beta < f_B(\alpha)\}| \leq \aleph_\alpha$. Thus, we may define an injective function g_α from $\{\beta : \beta < f_B(\alpha)\}$ into ω_α.

Assume $A \lessdot B$ (recall how we defined B above). Further, define

$$S_A = \{\alpha : f_A(\alpha) < f_B(\alpha)\} \subset \omega_1$$

and note that, by definition of $\lessdot$, the set S_A is stationary. Fix an element $\alpha \in S_A$. Now,

$$f_A(\alpha) < f_B(\alpha) \quad \text{(by definition of } S_A\text{)}$$

and hence

$$g_\alpha(f_A(\alpha)) < \omega_\alpha \quad \text{(by definition of } g_\alpha\text{)}$$

which will be useful later on. We denote $g_\alpha(f_A(\alpha))$ by $h_A(\alpha)$. If $w(\alpha) = \omega_\alpha$, then we see that h_A is w-regressive.

We now exploit the fact that S_A is stationary. Hence we may apply corollary 4.32 to S_A and h_A and hence obtain a stationary set $T_A \subset S_A$ as well as an ordinal $\gamma_A < \omega_1$ such that if $\alpha \in T_A$ then

$$h(\alpha) = g_\alpha(f_A(\alpha)) < \omega_{\gamma_A}$$

is true. As S_A is ω_1-stationary, it is ω_1-unbounded, and using the fact that ω_1 is regular, we see that $|S_A| = \aleph_1$. Hence there are $2^{\aleph_1}$-many subsets $T \subset S$. Similarly, by our construction, we see that $\gamma_A < \omega_1$, and hence there are at most $|\omega_1| = \aleph_1$ many γ. Hence we may conclude that there are at most

$$\begin{aligned}2^{\aleph_1} \times \aleph_1 &= \aleph_2 \times \aleph_1 \quad \text{(by our initial assumptions)}\\ &= \aleph_2\end{aligned}$$

many pairs (T_A, γ_A) that satisfy the conclusion of corollary 4.32. Using the fact that $\aleph_{\omega_1+1}$ is regular, we see that there must exist a pair (T, γ) such that

$$|\{A : A \lessdot B \wedge T = T_A \wedge \gamma = \gamma_A\}| \geq \aleph_{\omega_1+1}.$$

This holds by the second claim and our definition of B (informally, if we could attain $\aleph_{\omega_1+1}$ by taking unions of sets strictly smaller than $\aleph_{\omega_1+1}$, then this would, by definition, contradict the fact that $\aleph_{\omega_1+1}$ is regular; by our assumptions, there are only at most $\aleph_2$ pairs (T, γ) and $\{A : A \lessdot B\}$ has cardinality at least $\aleph_{\omega_1+1}$. As $\{A : A \lessdot B \wedge T = T_A \wedge \gamma = \gamma_A\}$ is a subset thereof, the result follows).

In the final step of the proof, recall that $\gamma < \omega_1$ implies $\gamma + 1 < \omega_1$. Hence

$$|\omega_\gamma|^{|T|} = \aleph_\gamma^{\aleph_1} = \max(\aleph_1^{\aleph_1}, \aleph_\gamma^{\aleph_\gamma}) = \max(2^{\aleph_\gamma}, 2^{\aleph_1}) = \max(\aleph_{\gamma+1}, \aleph_2) < \aleph_{\omega_1}$$

gives the number of functions from T into ω_γ (note that if we restrict h_A to T_A, then this h_A is such a function from T_A to ω_{γ_A}). The equality above follows directly from our initial assumption that CH holds below $\aleph_{\omega_1}$ and a simple application of cardinal arithmetic. Finally, as $\{A : A \lessdot B\}$ has cardinality at least $\aleph_{\omega_1+1}$, there must exist distinct sets A_1 and A_2 such that

$$A_1 \lessdot B,\ A_2 \lessdot B,\ T_{A_1} = T_{A_2} = T \text{ and } \gamma_{A_1} = \gamma_{A_2} = \gamma.$$

Fix any $\alpha \in T$. By definition of h_A, we have

$$h_{A_1}(\alpha) = h_{A_2}(\alpha)$$

which we can rewrite in its original form as

$$g(f_{A_1}(\alpha)) = g(f_{A_2}(\alpha)).$$

But g was defined to be injective. Thus we obtain

$$f_{A_1}(\alpha) = f_{A_2}(\alpha)$$

for all $\alpha \in T$, as α was chosen arbitrarily. Unfolding the definition of f_A, we hence obtain that $A_1 \cap \omega_\alpha = A_2 \cap \omega_\alpha$ for each $\alpha \in T$. As T is a stationary subset of ω_1, we hence see that $A_1 \cap \omega_\alpha = A_2 \cap \omega_\alpha$ for all $\alpha < \omega_1$, which implies that $A_1 = A_2$, contradicting our assumption that A_1 and A_2 are distinct.

Thus we have obtained the required contradiction, and the proof is complete. □

It is worth noting that the proof of Silver's theorem we give above is of a purely combinatorial nature. The original proof by Silver (see [Sil75] for reference) uses advanced techniques such as forcing.

It should be noted that Silver's original proof does not only apply to $\aleph_{\omega_1}$ but to any singular cardinal of uncountable cofinality. However as we remarked at the end of section 3.4.1, the result does not hold for $\aleph_\omega$. In [BP77, p. 113], Baumgartner and Prikry suggest this is due to the fact that the notion of closed unbounded sets is not particularly meaningful for sets of countable cofinality, as we observed, too.

For completeness, we give a generalisation of Silver's theorem (the proof is omitted, it can be found in [HH99, p. 244]):

Theorem 4.35 (Galvin-Hajnal-Theorem). *Let α be an ordinal. If $\aleph_\alpha$ is a singular strong limit cardinal of uncountable cofinality, then the following holds:*

$$2^{\aleph_\alpha} < \aleph_{(|\alpha|^{\mathrm{cf}(\aleph_\alpha)})^+}$$

4.5 What next?

Using Easton's theorem, we have seen that the continuum function on *regular cardinals* can take any values subject to the following constraints:

- if $\alpha \leq \beta$ then $2^{\aleph_\alpha} \leq 2^{\aleph_\beta}$;
- $\mathrm{cf}(2^{\aleph_\alpha}) > \aleph_\alpha$.

For *singular cardinals*, the situation is trickier. In the following exposition, we present results from [Jec95, p. 410].

For a regular cardinal $\aleph_\alpha$ we see that

$$\aleph_\alpha^{\mathrm{cf}(\aleph_\alpha)} = \aleph_\alpha^{\aleph_\alpha} = 2^{\aleph_\alpha}$$

whereas for a singular cardinal we only have

$$\aleph_\alpha^{\mathrm{cf}(\aleph_\alpha)} \leq 2^{\aleph_\alpha}.$$

Hence we need to consider the function that maps α to $\aleph_\alpha^{\mathrm{cf}(\aleph_\alpha)}$ in order to understand cardinal exponentiation. As we have shown in section 3.4, König's theorem allows us to assign further constraints to this inequality. Finally, it can be shown that in fact the only function that is crucial in order to determine cardinal exponentiation is the term $\kappa^{\mathrm{cf}(\kappa)}$. As shown above, this is simple for regular cardinals as it reduces to the continuum function. If

$$\kappa \leq 2^{\mathrm{cf}(\kappa)} \tag{$*$}$$

then

$$\begin{aligned} 2^{\mathrm{cf}(\kappa)} &\leq \kappa^{\mathrm{cf}(\kappa)} && \text{(as } \kappa > 2 \text{ by assumption)} \\ &\leq 2^{\mathrm{cf}(\kappa) \times \mathrm{cf}(\kappa)} && \text{(by } (*)) \\ &= 2^{\mathrm{cf}(\kappa)} \end{aligned}$$

and hence

$$\kappa^{\mathrm{cf}(\kappa)} = 2^{\mathrm{cf}(\kappa)}.$$

Thus the only remaining case which needs to be considered is the case of $\kappa > 2^{\mathrm{cf}(\kappa)}$. Of course, we may now use the result of Silver's theorem above to investigate this case.

The so-called *singular cardinals problem* (also known as *singular cardinal hypothesis*, or SCH for short) addresses this problem. SCH postulates that if κ is a cardinal for which $2^{\mathrm{cf}(\kappa)} < \kappa$ then $\kappa^{\mathrm{cf}(\kappa)} = \kappa^+$.

The interested reader may want to consult [Jec95] in detail, which provides an introduction to the theory that arises from investigating this hypothesis, which is called *possible cofinality theory* and is due to Shelah.

5 Constructibility

This section introduces Gödel's constructible universe. In the course of this report, we have not considered the question in which universe we work. This stems partly from the fact that ZFC does not tell us how to define sets, but merely how to verify whether some object is a set. As Gödel proved in his seminal paper [Göd38], the constructible universe provides us with an interesting theory that can be employed to prove far-reaching results such as the relative consistency of AC and ZF.

We will not make a case for replacing our ground universe by the constructible universe (this is called the *Axiom of Constructibility*, which we shall investigate in detail). The intention of this section is to introduce the structure of the constructible hierarchy as well as to illustrate its versatility and remarkable properties relating to infinitary combinatorics.

We commence by briefly recapping the language of set theory, LST, and the structure of its formulas. We introduce the notions of absoluteness and of definability, which will be crucial to the definition of constructible sets. In order to define the constructible universe, we will require a formal analogue of LST expressed exclusively in sets. This is what section 5.5 will be devoted to.

We will be working in ZF. As outlined in Gödel's original paper [Göd38], in subsection section 5.6 we will define the term *constructible*. It will act as our building block of the eponymous universe. That particular subsection will also focus on showing various powerful consequences of the definition of the hierarchy defined by constructible sets, which we shall denote by L. Among others, we will show that if Φ is an axiom of T where T is a subtheory of ZF then Φ^L holds (we say that L is an inner model of T). The special case in which T is ZF will feature prominently.

The main result of this section, however, will be the conclusion presented in section 5.7. By defining a suitable well-ordering on the class L, we will show that the Axiom of Choice holds in L. Hence we will deduce the relative consistency of ZF and AC.

This section on constructibility follows closely the notes by Keith Devlin, cf. [Dev17] for reference. As in the previous sections, unless otherwise stated, proofs have been found by the author.

5.1 The Language of Set Theory

As mentioned earlier, it will be vital to have a language we can work in. This will serve as our logical foundation. As we know from axiomatic set theory, we only require the binary relation symbol $\in$ to be able to capture all notions within set theory. These will be introduced here.

We require:

- one logical connective $\wedge$ (conjunction);
- the existential quantifier $\exists$;
- the logical symbol $\neg$ (negation);
- one binary relation symbol $\in$ (set membership);
- and countably many variables $v_0, v_1, \ldots$

For ease of readability, we will also use brackets in a rather flexible way. Legibility takes priority.

The atomic formulas of LST are given by

$$v_m = v_n$$

and

$$v_m \in v_n$$

for variables v_m and v_n.

We assign the usual shorthands and treat the following symbols as fundamental elements of the language (clearly, these can be derived from the logical symbols above very easily):

- the disjunction symbol $\vee$;
- the implication symbol $\rightarrow$;
- the logical equivalence symbol $\leftrightarrow$;
- and the universal quantifier $\forall$.

Further, we occasionally use the unique existential quantifier $\exists!$.

The language we will be using is induced by the symbols above in the natural way. For example, if Φ and Ψ are formulas of *LST* then so are $\Phi \wedge \Psi$ and $\Phi \vee \Psi$ as well as $\neg\Phi$. We also use the natural shorthands for (proper) set inclusion and unique existence. Further description is omitted.

Finally, bounded quantifiers will be used throughout, and the shorthand

$$\exists v_m \in v_n(\Phi) \quad \text{for} \quad \exists v_m(v_m \in v_n \wedge \Phi)$$

will be used.

Remark. *Note that, following our definition above, LST does not have any constant symbols. Once we have translated LST formulas into sets, we define $\mathscr{L}$ to be the set-wise defined analogue of LST. We will then extend the language $\mathscr{L}$ so that the elements of any set u will be defined to be constant symbols in $\mathscr{L}_u$.*

If we then consider the universal set V and hence the language $\mathscr{L}_V$ we have exactly what we need: a formal analogue of LST expressible exclusively in terms of sets in which every set within V is a constant symbol of $\mathscr{L}_V$.

The first-order logic that is given by the symbols above is what we call the Language of Set Theory, or, for short, *LST*.

Remark. *One crucial notational remark must be made at this point: throughout the pevious sections, we denoted ordered n-tuples using angle brackets; we will abandon this convention now and write ordered n-tuples using parentheses (e.g. $x = (1, 2)$ is the ordered pair), whereas angle brackets are reserved for sequences (e.g. $x = \langle 1, 2\rangle$ is the sequence with first element 1 and second element 2).*

5.2 The Axioms of ZFC

This section on constructibility is very heavy on notation and several conventions will be used that need prior explanation. We stick closely to the chosen nomenclature and conventions used by Devlin in [Dev17].

- For a finite number of variables $v_1, v_2, \ldots, v_n$, we simply write $\vec{v}$ to denote the sequence $(v_1, \ldots, v_n)$. We call n the length of $\vec{v}$. In order to indicate the length of $\vec{v}$, we write $\vec{v}^n$.
- When considering a formula Φ, we say that free variables are among $\vec{v}^n = (v_1, \ldots, v_n)$ if whenever i is such that v_i is free in Φ then $i \leq n$.
- If $\vec{v}^n = (v_1, v_2, \ldots, v_n)$, then
$$\forall \vec{v}^n$$
is a shorthand for
$$\forall v_1 \, \forall v_2 \, \ldots \, \forall v_n.$$
Similarly, we write
$$\exists \vec{v}^n$$
in place of
$$\exists v_1 \, \exists v_2 \, \ldots \, \exists v_n.$$

As mentioned previously, we aim to translate *LST* into sets. One of the main tools to use (the "letters" of our language, in some sense) will be ordered sets. Hence we introduce the following notational conventions:

- For an ordered tuple $x = (x_0, x_1, \ldots, x_{n-1})$, we write $(x)_i$ to denote the element x_i; hence this coordinate function is defined for $i < n$ only.
- We call n the length of the ordered tuple. In order to indicate the length of x, we write $(x)^n$, analogously to the case of sequences of variables above. When both the length as well as the coordinate are of interest, we may also write $(x)^n_i$ to denote x_i. Again, this is defined for $i < n$ only.
- Following these conventions, a sequence x is of length n (where n is a natural number) if and only if $\mathrm{dom}(x) = n$.

Throughout the following section, we will be mainly working in ZF, and consider the Axiom of Choice later on when we prove that the constructible universe L is an inner model of ZFC. For completeness, we state the theorems of ZF below:

Axioms of ZF

Extensionality:	E	$\forall x\, \forall y\, (\forall z\, (z \in x \leftrightarrow z \in y) \rightarrow x = y)$
Foundation:	F	$\forall x\, (\exists y\, (y \in x) \rightarrow \exists y\, (y \in x \wedge \neg \exists z\, (z \in x \wedge z \in y)))$
Pairing:	PA	$\forall x\, \forall y\, \exists z\, (x \in z \wedge y \in z)$
Union:	U	$\forall \mathscr{F}\, \exists A\, \forall Y\, \forall x\, (x \in Y \wedge Y \in \mathscr{F} \rightarrow x \in A)$
Replacement:	R	$\forall A\, \forall \vec{w}^n\, (\forall x \in A\, \exists! y\, (\Phi(x, y, A, \vec{w}^n)) \rightarrow \exists Y\, \forall x \in A\, \exists y \in Y (\Phi(x, y, A, \vec{w}^n)))$ for any LST-formula Φ with free variables among $x, y, A, \vec{w}^n$.
Separation:	S	$\forall z\, \forall \vec{w}^n\, \exists y\, \forall x\, (x \in y \leftrightarrow (x \in z \wedge \Phi(x, z, \vec{w}^n)))$ for any LST-formula Φ with free variables among $x, z, \vec{w}^n$.
Powerset:	P	$\forall x\, \exists y\, \forall z\, (z \subset x \rightarrow z \in y)$
Infinity:	I	$\exists x\, (\exists y\, (y \in x) \wedge \forall y \in x\, \exists z \in x (y \in z))$

Note that the Axioms of Replacement and Separation are axiom schemas, with one axiom for each formula Φ in LST. We will discuss the Axiom of Choice later. Further, note that we have used a couple of shorthands above in order to maintain readability (in P we used the subset symbol, for example). Whenever we use the LST-formulas in a formal manner, we shall omit using shorthands such as those described above.

5.3 The Cumulative Hierarchy

The axioms of ZFC are powerful statements when we would like to test whether a given mathematical object is a set – the only existential axioms are the *axiom of infinity* and the *empty set axiom*, which states that there is a set which happens to be empty (although it is not always stated as an axiom, it can be derived easily from the axiom of extensionality and replacement). In order to construct more sets, it seems reasonable to start with the empty set and apply the powerset and union axiom, for example:

Definition 5.1. The cumulative hierarchy of sets is defined inductively as follows: set

$$V_0 = \emptyset,$$

and define

$$V_{\alpha+1} = \mathcal{P}(V_\alpha).$$

Finally, if α is a limit ordinal, define

$$V_\alpha = \bigcup_{\beta<\alpha} V_\beta.$$

Now define the cumulative hierarchy, denoted by V, to be

$$V = \bigcup_{\alpha \in \mathbf{ON}} V_\alpha.$$

Occasionally, we may refer to a specific set V_α as the α-level of the cumulative hierarchy. We may just say level if there is no danger of confusion.

As can be seen easily, this hierarchy provides us with a rich collection of sets. One might be tempted to say this universe is "too rich"; we quickly run into problems when we consider this hierarchy.

In our above definition of the cumulative hierarchy, we have relied on the so-called unrestricted power set operator. In each step of our inductive definition, we considered the collection of all possible subsets of the previously obtained level of V. This unrestricted operator gives rise to problems we call independent of ZFC.

Definition 5.2. Let T be a set of LST-sentences. Then ϕ is called independent of T if

$$T \nvdash \phi$$

and

$$T \nvdash \neg\phi.$$

Most famously, Paul Cohen proved in 1966 that CH is independent of ZFC.

Definition 5.3. Let $x \in V$. Then we say that x is of rank α if α is the least ordinal for which $x \in V_{\alpha+1}$. It will be denoted by $\alpha = \text{rank}_V(x)$

Remark. *As the nomenclature indicates, the rank can be defined on any hierarchy. In fact we will later use the rank* rank_L *of elements of L in order to define a well-ordering on L and hence show that the Axiom of Choice holds in L.*

Why do we not define the rank to be the least ordinal of its related cumulative hierarchy level? In particular, why do we not define

$$\text{rank}(x) = \alpha \Leftrightarrow x \in V_\alpha \wedge \forall\beta < \alpha(x \notin V_\beta)$$

as our definition of rank? As it turns out after sparing a few moments' thoughts, we would quickly run into problems whenever we considered the case for sets whose "rank" (using our new definition) is a limit ordinal. Note that, by definition of the cumulative hierarchy, there is no set x for which

$$x \in V_\gamma \wedge \gamma \text{ is a limit ordinal} \wedge \forall\beta < \gamma(x \notin V_\beta).$$

Hence, in such cases, our definition of "rank" would be nonsensical.

It is clear that the rank of any set x is unique and that if $\text{rank}_V(x) = \alpha$, then $x \in V_\beta$ for any $\beta > \alpha$. Naturally the rank of a set is defined to be a meaningful identifier for a set's position within the cumulative hierarchy.

5.4 *LST* in Detail

In this subsection, we will introduce two new concepts. We begin by defining the *relativisation of an LST-formula to a class M*. In short, we will be able to bound unbounded quantifiers of a given LST-formula to a specific class. Along the way, we will state and prove the *Reflection Principle of Transitive Hierarchies*, a salient result necessary for the development of the subject matter later on.

Secondly, we will define how to classify formulas of LST. The Lévy hierarchy will help us do so by categorising the order and complexity of *unbounded* existential and universal quantifiers within each formula of LST. More importantly, however, is the number of interchanging blocks of like quantifiers, which we shall investigate in detail. Further, we will introduce the concept of absoluteness which is neatly intertwined with the reflection principle mentioned above.

5.4.1 Relativisation and the Reflection Principle

Let M be any class. Intuitively, the reflection principle states that whenever we have an LST-formula Φ that holds in V then there exists some $\alpha \in \mathbf{ON}$ such that Φ holds in V_α. Hence the nomenclature: the formula reflects down onto some initial segment of V.

Thus, there exists some minimal α' with this property. This is the strength of the reflection principle: we need not consider the entire hierarchy in order to examine any formula, but only a specific (and, in particular, smaller) initial segment of said hierarchy.

Remark. *Due to the translation process and hence equivalence of formulas expressed in (1) LST and (2) $\mathscr{L}_u$ for some fixed set u, we will adhere to the following convention:*

- *formulas of LST will be denoted by capital Greek letters (Φ, Ψ, etc.);*
- *formulas in $\mathscr{L}_u$ (note, these will be sets!) will be denoted by lower case Greek letters (ϕ, ψ, etc.).*

However, the reader is reminded that we aim to translate formulas from LST into sets, and hence, naturally, our goal is to preserve "meaning" between the translations. That is, formulas Φ and ϕ represent metamathematically the same statement.

We now fix a class M.

By definition, there exists an LST-formula Ξ such that

$$M = \{x : \Xi(x)\}.$$

Our aim in this section is to define a natural way of considering an LST-formula Φ restricted to M. That is, any unbounded quantifier within Φ should range over M only, and hence be bounded.

Definition 5.4. Let Φ be an LST-formula. Then we define the relativisation of Φ to M by the following:

- if Φ is atomic, then define Φ^M to be Φ;
- if Φ is of the form $\Psi \wedge \Theta$, then define Φ^M to be $\Psi^M \wedge \Theta^M$;
- if Φ is of the form $\neg\Psi$, then define Φ^M to be $\neg(\Psi^M)$.

Further, and most importantly:

- if Φ is of the form $\exists v_n(\Psi)$, then define Φ^M to be $\exists v_n \in M(\Psi^M)$.

The first three definitions given above are trivial. The focus should be on the final point of the definition which is crucial as it is the only one natively involving M and, more importantly, its definition. It is here that the quantifier-bounding takes place.

Remark. *The reader should note that we have used multiple shorthands in the presentation above. Formally, the following treatment is necessary: if Φ is $\exists v_n(\Psi)$, then define Φ^M to be $\exists v_n(\Xi(v_n) \wedge \Psi^M)$. Here we have used the formal definition of M and used the formal expression for bounded quantifiers. Hence we have used the fact that $v_n \in M$ if and only if $\Xi(v_n)$ holds.*

Hence we can state the following.

Corollary 5.5. *Let Φ be an LST-formula. Then so is the formula Φ^M.*

Proof. This follows immediately from the definition as well as the remark. □

It is noteworthy that we did not assume M to be a transitive class. The generalised reflection principle we will state shortly only applies to transitive classes, however.

The Generalised Reflection Theorem given with respect to ZF is due to Lévy and Montague.

Theorem 5.6 (The Generalised Reflection Principle)**.** *Consider a hierarchy* $(W_\alpha)_{\alpha\in\mathbf{ON}}$ *of transitive sets where* $W_\alpha = \{x : \Psi(x,\alpha)\}$ *for some LST-formula* Ψ*. Further, assume* $W := \bigcup_{\alpha\in\mathbf{ON}} W_\alpha$*. Suppose the hierarchy satisfies the following two conditions:*

- *if* $\alpha < \beta$ *then* $W_\alpha \subset W_\beta$*;*
- *if* γ *is a limit ordinal then* $W_\gamma = \bigcup_{\alpha<\delta} W_\alpha$*.*

Then the following holds: if $\Phi(\vec{v})$ *is an LST-formula with free variables among* $\vec{v}$ *then*

$$\forall\alpha\ \exists\beta < \alpha\ (\beta \textit{ is a limit ordinal} \wedge \forall\vec{v}\in W_\beta\ (\Phi^{W}(\vec{v}) \leftrightarrow \Phi^{W_\beta}(\vec{v}))) \tag{†}$$

is a theorem of ZF*.*

This result will be crucial for building the constructible universe. Intuitively, it states that any sentence that is true in W is also true on some initial segment of W, which, due to the two constraints given above, we can determine by W_β for some $\beta \in \mathbf{ON}$. Hence the nomenclature: the result "reflects down".

As mentioned above, it is necessary for the sets in the hierarchy to be transitive. The proof below will visualise this requirement. It is based on Devlin's presentation in [Dev17, pp. 25-26].

Proof. The proof comes in two stages: firstly, we decompose the formula Φ based on its complexity into formulas $(\Phi_i)_{i<n}$ and find a limit ordinal β as required. Using induction on n, we then verify that β satisfies the requirements set out in (†).

Step 1: For a given LST-formula $\Phi(\vec{v})$ with free variables among $\vec{v}$, consider a sequence of LST-formulas $(\Phi_i(\vec{x}_i))_{i<n}$ such that

- Φ_i is an atomic LST-formula; or
- Φ_{i+1} can be constructed from some Φ_j for $j \leq i$ using negation, conjunction, or by applying the existential quantifier (as per the definition of LST); and
- $\Phi_n = \Phi$.

It is clear that the sequence $\vec{x}_i$ crucially depends on the complexity of Φ_i and does therefore potentially change at each step of our sequence of LST-formulas.

We will obtain our value of β by defining functions f_i for each $i \leq n$ from the parameters $\vec{x}_i^n$ to $\mathbf{ON}$ that return 0 if the formula Φ_i belonging to the vector $\vec{x}_i^n$ is primitive, a negation, or a conjunction of previous formulas of the sequence. The only cases that are of importance are those for which Φ_i includes quantifiers. Hence we define

$$f_i(\vec{x}_i) = \begin{cases} 0 & \text{if } \Phi_i \text{ is primitive, a negation,} \\ & \text{or a conjunction of some } \Phi_j, \Phi_k \text{ for } j,k < i \\ \beta & \text{if } \Phi \text{ is of the form } \exists y\ \Phi_j(y,\vec{x}_i) \text{ for some } j < i \text{ and } \beta \text{ is the least ordinal} \\ & \text{for which } \exists y \in W\ \Phi_j^W(y,\vec{x}_i) \to \exists y \in W_\beta\ \Phi_j^W(y,\vec{x}_i) \text{ holds.} \\ & \text{We denote this formulas by } (**). \end{cases}$$

for all $i \leq n$. By our definition os LST, this definition is clearly exhaustive and hence well-defined.

Fix any α. We now consider a limit ordinal β such that $\forall\vec{x}_i \in W_\beta\ (f_i(\vec{x}_i)) < \beta)$ for all $i \leq n$. By the axiom of replacement with a suitable LST-formula, we can easily deduce the existence of such an ordinal β (one should note, however, that the function we need to define in order to appeal to the axiom of replacement is cumbersome to write out formally).

Step 2: We now proceed by induction on i and iterate over the complexity of Φ (i.e. we consider Φ_i for each $i \leq n$) and prove that

$$\Phi_i^W(\vec{x}_i) \leftrightarrow \Phi_i^{W_\beta}(\vec{x}_i)$$

for any $\vec{x}_i \in W_\beta$. This clearly proves the result.

Hence assume $\vec{x}_i \in W_\beta$.

- The base case $i = 0$ is trivial as Φ_0 must be primitive, and hence the formulas coincide (the relativisation does not affect the formulas).
- For the inductive step, firstly consider formulas Φ_i that are primitive or of the form $\neg\Phi_j$ or $\Phi_j \wedge \Phi_k$ for some $j, k < i$. By the same reasoning as for the base case, it is clear that in all such cases the inductive step is trivial again as conjunction and negation do not introduce any quantifiers that need to be taken care of.

  The only non-trivial case arises when Φ_i is of the form $\exists y \; \Phi_j(y, \vec{x}_i)$ for some $j < i$, which we shall consider now. We prove both implications independently:

  - For the left-to-right direction, suppose that $\exists y \in W \; \Phi_j^W(y, \vec{x}_i)$ holds. We need to show that then $\exists y \in W_\beta \; \Phi_j^{W_\beta}(y, \vec{x}_i)$ is also true. We can now make use of our function f_i: by definition, we see that $f_i(\vec{x}_i) < \beta$ and, further, by $(**)$, we have
    $$\exists y \in W_\beta \; (\Phi_j^W(y, \vec{x}_i)).$$
    Note that both y and $\vec{x}_i$ are now elements of W_β. Further, as $j < i$, we may now apply the inductive hypothesis to the formula Φ_j^W and hence obtain
    $$\exists y \in W_\beta \; (\Phi_j^{W_\beta}(y, \vec{x}_i))$$
    which is, by definition, $\Phi_i^{W_\beta}(\vec{x}_i)$, as required.
  - For the other direction, suppose that $\Phi_i^{W_\beta}(\vec{x}_i)$ holds, which we may rewrite as
    $$\exists y \in W_\beta \; (\Phi_j^{W_\beta}(y, \vec{x}_i)).$$
    We may apply the induction hypothesis immediately to $\Phi_j^{W_\beta}(y, \vec{x}_i)$ and hence obtain
    $$\exists y \in W_\beta \; (\Phi_j^W(y, \vec{x}_i)).$$
    So we only remain to show that a suitable y exists in W. But this follows trivially as $W_\beta \subset W$ by the initial assumptions. Hence the inductive step is proven.

Hence the proof is complete. □

The following lemma will allow us to deduce that we may apply the Generalised Reflection Principle to V:

Lemma 5.7. *The cumulative hierarchy V satisfies the condition in the hypothesis of the Generalised Reflection Principle.*

Proof. We prove the result in reverse order: for transitivity, we proceed by induction:

- The result is trivial for $\alpha = 0$ as $V_0 = \emptyset$.
- Assume V_α is transitive and suppose $x \in V_{\alpha+1}$. By definition, $V_{\alpha+1} = \mathcal{P}(V_\alpha)$, and hence x is a subset of V_α. As such, each element $y \in x$ is in fact an element of V_α. But by our assumption, V_α is transitive, and hence $y \subset V_\alpha$. Thus, x is in fact a set of subsets of V_α, and as $V_{\alpha+1}$ is the set of all subsets of V_α, we have $x \subset V_\alpha$.
- Now assume V_α is transitive for all $\alpha < \beta$, and suppose that β is a limit ordinal. Let $x \in V_\beta = \bigcup_{\gamma<\beta} V_\gamma$. Hence $x \in V_{\gamma'}$ for some $\gamma' < \beta$. By our assumption, $V_{\gamma'}$ is transitive, and thus $x \subset V_\gamma$, whence $x \subset V_\beta$, as required.

Now, in order to prove that for all $\alpha \leq \beta$ we have $V_\alpha \subset V_\beta$, we again use induction and the just proven transitivity:

- Again, the result is trivial for $\alpha = 0$.
- Assume that $x \in V_\alpha$. We need to show that $x \in V_{\alpha+1}$. By transitivity, $x \subset V_\alpha$, and hence $x \in \mathcal{P}(V_\alpha)$. But $\mathcal{P}(V_\alpha) = V_{\alpha+1}$, as required.
- The limit case is trivial as it follows directly from the definition of V.

Hence the proof is complete. □

The following result is given without proof in [Dev17, p. 26].

Theorem 5.8 (The Reflection Principle)**.** *Let $\Phi(\vec{v})$ be an LST-formula with free variables among $\vec{v}$. Then the sentence*

$$\forall\alpha\, \exists\beta < \alpha\ (\beta \textit{ is a limit ordinal} \wedge \forall\vec{v} \in V_\beta\ (\Phi(\vec{v}) \leftrightarrow \Phi^{V_\beta}(\vec{v}))) \qquad (\dagger')$$

is provable in ZF*.*

Proof. By the previous lemma we know that V is transitive, as required. Hence the Generalised Reflection Principle is applicable. □

We will consider applications of the Reflection Principle shortly. However, we will introduce the crucial notion of *absoluteness* first.

5.4.2 Absoluteness and the Lévy Hierarchy

A notion similar to the Reflection Principle is the following: we fix a transitive class M and consider an LST-formula $\Phi(\vec{x})$. If we consider the relativisation of Φ to M, we may wonder under what conditions Φ^M holds.

The following two definitions formalise this idea.

Fix a transitive class M and an LST-formula $\Phi(\vec{x})$.

Definition 5.9. We say that $\Phi(\vec{x})$ is downward absolute for M (or D-absolute) if

$$\forall\vec{x} \in M\ (\Phi(\vec{x}) \rightarrow \Phi^M(\vec{x})).$$

Hence, $\Phi(\vec{x})$ is D-absolute if whenever $\Phi(\vec{x})$ holds for some $\vec{x} \in M$, then the relativisation $\Phi^M(\vec{x})$ also holds. In other words, the formula $\Phi(\vec{x})$ is D-absolute if and only if we can somehow restrict $\Phi(\vec{x})$ to M and, at the same time, retain its truth value. This approach is analogous to the Reflection Principle: the formula reflects down onto M.

As the nomenclature suggests, there exists an upward-analogue:

Definition 5.10. We say that $\Phi(\vec{x})$ is upward absolute for M (or U-absolute) if

$$\forall \vec{x} \in M\ (\Phi^M(\vec{x}) \to \Phi(\vec{x})).$$

In this case, the reflection is upwards: an U-absolute formula $\Phi(\vec{x})$ satisfies that if $\Phi^M(\vec{x})$ holds for all $\vec{x} \in M$ then in fact the non-relativised formula $\Phi(\vec{x})$ also holds for each $\vec{x} \in M$.

Finally, we define absoluteness:

Definition 5.11. We call $\Phi(\vec{x})$ absolute for M if $\Phi(\vec{x})$ is both upward and downward absolute.

Kunen defines absoluteness on all classes and not only on transitive classes (see [Kun80, p. 117]). As we will be exclusively working with transitive classes, we adhere to the less general definition provided by Devlin. For more details on the importance of transitivity, see below. Similarly, in [Put63], Putnam calls such formulas invariant.

We can now briefly return to the Reflection Principle and a simple application thereof which yields a nice result:

Corollary 5.12. ZF *is not finitely axiomatisable.*

This result and some explanations on the validity as well as the proof itself can be found in [Kun80, p. 138]. Before we give the proof, however, we require the following lemmata:

Lemma 5.13. *The LST-formula "$x = \mathrm{rank}_V(y)$" is absolute for transitive models.*

The proof is omitted; the proof (as presented in [Kun80, p. 129]) uses the fact that the function rank_V is defined recursively, and that absolute notions are closed under composition (cf. [Kun80, p. 121] for further details).

Lemma 5.14. *Let M be a transitive model of* ZF*. Then*

$$V_\alpha^M = V_\alpha \cap M$$

for all $\alpha \in M$.

The result as well as the proof can be found in [Kun80, p. 130]; we extend the proof by additional explanation below.

Proof. Observe that each level of the cumulative hierarchy is determined by the function rank_V. Indeed,

$$V_\alpha = \{x : \mathrm{rank}_V(x) < \alpha\}.$$

Hence $x \in M$ is an element of V_α if and only if $x \in V_\alpha^M$, which is the same as saying

$$V_\alpha^M = V_\alpha \cap M$$

as required. □

We are now ready to present the proof:

Proof of corollary 5.12: In view of a contradiction, assume ZF can be axiomatised by the LST-formulas $\Phi_1, \ldots, \Phi_n$. That is, each theorem of ZF can be obtained by constructing sentences from the LST-formulas above. Now consider the sentence $\Phi_1 \wedge \ldots \wedge \Phi_n$ and denote it by Φ. It now follows trivially that ZF is axiomatised by Φ. Hence we can apply the Reflection principle and find the least ordinal α such that Φ^{V_α} holds. Using the claim we see that V_α is a model of ZF. Hence, in particular, the absoluteness of rank, for example, holds within V_α. Take any $\beta \in V_\alpha$. By the previous lemma, we now have

$$V_\beta^{V_\alpha} = V_\beta \cap V_\alpha = V_\beta.$$

Hence we see that V_β is absolute for V_α. As ZF proves the Generalised Reflection principle, we see that

$$\exists \beta \ \Phi^{V_\beta}.$$

By the absoluteness proven above, we hence must have

$$\exists \beta < \alpha \ \Phi^{V_\beta}$$

which contradicts the minimality of α. □

Remark. *The reader might wonder about the consequences of the lack of absoluteness of certain formulas; Skolem's paradox sheds some light on this. Roughly, it postulates that a countable model of* ZF *can prove the existence of uncountable cardinals (assuming that* ZF *is consistent). This is possible due to the fact that the LST-formula "x is uncountable" is not absolute. The interested reader might want to consult [Res66], for example, for further details.*

Determining whether a given formula Φ is absolute in a given transitive class M is non-trivial. In the simple case of atomic formulas the result follows immediately.

Proposition 5.15. *Atomic formulas are absolute for all transitive classes M.*

Proof. The proof follows immediately from the definition of relativisation. Indeed, any atomic LST-formula Φ is, by definition, identical to its relativisation Φ^M. □

In general, however, we will have to examine the logical complexity of Φ. In particular, we are interested in the number and order of unbounded quantifiers occurring in Φ.

One method of constructing the *Lévy hierarchy* is by rewriting Φ so that it is of a form with all its unbounded quantifiers at the front. The following theorem is required:

Theorem 5.16 (Prenex Normal Form Theorem)**.** *Let Φ be a formula in LST. Then Φ is provably equivalent to some formula of the form*

$$Q_1 \vec{x}_1 \ldots Q_n \vec{x}_n \Psi$$

where Ψ is a quantifier-free formula, $\vec{x}_i$ are variables, and each Q_i is either an existential or a universal quantifier.

Remark. *By our notation, each sequence of variables $\vec{x}_i^{n_i}$ can be written as a sequence $\langle x_{i_1}, x_{i_2}, \ldots, x_{i_n} \rangle$. In particular, the length of $\vec{x}_i$ depends on i.*

Proof. The proof is omitted; details on the theorem as well as some additional results concerning the prenex normal form can be found in [CK62] for interest. □

We now introduce the definitions that are paramount to the following sections:

Definition 5.17. Let Φ be an *LST*-formula. We define the quantifier complexity of Φ recursively: if Φ has no unbounded quantifiers, then we say that

$$\Phi \text{ is } \Sigma_0 \text{ (and } \Pi_0). \tag{$*$}$$

For any $n > 0$, we define the complexity as follows:

- if Φ is of the form $\exists \vec{x}\ \Phi'(\vec{x})$ where $\Phi'(\vec{x})$ is Π_{n-1}, then we say

$$\Phi \text{ is } \Sigma_n;$$

- if Φ is of the form $\forall \vec{x} \Phi'(\vec{x})$ where $\Phi'(\vec{x})$ is Σ_{n-1}, then we say

$$\Phi \text{ is } \Pi_n.$$

Remark. *As indicated above, we are only interested in unbounded quantifiers. Hence, if Φ is Σ_0 and not atomic, it contains bounded quantifiers (existential or universal or both).*

Further, note that the blocks of existential and universal quantifiers alternate. Recall the notational convention of writing $\exists \vec{x}^n$ in place of $\exists x_1, \ldots, \exists_n$, which finds repeated application throughout the following sections.

Example 26. *We give a few examples of LST-formulas of different complexities:*

- *The formula*

$$\forall x\ \exists y \in x\ (y \subset x)$$

  is Π_1: the formula $\exists y \in x\ (y \subset x)$ is Σ_0 (there are no unbounded quantifiers involved) and the only preceding unbounded quantifier is universal.

- *The formula*

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

  is Σ_2: $y \notin x$ is Σ_0, the recursive definition tells us that $\forall y\ (y \notin x)$ is hence Π_1, and thus $\exists x\ \forall y\ (y \notin x)$ is Σ_2. (This is the Empty Set Axiom.)

- *The formula*

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

  is Π_2: the formula $\exists z\ (x \in z \wedge y \in x)$ is Σ_1, and the block of repeated universal quantifiers counts only once as per our definition. (This is the Pairing Axiom.)

- *The formula*

$$\forall x\ (\exists y\ (\Psi(x, y)) \to \Phi(x))$$

  does not permit its complexity to be read off as easily. However, note that we can write an equivalent formula

$$\forall x\ \forall y\ (\neg\Psi(x, y) \vee \Phi(x))$$

  which is in Prenex Normal Form. It is now clear that the formula is Π_1.

These examples visualise the structure of our definition of quantifier complexity: let Φ be an *LST*-formula. The frontmost unbounded quantifier of Φ is an existential quantifier if and only if Φ is Σ_n for some positive integer n. The analogous equivalence holds for universal quantifiers and the Π-notation.

Remark. *Consider the formula* Φ *defined by* $\forall x\,(x \in y \rightarrow x \in z)$*. Now,* Φ *is clearly* Π_1*. However, equivalently, we can easily write* $\forall x \in y\,(x \in z)$*, which is* Σ_0*. In order to resolve this ambiguity, we define the quantifier complexity to be the least* n *such that* Φ *is* Σ_n *(or* Π_n*, respectively).*

In the course of this section, we will build up a collection of semantic statements in *LST* and their respective formal analogues. Indeed, if we consider a formula in *LST* such as "x is an ordinal", then we have an understanding of its meaning without explicitly consulting the rendering in *LST*. However, it is undoubtedly necessary to examine the formula expressed in *LST* in order to determine its level in the Lévy hierarchy.

Below we consider simple formulas and their natural rendering in *LST* (these examples below follow Devlin's listing in [Dev17, p. 28]). Further down the list, we use shortcuts for *LST*-formulas we have above shown to be Σ_0 so as to improve readability.

Of course, this list is no way exhaustive.

The following *LST*-formulas are Σ_0:

Formula	**Rendering in *LST***
x is a subset of y	$\forall u \in x\,(u \in y)$
x is the set $\{y\}$	$\forall u \in x\,(u = y)$
z is the union of x and y	$\forall u \in x\,\forall v \in y\,(u \in z \wedge v \in z)$ $\wedge\ \forall w \in z\,(w \in x \vee w \in y)$
z is the intersection of x and y	$\forall u \in x\,(u \in z \leftrightarrow u \in y)$
x is the set $\bigcup y$	$\forall u \in x\,\exists v \in y\,(u \in v)$ $\wedge\ \forall u \in y\,\forall v \in z\,(v \in x)$
z is the set difference of x and y	$\forall u \in z\,(x \in z \wedge \neg(y \in z))$
x is a transitive set	$\forall u \in x\,\forall v \in u\,(v \in x)$
x is the ordered pair (y_1, y_2)	x is the union of $\{\{y_1\}\}$ and $\{\{y_1, y_2\}\}$
x is a relation on y	$\forall u \in x\,\exists a \in y\,\exists b \in y\,(u = (a, b))$
x is an ordinal	x is a transitive set and well-ordered with respect to $\in$
x is a successor ordinal	x is an ordinal and there exists y such that y is an ordinal and $x = y \cup \{y\}$
x is a limit ordinal	x is an ordinal and $\neg$(x is a successor ordinal)
x is a function on y	x is a relation and if $(a, b) \in x$ then whenever $(a, c) \in x$ we have $b = c$
x is the domain of a function y	the set y is a function and for every $(a, b) \in y$ we have $a \in x$
x is the range of a function y	the set y is a function and for every $(a, b) \in y$ we have $b \in x$
x is a sequence of length n	the set x is a function and $\lvert\operatorname{dom}(x)\rvert = n$

From these, it is clear that we can also write the formulas "x is the n-tuple $(y_1, \ldots, y_n)$" and "x is the sequence $\langle y_1, \ldots, y_n \rangle$" in terms of Σ_0 *LST*-formulas, for example.

We will be using many of these examples later on when the actual coding takes place. However, as indicated above, having written down the natural *LST*-rendering once allows us to be more flexible in the way we deal with them (for instance, there is no need to write out the *LST*-details every single time).

The following notion will be very useful:

Definition 5.18. Let Φ be an *LST*-formula. Then we say that $\underline{\Phi \text{ is } \Sigma_n^{\mathrm{ZF}}}$ if

$$\mathrm{ZF} \vdash \Phi \leftrightarrow \Psi$$

for some Σ_n formula Ψ. Similarly, we say that $\underline{\Phi \text{ is } \Pi_n^{\mathrm{ZF}}}$ if

$$\mathrm{ZF} \vdash \Phi \leftrightarrow \Psi$$

for some Π_n formula Ψ.

We naturally extend these definitions to subtheories T of ZF, and hence use the notation Σ_n^T and Π_n^T in such cases. Note that ZF is a subtheory of itself.

The structure of our Σ and Π-notation begs the question whether there are formulas that are both Σ_n^{ZF} and Π_n^{ZF} for some $n > 0$ (by our discussion earlier, the result is trivial for $n = 0$). As we will show exemplarily with the following result, such formulas indeed exist:

Proposition 5.19. *Let Φ denote the LST-formula "x is a well-founded relation on y". Then Φ is both Σ_1^{ZF} and Π_1^{ZF}.*

Before we give the proof, we need to derive a few closure properties of the Σ^T and Π^T-notation introduced above.

Of course, we want as many formulas as possible to be absolute with respect to a specific axiom system. In particular, constructing new absolute formulas from old ones is crucial; we will then be able to consider a respectable class of formulas and need not worry about transitivity.

Luckily, as the following results shows, absolute formulas behave well under many of the logical operations we can apply. The result is taken from [Dev17, p. 29].

Lemma 5.20. *For any subtheory T of* ZF*, the following hold:*

(i) if Φ and Ψ are Σ_0^T, then so are $\Phi \wedge \Psi$, $\Phi \vee \Psi$, and $\neg\Phi$;

(ii) if Φ is Σ_n^T, then $\neg\Phi$ is Π_n^T. Similarly, if Φ is Π_n^T, then $\neg\Phi$ is Σ_n^T;

(iii) Φ is Δ_n^T if and only if Φ is Σ_n^T and $\neg\Phi$ is Σ_n^T;

(iv) if Φ and Ψ are Σ_n^T, then so are $\Phi \wedge \Psi$, $\Phi \vee \Psi$, $\exists x\ \Phi$ and $\exists x \in z\ \Phi$;

(v) if Φ and Ψ are Π_n^T, then so are $\Phi \wedge \Psi$, $\Phi \vee \Psi$, $\forall x\ \Phi$ and $\forall x \in z\ \Phi$;

(vi) if Φ and Ψ are Δ_n^T, then so are $\Phi \wedge \Psi$, $\Phi \vee \Psi$, and $\neg\Phi$;

(vii) if $m < n$ then any formula Φ that is Σ_m^T or Π_m^T is also Δ_n^T.

The results given in the lemma seem trivial, however, they will be immensely useful. Closure under taking conjunctions, disjunctions, negations, and introducing unbounded quantifiers (existential ones for Σ and universal ones for Π) will allow us to build more complex formulas from simple ones while preserving the quantifier complexity. This will, in turn, preserve absoluteness, as we will prove in theorem 5.22.

Remark. *For the sake of completeness, it should be mentioned that the results above in fact hold for all LST-theories that include the axioms of predicate logic for LST.*

In its original presentation, the proofs are omitted; we present a few below in order to illustrate the simple reasoning.

Proof. All of the proofs above follow directly from the definition of Σ_n^T and Π_n^T notation introduced above.

For part (iii) for example, note that if $\neg\Phi$ is Σ_n^T, then we can apply de Morgan's laws to the leading block of existential quantifiers and hence express Φ equivalent to a Π_n^T formula, as required.

Part (vii) is only a quirk of the definition. However, as we shall see later, it will be of immense importance when we consider definability. □

This proof of proposition 5.19 follows closely Devlin's approach outlined in [Dev17, p. 29]; further comments have been added:

Proof of proposition 5.19. Note that the formula $\Phi(x, y)$ saying "x is a well-founded relation on y" can be decomposed into formulas Φ_1 and Φ_2 where Φ_1 states "x is a relation on y" and Φ_2 states "x is well-founded". Clearly, we then have that Φ is equivalent to $\Phi_1 \wedge \Phi_2$.

By the table above, we see that the formula "x is a binary relation on y" is Σ_0. Hence, it suffices to consider the *LST*-formula "x is well-founded", which we shall denote by Ψ. Let E be any binary relation. By definition of well-foundedness, we have that $\Psi(E, X)$ can be written as

$$\forall A\ (A \subset X \wedge A \neq \emptyset \rightarrow \exists a \in A\ \forall x \in A\ \neg(xEa)).$$

Another way of describing well-foundedness is the following: consider a function from X to the class of ordinals. If f preserves the binary relation E, then we may use the fact that the range of f, the class of ordinals, is well-ordered. Hence if such a function exists, then E is well-founded.

One can indeed formally prove that

$$\mathrm{ZF} \vdash \Phi(E, X) \leftrightarrow \exists f\ (f \colon X \rightarrow \mathbf{ON} \wedge \forall x, y \in X\ (xEy \rightarrow f(x) < f(y))).$$

Now we have two renderings of the formula Ψ, the first one is clearly Π_1, whereas we have proven the second one above to be provably equivalent to a Σ_1 formula with respect to ZF. Hence the result follows. □

Now, the following definition will be of immense use:

Definition 5.21. Let Φ be an *LST*-formula and let T be a subtheory of ZF. If Φ is both Σ_n^{ZF} and Π_n^{ZF}, then we say that

$$\Phi \text{ is } \Delta_n^{\mathrm{ZF}}.$$

Remark. *The notion of Δ_n formulas does not exist for $n \geq 1$. We need an underlying theory in order to verify whether the associated Σ_n and Π_n formulas are provably equivalent; this is impossible without such a theory.*

Our primary goal when building the Lévy hierarchy was to relate absoluteness to our construction based on quantifier complexity. The following statements will produce this connection.

Theorem 5.22. *Let T be some subtheory of* ZF *and let M be a transitive class. If the relativisation Ψ^M holds for all $\Psi \in T$, and if Φ is some LST-formula, then the following also hold:*

- *if Φ is Σ_0^T, then Φ is absolute;*

- *if* Φ *is* Σ_1^T*, then* Φ *is U-absolute;*
- *if* Φ *is* Π_1^T*, then* Φ *is D-absolute;*
- *if* Φ *is* Δ_1^T*, then* Φ *is absolute,*

where absoluteness is determined for M*.*

This theorem produces the link between absoluteness and the Lévy hierarchy that we were looking for in the first place. Indeed, its statement is surprising as it connects the syntax of formulas to their absoluteness (which is a statement about "truth").

The result is given in [Dev17, p. 27].

Proof. We prove (i) exemplarily; the latter cases are similar. Details for all of the proofs (the one we present below is taken from the same source) can be found in [Dev17, pp. 27-8].

Let T and $\Phi(\vec{v})$ with its free variables being among $\vec{v}$ be given. Assume $\Psi(\vec{v})$ is a Σ_0 formula such that

$$T \vdash \Phi \leftrightarrow \Psi.$$

Now, we do not know anything about the quantifier complexity of Φ. However, the following trick will help us simplify the proof: if T proves $\Phi \leftrightarrow \Psi$, then it does also prove $\forall \vec{v}\ (\Phi \leftrightarrow \Psi)$, and, as Φ^M holds for all $\Phi \in T$ by assumption, we also see that $(\forall \vec{v}\ (\Phi \leftrightarrow \Psi))^M$ holds. By definition of relativisation, we hence have

$$\forall \vec{v} \in M\ (\Phi^M \leftrightarrow \Psi^M).$$

Thus, in order to prove the result, we may assume that Φ is a Σ_0 formula itself.

We now prove the result by induction on the complexity of Φ.

- If Φ is primitive, then the result is immediate by definition.
- Assume Φ is of the form $\Phi_1 \wedge \Phi_2$ or of the form $\neg\Phi_1$. Using the definition again, absoluteness is guaranteed immediately for Φ following the definition yet again.

  The only non-trivial case is the following: assume $\Phi(y, \vec{v})$ is of the form $\exists x \in y\ (\Psi(x, y, \vec{v}))$ where Ψ is absolute. Let $y, \vec{v} \in M$. There are two directions to consider:

  - if $\Phi(y, \vec{v})^M$ holds, then, by assumption, we have $(\exists x \in y\ (\Psi(x, y, \vec{v})))^M$. By unwrapping the relativisation, we obtain

    $$\exists x \in M\ (x \in y \wedge \Psi^M(x, y, \vec{v}))$$

    But by the inductive hypothesis, we see that $\Psi(x, y, \vec{v})$ holds as it is absolute. Hence we may drop the relativisation and note that

    $$\exists x \in y\ (\Psi(x, y, \vec{v}))$$

    holds, which is precisely $\Phi(y, \vec{v})$, as required to show U-absoluteness.
  - For the other direction, we assume that $\Phi(y, \vec{v})$ holds, which, by definition, we may rewrite as

    $$\exists x \in y\ (\Psi(x, y, \vec{v})).$$

We apply the same trick as before and exploit the transitivity of M: as $y \in M$ and $x \in y$, we see immediately that $x \in M$. Thus we can conclude, as $\Psi(x, y, \vec{v})^M$ holds by the inductive hypothesis, that in fact

$$\exists x \in M\ (\exists x \in y \wedge (\Psi^M(x, y, \vec{v})))$$

is true as well. Therefore, we have $\Phi(y, \vec{v})^M$, as required.

The case in which Φ is of the form $\forall x \in y\ (\Psi(x, y, \vec{v}))$ is very similar (we must write, for example, $\forall x \in M\ (x \in y \to \Psi^M(x, y, \vec{v}))$).

Hence the result is proven by induction. □

Let T be ZF. If M is a transitive proper class and satisfies the hypotheses in theorem 5.22, i.e. Φ^M holds for each ZF-axiom Φ, then we call M an inner model of ZF. As we shall see, the constructible universe L is an inner model of ZF; we will investigate this case in section 5.6 (where we will also formally state the definition of inner models and its nomenclature again).

Remark. *Assume M is a transitive class, moreover, an inner model of* ZF*. If Φ is Δ_1^{ZF}, then it holds in M if and only if it holds in "the real world". This means, Φ yields the same interpretation in every single inner model of* ZF*.*

Lemma 5.23. *Assume the axioms of* ZF *and let M be a transitive class. Then the Axiom of Extensionality,* E, is absolute for M.

Proof. Let $\Phi(x, y)$ be the *LST*-formula $\forall z\ ((z \in x \leftrightarrow z \in y) \to x = y)$. We verify the definition, hence we need to show that $\Phi(x, y)$ is both U-absolute and D-absolute.

For U-absoluteness, consider Φ^M and note that

$$\forall x \in M\ \forall y \in M\ (\forall z\ (z \in x \leftrightarrow z \in y) \to x = y)^M$$

is the same as

$$\forall x \in M\ \forall y \in M\ (\forall z \in M\ (z \in x \leftrightarrow z \in y) \to x = y).$$

Now assume Φ^M holds. Note that, as M is transitive, we have $x \subset M$ and $y \subset M$. Suppose we are given elements $x, y \in M$, and hence we remain to show that

$$\forall z \in M\ ((z \in x \leftrightarrow z \in y) \to x = y).$$

By transitivity of M, the formula reduces to

$$\forall z\ ((z \in x \leftrightarrow z \in y) \to x = y). \tag{$*$}$$

This holds since transitivity implies that we do not need to range our quantifier acting on z over the entire class M but only over the subsets x and y in order to determine membership. Clearly, if $z \in x$ then z must be an element of M, and similarly for y. Hence we may drop the bound on the quantifier entirely.

Note that the resulting formula $(*)$ is in fact $\Phi(x, y)$. Hence we have shown U-absoluteness, as required.

For the other direction we need to show that if for all $x, y \in M$ the axiom of extensionality holds, then it also holds relativised to M. But this is clearly true as the non-relativised formula $\Phi(x, y)$ holds by the axiom of extensionality itself (it holds, in some sense, in "the real world"). Hence the proof is complete. □

Absoluteness is a crucial notion. It gives insight into the validity of a given formula within the universe as opposed to within the subtheory only. It is clear that any formula that does not contain any quantifiers is Σ_0 (this follows immediately from the definition). However, as Kunen points out in [Kun80, pp. 117-8], even simple Σ_0 formulas including bounded quantifiers can fail to be absolute:

Example 27. *Let $M = \{1, \{\{1\}\}\}$, and let $\Phi(x, y)$ be the LST-formula for $x \subset y$. Note that Φ is a Σ_0-formula (we have shown this in the table above). Further, $\Phi(\{\{1\}\}, 1)^M$ clearly holds: by definition, we need to check that*

$$(\forall z\ (z \in \{\{1\}\} \to z \in 1))^M$$

which is the same as

$$\forall z \in M\ (z \in \{\{1\}\} \to z \in 1)$$

holds. But as the M only contain two elements, $\{\{1\}\}$ and 1, and since neither is an element of $\{\{1\}\}$, the left hand side of the implication is always false and hence the implication itself is true. On the other hand, clearly $\Phi(\{\{1\}\}, 1)$, i.e. $\{\{1\}\} \subset 1$, is false as $\{1\}$ is not an element of 1. Hence the formula $\Phi(x, y)$ is not absolute for M.

Although the formula Φ is Σ_0 and hence simple in terms of our notion of quantifier complexity, it fails to be absolute for M. Our aim in utilising the Lévy hierarchy, however, is to build a rich class of statements that describe as many mathematical notions as possible while being absolute. Absoluteness guarantees that the truth value of sentences holds in all (transitive) models.

The problematic above with the class M and the subset-formula Φ stems from the fact that M is not transitive. In order to verify that "simple" formulas such as Φ above are indeed absolute for M, we will always consider transitive classes.

The ZF-axioms were the result of an attempt to provide a system that gives rise to a mathematical universe behaving most closely to what we expect mathematics to be like. The failure of absoluteness undermines this goal and yields different interpretations of notions between models. In some cases, this is wanted: the method of *forcing* is built upon exactly this idea. In other cases, however, we run into unwanted predicaments: the formula "x is a cardinal", for example, is not absolute, and hence we can find inner models M of ZF in which ω_1^M is not a cardinal.

Proposition 5.24. *Assume the axioms of* ZF *and let M be a transitive class. Then the subset-formula $\Phi(x, y)$ as described above is absolute for M.*

The proof is very similar to the proof of lemma 5.23.

Proof. We verify the definition, hence we need to show both U-absoluteness and D-absoluteness for M. Let $x, y \in M$ be given.

For U-absoluteness, we need to show that if

$$\forall z \in M\ (z \in x \to z \in y) \tag{†}$$

then

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

But as before, as M is transitive, we have that both x and y are subsets of M, and hence if z is an element of x it must also be an element of M. Thus, (†) reduces to

$$\forall z\ (z \in x \to z \in y), \tag{‡}$$

as required.

For the other direction, we remain to show that if (‡) holds then

$$\forall z \in M \; (z \in x \rightarrow z \in y)$$

is also true. But this follows immediately by definition. □

We will now leave absoluteness and Lévy hierarchy behind in order to finally focus on the construction of our set-wise defined analogue of LST. However, as we shall see shortly, the notions we have just defined will be of paramount importance throughout, and we will continuously refer to several results form this section when necessary.

5.4.3 Aside: The Mostowski Collapsing Lemma

In the theory we have developed above, we have mostly worked with transitive classes directly. Indeed, for the definition of absoluteness, we insisted on exclusively dealing with transitive classes in the first place.

Of course, not every class is transitive, and the consequences of a non-transitive class on absoluteness of LST-formulas can be severe, as example 27 showed.

The following theorem, which we include for completeness, helps in that respect:

Theorem 5.25 (Mostowski Collapsing Lemma)**.** *Let X be a set such that X models* E*, the axiom of extensionality (such sets are also called extensional). Then there is a unique transitive set M and a unique bijection π such that*

$$\langle X, \in \rangle \cong \langle M, \in \rangle \, .$$

That is, the structures $\langle X, \in \rangle$ and $\langle M, \in \rangle$ are isomorphic. We call M the transitivisation of X.

Using this theorem, we may pass from one model into another and thereby obtain a transitive set. As we have illustrated the versatility of transitive sets (and classes, in particular) in terms of absoluteness results, this is a highly useful theorem.

Proof. The proof is omitted, the interested reader can find details in [Dev17, pp. 22-3] □

5.5 The Language $\mathscr{L}_V$

The task in this subsection will be to translate LST-formulas into sets. This procedure will be done in multiple steps: we begin by defining *sequences* that will uniquely identify the atomic formulas of LST. Later, we will define how to build LST-formulas of higher complexity using sets, analogously mimicking the way we construct formulas in LST (i.e. using repeated negation and conjunction, and by introducing existential quantifiers).

Throughout this section, we will try and always consider both the syntactic definition of our construction as well as the semantic interpretation in our constructed language $\mathscr{L}$.

Remark. *As mentioned earlier, the use of somewhat complicated notation will be prevalent in this section. However, the way in which we code LST-formulas into sets is not paramount to advancing the theory. In fact, so long as our encoding is consistent, it does not matter at all.*

We begin by defining the notation we shall use for sequences as needed throughout. Note that we will mostly be working with *finite sequences.*

We adopt Devlin's notation for the sake of good readability and conformity.

- A sequence with domain 1 and element x is denoted by

  $$\langle x \rangle.$$

  Recursively, we now define a sequence with domain n with elements $x_0, \dots, x_{n-1}$ to be denoted by

  $$\langle x_0, \dots, x_{n-1} \rangle.$$

- Two sequences s and t can be concatenated: if

  $$s = \langle x_0, \dots, x_{m-1} \rangle$$

  and

  $$t = \langle y_0, \dots, y_{n-1} \rangle$$

  then we write

  $$s^\frown t := \langle x_0, \dots, x_{m-1}, y_0, \dots, y_{n-1} \rangle$$

  and hence $s^\frown t$ has domain $m + n$.

- If a sequence s is finite, then we denote the greatest element of $\mathrm{dom}(s)$ by $||s||$. That is, if $s = \langle x_0, \dots, x_{n-1} \rangle$, i.e. s has domain $n = \{0, 1, \dots, n-1\}$, then

  $$||s|| = n - 1$$

  and hence

  $$x_{||s||} = x_{n-1}$$

  denotes the last element of the sequence s.

Remark. *Note that following our definition we do not consider a sequence with domain n to be an n-tuple. This distinction will be necessary later.*

We will now begin to code our language into sets. Our primary goal is to obtain a consistent but also reasonably convenient method of coding LST-formulas.

Secondly, we aim to build a dictionary of LST-formulas and their translations into sets. This will render the actual translation process much clearer.

Remark. *This work is only preliminary to the theory of the constructible universe. Note that once we have completed the translation process, i.e. once we are able to express any LST-formula as a unique set, we will be able to formalise several metamathematical ideas (such as absoluteness and satisfaction) in set theory; we compare external formulas (these are formulas in LST) with internal formulas (these are sets).*

We would like to give one final remark before we begin the actual coding process: of course, the translation we are about to do is completely analogous to Gödel's famous process of identifying elements of a language by natural numbers (using *Gödel numbers*). Our procedure is only different insofar as it concerns exclusively sets (which will be, incidentally, in most cases natural numbers as well).

5.5.1 Coding *LST*-formulas

We may now begin outlining our construction. Each coding definition has three steps: firstly, we *define* a set X in $\mathscr{L}_V$ to be interpreted in some specific way (for instance, sets of a special form will be interpreted as variables in $\mathscr{L}_V$). Secondly, we *find* an LST-formula Φ that determines the construction of X. Finally, we *prove* that Φ holds if and only if X is as desired, i.e. if and only if the semantic interpretation of Φ in $\mathscr{L}_V$ coincides with our semantic definition of X.

As in previous sections, we may not write out the LST-formula in detail but use the general (unambiguous) shorthands we have used before. This crucially relates to the fact that, in most of our cases, the LST-rendering is obvious.

As outlined in the definition of LST, we need to code variables. This will be our starting point.

Code 5.26. *If a set x is an ordered pair of the form*

Set definition:
$$(2, n)$$

for any $n \in \omega$ then we say that <u>x is a variable in $\mathscr{L}_V$</u>. We denote x by v_n. The associated LST-formula is

LST-formula:	Vbl(x)
$$x \text{ is an ordered pair} \wedge (x)^2_0 = 2 \wedge (x)^2_1 \in \omega.$$	

Proof. We need to prove that a set x is a variable in $\mathscr{L}_V$ if and only if Vbl(x). This clearly follows from the definition. □

Most of these equivalence proofs will follow immediately from the definition, as seen in the previous proof. Hence we will omit the proof and only focus on those which are not immediate.

Remark. *The reader should be aware that when we say $x = (2, n)$ is a variable in $\mathscr{L}_V$ this really is a shorthand for saying that we interpret x as the variable v_n in $\mathscr{L}_V$.*

As we remarked earlier, LST itself does not have any constant symbols. As per our plan set out to construct $\mathscr{L}_V$, we require every element x of V to be a constant symbol in $\mathscr{L}_V$. Clearly, simply taking x as the constant symbol defeats the purpose of our coding process (what if $x = (2, 1)$, for example?).

Hence we continue in a fashion analogous to the case above and make the following definition:

Code 5.27. *If a set x is an ordered pair of the form*

Set definition:
$$(3, y)$$

for any set y then we say that <u>x is the constant symbol y of $\mathscr{L}_V$</u>. We denote x by $\mathring{y}$. The associated LST-formula is

LST-formula:	Const(x)

$$x \text{ is an ordered pair} \wedge (x)^2_0 = 3.$$

One note of caution regarding notation is crucial: above, we have used the overhead-circle-notation in order to indicate constant symbols. The usage of this notation will not be limited to constant symbols. In fact we shall extend it to special symbols in the language, such as predicate symbols, as well.

Remark. *As indicated before, we do not necessarily write out the entire LST-formula formally. Rather, we use many of the shorthands we derived in the section on absoluteness. It is clear by the table at the end of section 5.4.2 that we may rewrite each informal sentence as an LST-analogue very easily.*

Having variables and constant symbols set up, we are ready to code atomic *LST*-formulas. As atomic formulas also include bracket symbols of *LST*, we need to uniquely identify these with sets as well.

Code 5.28. *The set* 1 *is to be interpreted as the open bracket LST-symbol* (*. Similarly, we interpret* 2 *as the closed bracket LST-symbol*)*.*

From now on, we will not explain every single symbol in detail but give the set which is supposed to describe the required expression directly and then add further comments on its structure.

Code 5.29. *If a set x is a sequence of the form*

Set definition:

$$\langle 0, 4, y_1, y_2, 1 \rangle$$

for any $\mathscr{L}_V$-variables or $\mathscr{L}_V$-constants y_1, y_2 then we say that x describes membership of y_1 in y_2 in $\mathscr{L}_V$. We write $y_1 \in y_2$. The associated LST-formula is

LST-formula:	Memb(x)

$$x \text{ is a sequence of length } 5 \wedge (x)^5_0 = 0 \wedge (x)^5_1 = 4 \wedge (x)^5_2 = y_1 \wedge (x)^5_3 = y_2 \wedge (x)^5_4 = 1$$
$$\wedge\, (\text{Const}(y_1) \vee \text{Vbl}(y_1)) \wedge (\text{Const}(y_2) \vee \text{Vbl}(y_2)).$$

Analogously, we describe equality of y_1 and y_2 in $\mathscr{L}_V$ by

Set definition:

$$\langle 0, 5, y_1, y_2, 1 \rangle.$$

and denote the associated LST-formula by Equal(x)*.*

Having defined the blocks making up atomic formulas, we are now ready to describe the formula "a set x is an atomic formula" in $\mathscr{L}_V$.

Code 5.30. *If a set x is a sequence of the form*

$$\langle 0, 4, y_1, y_2, 1 \rangle$$

or of the form

$$\langle 0, 5, y_1, y_2, 1 \rangle$$

for any $\mathscr{L}_V$-constants or $\mathscr{L}_V$-variables y_1 and y_2 then we say that x is an atomic formula in $\mathscr{L}_V$. The associated LST-formula is

LST-formula:	$\mathrm{AFml}(x)$
$\mathrm{Memb}(x) \vee \mathrm{Equal}(x).$	

Proof. By the definition of atomic formulas in $\mathscr{L}_V$, it is clear that we have obtained equivalence of $\mathrm{AFml}(x)$ and the sentence "x is an atomic formula of $\mathscr{L}_V$", as required. □

Here, we have made the crucial step away from LST into our new language $\mathscr{L}_V$: in order to check whether a given $\mathscr{L}_V$-formula x (note this is a set!) is atomic, we need to check that it is of the required form given above which can be expressed in sets exclusively. That is, we now have all the components, defined in terms of sets so that we can build new formulas. However, the way in which we perform this check is described in a metamathematical way: we find a suitable LST-formula.

From now on, we will write down formulas in $\mathscr{L}_V$ and denote them by lower case letters of the Greek alphabet, as mentioned previously.

As per the definition of LST, our next goal is to write any valid LST-formula in $\mathscr{L}_V$ (and hence build the required analogue).

Code 5.31. *Let ϕ and ψ be $\mathscr{L}_V$-formulas. Then we define*

$$\phi \wedge \psi$$

by

$$\langle 0, 6\rangle^\frown \phi^\frown \psi^\frown \langle 1\rangle.$$

Similarly, we define

$$\neg\phi$$

by

$$\langle 0, 7\rangle^\frown \phi^\frown \langle 1\rangle$$

and

$$\exists u\ \phi$$

by

$$\langle 0, 8, u\rangle^\frown \phi^\frown \langle 1\rangle$$

where u is an $\mathscr{L}_V$-variable.

Note the use of $6, 7$ and 8 as unique indicators for the semantic interpretation of the formulas, respectively. As before, the equivalence between LST and $\mathscr{L}_V$ follows immediately from the previous results.

Remark. *One final remark on the definitions above may be permitted: note that, as was our plan all along, each of the terms in the definitions above are sets. We have successfully translated LST-formulas into sets, and are now able to write any LST-formula we like in a consistent and unique way in terms of sets and hence in $\mathscr{L}_V$, as required.*

5.5.2 Taking the Construction Further

As mentioned in the previous sections, we need to guarantee absoluteness in order to be able to construct a strong theory as we intend to. The connection between absoluteness and the level in the Lévy hierarchy as outlined in theorem 5.22 produces this link.

We have made the required definitions in order to be able to construct $\mathscr{L}_V$-formulas in an analogous fashion to the construction of LST-formulas. Our next aim is to define an LST-formula that determines whether a given set is indeed a formula in $\mathscr{L}_V$. In order to guarantee absoluteness, we would like this formula to be of a rather low complexity. As we shall see, the resulting formula will be Σ_1.

We will define the required LST-formula in multiple steps. By definition, we know how formulas (i.e. sets) in $\mathscr{L}_V$ are constructed. Hence, the most natural way of checking whether a given set is an $\mathscr{L}_V$-formula is to verify whether it "looks" like one.

Remark. *As outlined previously, whenever defining new formulas we may now ignore the set definition and focus on the LST-analogue that guarantees equivalence.*

Firstly, any set that is a formula is a finite sequence. Hence we define the following:

Definition 5.32. Let x be a set. Then

LST-formula:	$\mathrm{FSeq}(x)$

$$x \text{ is a sequence} \wedge \forall u \in \mathrm{dom}(x)\ (u \text{ is a natural number}) \\ \wedge\ \exists v \in \mathrm{dom}(x)\ \forall u \in \mathrm{dom}(x)\ (u \in v \vee u = v)$$

holds if and only if "x is a finite sequence"

The equivalence is immediate.

Remark. *Notationally, we will regard sequences as functions from now on. This renders the various LST-formulas much more readable and allows natural shorthand such as the below used $x(0)$, for example.*

Now, for each type of atomic $\mathscr{L}_V$-formula ϕ, we need an associated *LST*-formula that determines whether a given set x is an atomic $\mathscr{L}_V$-formula:

Definition 5.33. Let x be a set and let v_1, v_2 be $\mathscr{L}_V$-variables or constants. Then define

LST-formula:	$F_{\in}(x, v_1, v_2)$

$$\mathrm{FSeq}(x) \wedge \mathrm{dom}(x) = 5 \wedge x(0) = 0 \wedge x(4) = 1 \wedge x(1) = 4 \wedge x(2) = v_1 \wedge x(3) = v_2.$$

Now $F_{\in}(x, v_1, v_2)$ holds if and only if x is the $\mathscr{L}_V$-formula $(v_1 \in v_2)$. Similarly, we define

LST-formula:	$F_{=}(x, v_1, v_2)$

$$\mathrm{FSeq}(x) \wedge \mathrm{dom}(x) = 5 \wedge x(0) = 0 \wedge x(4) = 1 \wedge x(1) = 5 \wedge x(2) = v_1 \wedge x(3) = v_2.$$

for the associated $\mathscr{L}_V$-formula $(v_1 = v_2)$.

In the definitions above, the only difference can be found in the second coordinate of the five-element sequence, as it is necessary in order to be consistent with code 5.29.

Remark. *One final reminder: recall that if we say that x is the $\mathscr{L}_V$-formula $(v_1 \in v_2)$, then we mean we interpret the set x in $\mathscr{L}_V$ in such a way.*

We define analogous formulas for the remaining two types of atomic formulas:

Definition 5.34. Let x, y, z be sets. Then define

LST-formula:	$F_{\wedge}(x, y, z)$

$$\mathrm{FSeq}(x) \wedge \mathrm{FSeq}(y) \wedge \mathrm{FSeq}(z) \\ \wedge\ \mathrm{dom}(x) = \mathrm{dom}(y) + \mathrm{dom}(z) + 3 \\ \wedge\ x(0) = 0 \wedge x(||x||) = 1 \wedge x(1) = 6 \\ \wedge\ \forall i \in \mathrm{dom}(y)\ (x(i+2) = y(i)) \\ \wedge\ \forall i \in \mathrm{dom}(z)\ (x(i + \mathrm{dom}(y) + 2) = z(i)).$$

Now $F_{\wedge}(x, y, z)$ holds if and only if y and z are $\mathscr{L}_V$-formulas and x is the $\mathscr{L}_V$-formula $(y \wedge z)$.

Proof of equivalence: We explain the structure of the LST-formula above in detail: for x to be the $\mathscr{L}_V$-formula $y \wedge z$, we require x, y and z to be $\mathscr{L}_V$-formulas in the first place. Hence they are all finite sequences, which explains the first line of the LST-formula.

As we construct new formulas by concatenating respective sequences, the domain (and hence the length) of x depends on y and z. Thus $\mathrm{dom}(x) = \mathrm{dom}(y) + \mathrm{dom}(z) + 3$. The three extra elements of x are the open bracket (identified by a 0 in the first position of x), the identifier 6 as per code 5.31, and the closed bracket in the last position. Note that $x(\mathrm{dom}(y) + \mathrm{dom}(z) + 2) = x(||x||)$ as x has length $\mathrm{dom}(y) + \mathrm{dom}(z) + 3$, which explains $x(||x||) = 1$ as the code for the closed bracket.

The fourth and fifth line explain the concatenation of the sequences y and z outlined in code 5.31.

The equivalence now follows from the uniqueness of x given $\mathscr{L}_V$-formulas y and z. □

In a very similar fashion we define the remaining simple constructions:

Definition 5.35. Let x, y be sets. Then define

LST-formula: $F_\neg(x, y)$

$$\begin{gathered} \mathrm{FSeq}(x) \wedge \mathrm{FSeq}(y) \\ \wedge\ \mathrm{dom}(x) = \mathrm{dom}(y) + 3 \\ \wedge\ x(0) = 0 \wedge x(||x||) = 1 \wedge x(1) = 7 \\ \wedge\ \forall i \in \mathrm{dom}(y)\ (x(i+2) = y(i)). \end{gathered}$$

Now $F_\neg(x, y)$ holds if and only if y is an $\mathscr{L}_V$-formula and x is the $\mathscr{L}_V$-formula $(\neg y)$.

Definition 5.36. Let x, u, y be sets. Then define

LST-formula: $F_\exists(x, u, y)$

$$\begin{gathered} \mathrm{FSeq}(x) \wedge \mathrm{FSeq}(y) \\ \wedge\ \mathrm{dom}(x) = \mathrm{dom}(y) + 4 \\ \wedge\ x(0) = 0 \wedge x(||x||) = 1 \wedge x(1) = 8 \\ \wedge\ x(2) = u \wedge \forall i \in \mathrm{dom}(y)\ (x(i+3) = y(i)). \end{gathered}$$

Now $F_\exists(x, u, y)$ holds if and only if y is an $\mathscr{L}_V$-formula, u is an $\mathscr{L}_V$-variable, and x is the $\mathscr{L}_V$-formula $(\exists u\ (y))$.

We give an example of the translation (or "encoding") os an LST-formula into $\mathscr{L}_V$.

Example 28. *Consider the LST'-formula (here, LST' denotes the usual language of set theory but also includes the constant symbols* 5 *and* 7*)*

$$\exists x\ (x = 5 \wedge x = 7).$$

We work our way inside out and decompose the formula into the components

$$\exists x \qquad x = 5 \qquad x = 7.$$

We can now use the machinery we have developed and work out the unique sets x, y and z that we interpret as $\exists x$, $x = 5$ and $x = 7$, respectively.

- *Firstly, note that x is a variable in LST'. Hence we need to express x as a variable in $\mathscr{L}_V$ first, which we can do using the sequence $\langle 2, 0\rangle$, for example.*

- *The numbers 5 and 7 are constant symbols in LST', and hence, in order to express them in $\mathscr{L}_V$, we write them as $\langle 3, 5\rangle$ and $\langle 3, 7\rangle$, first. (Usually, we would express these sequences as $\mathring{5}$ and $\mathring{7}$, but in order to visualise the construction we shall refrain from doing so in this particular instance.)*

- *The equality $x = 5$, for instance, is given by the sequence*

$$\langle 0, 5, (2, 0), (3, 5), 1\rangle.$$

  We concatenate the respective sequences and embrace them using brackets (i.e. 0 and 1 in the first and last position, respectively), and the unique identifier for equality, 5, in the second position. Note that $(2, 0)$ and $(3, 5)$ are ordered pairs; no concatenation takes place here!

- *In order to obtain the conjunction $(x = 5 \wedge x = 7)$, we again concatenate and hence obtain*

$$\langle 0, 6\rangle^\frown\langle 0, 5, (2, 0), (3, 5), 1\rangle^\frown\langle 0, 5, (2, 0), (3, 7), 1\rangle^\frown\langle 1\rangle$$

  and hence

$$\langle 0, 6, 0, 5, (2, 0), (3, 5), 1, 0, 5, (2, 0), (3, 7), 1, 1\rangle.$$

- *For the existence part, let ϕ be any $\mathscr{L}_V$-formula (i.e. a set). Then, identifying x with $(2, 0)$ already, $\exists x\ \phi$ is coded as*

$$\langle 0, 8, (2, 0)\rangle^\frown\phi^\frown\langle 1\rangle$$

  according to our construction.

Finally we can concatenate all the sequences and hence obtain

$$\langle 0, 8, (2, 0)\rangle^\frown\langle 0, 6, 0, 5, (2, 0), (3, 5), 1, 0, 5, (2, 0), (3, 7), 1, 1\rangle^\frown\langle 1\rangle$$

or, in its concatenated form,

$$\langle 0, 8, (2, 0), 0, 6, 0, 5, (2, 0), (3, 5), 1, 0, 5, (2, 0), (3, 7), 1, 1, 1\rangle$$

which is the required (and unique) set, as necessary.

It is clear from the example above that this construction is not very convenient to work with directly. Even for a fairly simple formula as the one given above, the code is quite long and fairly cumbersome to decipher. But we will of course never work with the codes directly: the whole point about defining this consistent set of codes is to allow us to use the shorthands we have defined ($x \in y$, for all x, etc.) but to implicitly deal with sets. This example is merely an illustration of the procedures that are necessary in order to perform the encoding.

The underlying concept based on which we have defined atomic $\mathscr{L}_V$-formulas and the construction of $\mathscr{L}_V$-formulas of higher complexity was supposed to be analogous to the construction of LST-formulas. Using the definitions above we have succeeded in doing so. In order to find an LST-formula that determines whether a given set x is indeed an $\mathscr{L}_V$-formula, we consider the following reasoning: by the definition of any formula ϕ of $\mathscr{L}_V$, it has been constructed by successive application of negation, conjunction, or existential quantification. Thus we can find a sequence $\phi_0, \phi_1, \ldots, \phi_n$ such that $\phi_n = \phi_0$ and such that each ϕ_i is either atomic or generated using (some of the) previous formulas ϕ_j for $j < i$.

This idea will be fundamental to the following definition. As before, the structure of our construction allows us to deduce the formula determining whether a given set is an $\mathscr{L}_V$-formula quite easily:

Definition 5.37. Let x, y be sets. Then define

LST-formula: $\mathrm{Build}(x, y)$

$$\begin{aligned} \mathrm{FSeq}(y) \wedge y(||y||) = x \wedge \forall i \in \mathrm{dom}(y)\ [(&\mathrm{AFml}(y_i)) \\ &\vee \exists j, k \in i\ (F_\wedge(y_i, y_j, y_k)) \\ &\vee \exists j \in i\ (F_\neg(y_i, y_j)) \\ &\vee \exists j \in i\ \exists u \in \mathrm{ran}(x)\ (\mathrm{Vbl}(u) \wedge F_\exists(y_i, u, y_j))]. \end{aligned}$$

Now $\mathrm{Build}(x, y)$ holds if and only if y is a finite sequence of $\mathscr{L}_V$-formulas and x can be constructed in a recursive manner from the formulas in y using the rules we have defined previously.

Proof of equivalence: The proof follows immediately from the discussion preceding the definition. □

From the above definition we can deduce the main result of this subsection: a set ϕ is a formula in $\mathscr{L}_V$ if and only if $\exists \psi\ (\mathrm{Build}(\phi, \psi))$.

Definition 5.38. Let x be a set. Then define

LST-formula: $\mathrm{Fml}(x)$

$$\exists y\ (\mathrm{Build}(x, y)).$$

Remark. *Up to this point, we have referred to all formulas in $\mathscr{L}_V$ as sets. As we are now able to build new formulas and check whether they are indeed formulas, we will now begin denoting $\mathscr{L}_V$-formulas by lower case Greek letters.*

5.5.3 Quantifier Complexity of *LST*-formulas defining $\mathscr{L}_V$

In this section, we will consider the *LST*-formulas we have defined previously and examine their quantifier complexity. Of course, in order to guarantee absoluteness, we would like as many as possible to be Σ_0.

As is obvious from the last remark in the previous subsection, we see that the *LST*-formula "a set x is a formula in $\mathscr{L}_V$" is at least Σ_1. Indeed, we require the unbounded quantifier in order to verify whether a suitable sequence of $\mathscr{L}_V$-formulas exists so that we can build x from them.

Our aim is to build complex Σ_0 formulas from simple ones. Lemma 5.20 will be of immense help. It will be used repeatedly throughout.

Proposition 5.39. *The formulas* $\mathrm{Vbl}(x), \mathrm{Const}(x), \mathrm{Memb}(x)$ *and* $\mathrm{Equal}(x)$ *are all* Σ_0 *formulas of LST.*

Proof. The proofs follow immediately from the respective definitions in *LST* as well as the table in section 5.4.2. □

Proposition 5.40. *The LST-formula* FSeq(x) *is* Σ_0.

The proof is almost immediate. We extend the proof by Devlin given in [Dev17, p. 33].

Proof. By definition, the formula is given by

$$x \text{ is a sequence} \wedge \forall u \in \operatorname{dom}(x)\ (u \text{ is a natural number})$$
$$\wedge\ \exists v \in \operatorname{dom}(x)\ \forall u \in \operatorname{dom}(x)\ (u \in v \vee u = v).$$

We have shown most of these to be Σ_0 already. Indeed, the only part we need to consider in detail is the bounded quantifier

$$\forall u \in \operatorname{dom}(x)\ (\phi(u)),$$

which we need to write out in *LST*. In order to do so, we recall the definition of a sequence: any sequence x is in fact a function and hence a set of ordered pairs. So, rather than quantifying over the domain of x, it suffices to quantify over x itself, and remark that any such element $y \in x$ is of the form (y_0, y_1). Clearly, we are interested in the value of the sequence (or function), and hence we may write

$$\forall y \in x\ (\phi((y)_1))$$

where, according to our definition, $(y)_1$ denotes y_1, as required.

Finally, we remain to show that if $\phi(y_0, \vec{w})$ is Σ_0, then so is $\phi((y)_0, \vec{w})$ for some ordered pair $y = (y_0, y_1)$. We rewrite $\phi((y)_0, \vec{w})$ as

$$\exists u \in y\ \exists a \in u\ \exists b \in u\ (y = (a, b) \wedge \phi(a, \vec{w})).$$

This follows directly from the definition of the Kuratowski pair which, as a reminder, is defined by

$$y = (y_0, y_1) \Leftrightarrow y = \{\{y_0\}, \{y_0, y_1\}\}.$$

The case for $(x)_1$ is very similar; we simply replace $\phi(a, \vec{w})$ by $\phi(b, \vec{w})$. This formula above captures exactly what we want to express, and is further clearly Σ_0, which completes the proof. □

Using the latter half of the previous proof the following corollary is immediate.

Corollary 5.41. *All the $F_\bullet$-LST-formulas from section 5.5.2 are* Σ_0.

Finally, we verify the complexity of Build(ϕ, ψ):

Corollary 5.42. *The LST-formula* Build(ϕ, ψ) *is* Σ_0.

The proof given here is a more detailed version of Devlin's approach which can be found in [Dev17, p. 35].

Proof. It is not necessary to examine the entire formula in detail. The fact that the formula is Σ_0 follows directly from lemma 5.20 as well as from the latter half of the proof of proposition 5.40. Indeed, consider the formula

$$\forall i \in \mathrm{dom}(\psi)\ \exists j \in i\ (F_\neg(\psi_i, \psi_j)).$$

Intuitively, this formula should indeed be Σ_0, but in order to verify this we need to rephrase the formula and get rid of the shorthands ψ_i. This is easily done: the formula can be rewritten as

$$\forall i \in \mathrm{dom}(\psi)\ \exists j \in i\ \exists a \in \mathrm{ran}(\psi)\ (F_\neg(a, b) \wedge a = \psi_i \wedge b = \psi_j)$$

and hence is is clearly Σ_0. Similarly, for

$$\forall i \in \mathrm{dom}(\psi)\ \exists j \in i\ \exists u \in \mathrm{ran}(\phi)\ (\mathrm{Vbl}(u) \vee F_\exists(\psi_i, u, \psi_j))$$

we can find an equivalent formula

$$\forall i \in \mathrm{dom}(\psi)\ \exists j \in i\ \exists u \in \mathrm{ran}(\phi)\ \exists a, b \in \mathrm{ran}(\psi)(\mathrm{Vbl}(u) \vee (F_\exists(a, u, b) \wedge a = \psi_i \wedge b = \psi_j)).$$

Another formula,

$$\forall i \in \mathrm{dom}(\psi)\ \exists j, k \in i\ (F_\wedge(\psi_i, \psi_j, \psi_k))$$

can be rewritten as

$$\forall i \in \mathrm{dom}(\psi)\ \exists j, k \in i\ \exists a, b, c \in \mathrm{ran}(\psi)\ (F_\wedge(a, b, c) \wedge a = \psi_i \wedge b = \psi_j \wedge c = \psi_k)$$

which, after applying lemma 5.20, is also Σ_0. By a completely analogous reasoning we can show that $\forall i \in \mathrm{dom}(\psi)\ (\mathrm{AFml}(\psi_i))$ is also Σ_0, which completes the proof. □

Remark. *We used a couple of shorthands here. Writing out the entire formula would be very cumbersome. As a reminder we would like to note, however, that ψ_i really is short for $\psi(i)$, where ψ is by definition a sequence, and hence a function, in particular.*

The desired corollary follows (as Devlin remarks in [Dev17, p. 35]):

Corollary 5.43. *The LST-formula "x is a formula in $\mathscr{L}_V$" is Σ_1.*

Proof. The formula is given by

$$\exists y\ (\mathrm{Build}(x, y))$$

which is Σ_1 by the previous corollary and lemma 5.20. □

The fact that $\mathrm{Fml}(x)$ is Σ_1 gives rise to a problem: our main goal is to produce absoluteness results so that our theory is applicable as possible. For Σ_0 formulas, the absoluteness follows from theorem 5.22 (note we still assume all classes we are working with are transitive). For a Σ_1 formula such as $\mathrm{Fml}(x)$ we need to verify that it is also Π_1 in order to deduce its absoluteness. The hypotheses in theorem 5.22 require us to work with a subtheory of ZF. Naturally, in order to guarantee our absoluteness results to be as widely applicable as possible, we would like our theory, T say, to be rather weak; then, any transitive model (and hence any model of a stronger theory than T) will preserve the absoluteness.

In his text [Dev17], Devlin introduces a weaker theory called *Basic Set Theory*, or BS for short. The aside in section 5.5.6 gives further details on BS and its role. However, in this report, we will carry on working in ZF for ZF provides us with everything we need in order to develop the theory without getting hung up on technical details.

5.5.4 Defining Satisfaction

In the course of our development, we have arrived at a crucial stage: it is now time to translate the notion of "truth" from *LST* into $\mathscr{L}_V$.

Our aim is to build the constructible universe from sets that can be defined (or "constructed") from sets we have constructed before (intuitively, such sets will be simple as we can relate them to a formula – more on this later). Hence, we may only have access to some sets, and not all sets that V has to offer. The following definition allows us to consider the formal language $\mathscr{L}$ augmented with a restricted number of individual constant symbols.

Definition 5.44. Let u be a set. Then we define the sublanguage $\mathscr{L}_u$ to be the language $\mathscr{L}_V$ excluding the constant symbols $\mathring{v}$ for each $v \notin u$. In particular, $\mathring{u} \notin \mathscr{L}_u$.

If u is the empty set, we simply write $\mathscr{L}$.

Remark. *The sublanguage $\mathscr{L}_u$ gives rise to all those $\mathscr{L}_V$-formulas whose analogue in LST is relativised with respect to u.*

For each such sublanguage $\mathscr{L}_u$, we now define analogues of $\mathrm{Const}(x), \mathrm{AFml}(x)$, etc. verifying membership of x to u:

Definition 5.45. Let u be a set. Then define

LST-formula:	$\mathrm{Const}(x, u)$
$\mathrm{Const}(x) \wedge (x)_1 \in u.$	

We may also write $\mathrm{Const}_u(x)$ for the sake of improved readability.

Of course, the choice of u only influences the constant symbols existing in $\mathscr{L}_u$. Hence, in order to adjust existing formulas, such as $\mathrm{AFml}(x)$, to take into account this difference, it suffices to replace $\mathrm{Const}(x)$ by $\mathrm{Const}_u(x)$ in all such formulas, and hence obtain an analogous statement defined in $\mathscr{L}_u$.

Definition 5.46. Let u be a non-empty set. Then define $\mathrm{AFml}(x, u)$ to be the formula $\mathrm{AFml}(x)$ with all instances of $\mathrm{Const}(y)$ appearing in $\mathrm{AFml}(x)$ replaced by $\mathrm{Const}_u(y)$. Similarly, define $\mathrm{Fml}(x, u)$.

Lemma 5.47. *The formulas $\mathrm{AFml}_u(x)$ and $\mathrm{Const}_u(x)$ are both Σ_0.*

Proof. Both results follow from the respective construction of $\mathrm{AFml}_u(x)$ and $\mathrm{Const}_u(x)$ as well as from the fact that $\mathrm{AFml}(x)$ and $\mathrm{Const}(x)$ are both Σ_0. □

In order to define satisfaction, we need to find a procedure by which we are able to verify whether, for any set u, a given formula ϕ is true in the structure $\langle u, \in\rangle$. Clearly, theorems in a given language must be constructed using the atomic formulas, and hence once we have found a way of checking whether a given formula can be constructed from the atomic formulas of $\mathscr{L}_u$, we are halfway there. As all $\mathscr{L}_V$-formulas are sets, and, in particular, sequences, it suffices to check whether there is a recursive procedure that yields the required sequence.

As such, we would like to investigate the finite sequences in $\mathscr{L}_u$. In order to guarantee absoluteness, the following definition and proposition will be required:

Definition 5.48. Let u, a and n be sets. Define

LST-formula:	$\mathrm{Seq}(u, a, n)$

$$\exists f\ (\mathrm{FSeq}(f)\ \wedge\ n \text{ is a natural number}$$
$$\wedge\ \mathrm{dom}(f) = n \wedge u = \bigcup \mathrm{ran}(f)$$
$$\wedge\ \forall i \in \mathrm{dom}(f)\ \forall x \in f(i)\ (\mathrm{FSeq}(x) \wedge \mathrm{dom}(x) = i \wedge \forall j \in i\ (x(j) \in a))$$
$$\wedge\ \forall i \in \mathrm{dom}(f)\ \forall j \in i\ \forall x \in f(j)\ \forall p \in a\ (i = j+1 \rightarrow (x \cup \{(p, i)\} \in f(i)))))$$

Now $\mathrm{Seq}(u, a, n)$ holds if and only if u is the set set of all sequences of length at most n exclusively comprising elements of a.

Proposition 5.49. *The LST-formula* $\mathrm{Seq}(x)$ *is* Δ_1^{ZF}.

We describe the proof in [Dev17, p. 37].

Proof. From the definition, it is clear that $\mathrm{Seq}(u, a, n)$ is Σ_1. As the set "constructed" by $\mathrm{Seq}(u, a, n)$ is a set of finite sets, and since all finite sets exist in ZF, we have verified that $\mathrm{Seq}(u, a, n)$ is Σ_1^{ZF}. For absoluteness with respect to ZF, we now need to find a Π_1 formula $\phi(u, a, n)$ so that

$$\mathrm{ZF} \vdash \mathrm{Seq}(u, a, n) \leftrightarrow \phi(u, a, n).$$

It is easily verified that

$$\mathrm{ZF} \vdash \mathrm{Seq}(u, a, n) \leftrightarrow (n \text{ is a natural number} \wedge \forall z\ (\mathrm{Seq}(z, a, n) \rightarrow z = u))$$

which is Π_1, as required. □

Remark. *Note that the right hand side of the equivalence is* Π_1 *although* $\mathrm{Seq}(u, a, n)$ *is* Σ_1. *This holds as the formula* $\mathrm{Seq}(u, a, n) \rightarrow z = u$ *is clearly equivalent to* $\neg\,\mathrm{Seq}(u, a, n) \vee z = u$. *Now, by applying lemma 5.20, we obtain the result.*

Theorem 5.50. *The LST-formula* $\mathrm{Fml}(x)$ *is* Δ_1^{ZF}.

Proof. The proof is omitted. Details can be found in [Dev17, pp. 37-8]. □

We may extend this formula in the same way as above and hence obtain the following result:

Corollary 5.51. *The formula* $\mathrm{Fml}(x, u)$ *is* Δ_1^{ZF}.

In order to progress to the actual theory, we take a shortcut at this point and omit writing out the formulas for the following ingredients of the *LST*-analogue. Instead, we give the following theorem without proof:

Theorem 5.52. *Let* ϕ *and* x *be sets. Then there exists a* Δ_1^{ZF} *LST-formula* $\mathrm{Fr}(\phi, x)$ *such that*

$$\mathrm{Fr}(\phi, x) \leftrightarrow\ \phi \textit{ is an } \mathscr{L}_V\textit{-formula and } x \textit{ is the set of variables occurring free in } \phi.$$

Similarly, let ϕ', ϕ, v *and* t *be sets. Then there exists a* Δ_1^{ZF} *LST-formula* $\mathrm{Sub}(\phi, \phi, v, t)$ *such that*

$$\mathrm{Sub}(\phi', \phi, v, t) \leftrightarrow \phi' \textit{ is the } \mathscr{L}_V\textit{-formula obtained by replacing every instance of the variable } v \textit{ in the } \mathscr{L}_V\textit{-formula } \phi \textit{ by the constant } t.$$

Proof. The proof is omitted. Details can be found in [Dev17, pp. 38-40].

A brief outline is given by the following: in both cases, we define an LST-formula that is closest to our intuitive way of translating the semantics into LST. Both these formulas are Σ_1.

- For $\mathrm{Fr}(\phi, x)$, we iterate over the components making up ϕ (we can do this using $\mathrm{Build}(\phi, \psi)$, for example). At each stage (i.e. at each member of ψ), consider the free variables. Now, whenever we move to a lower stage within the iteration, we remove a tentatively free variable if it occurs in one of the members of ϕ as a bound variable. Once we have completed the iteration we are left with exactly those variables that are free in ϕ.
- For $\mathrm{Sub}(\phi', \phi, v, t)$, we define an auxiliary formula $S(\phi', \phi, v, t)$, that determines substitution for atomic formulas. Then, using a similar approach to Fr, we consider a sequence (using Build) that constructs ϕ, and at each stage containing an atomic formula, we check the substitution using S.

Finally, as before, we find a Π_1 formula in LST in order to prove the Δ_1^{ZF} property. □

Remark. *At this point, it is not clear that investigating the structure of formulas in $\mathscr{L}_u$ will give any insight on the truth of such a formula. Absoluteness, however, maintains this crucial link, and, similarly, a vital theorem later will indeed verify the analogy and successful translation of "truth" from $\mathscr{L}_u$ into LST.*

Finally, we are able to define truth in $\mathscr{L}_V$.

We follow the approach Devlin outlines in [Dev17]: we aim to define an LST-formula $\mathrm{Sat}(u, \phi)$ such that $\mathrm{Sat}(u, \phi)$ holds if and only if the $\mathscr{L}_u$-formula ϕ is a sentence and is true in the structure $\langle u, \in \rangle$.

Theorem 5.53. *Let ϕ and x be sets. Then there exists a Δ_1^{ZF} LST-formula* $\mathrm{Sat}(\phi, x)$ *such that*

$$\mathrm{Sat}(\phi, x) \leftrightarrow \phi \textit{ is a sentence in } \mathscr{L}_x \textit{ which is true in the structure } \langle x, \in \rangle.$$

Proof. We only give an outline of the proof presented by Devlin, details can be found in [Dev17, pp. 40-1].

We define functions f and g with domain ω recursively such that $f(0)$ is the set of all atomic formulas of $\mathscr{L}_x$ and $f(i+1)$ yields the set of all formulas obtained from applying one of the formula construction rules (conjunction, negation, existential quantification) to the $\mathscr{L}_x$-formulas in $f(i)$.

Clearly, not all of these formulas are sentences. Here, we use the function g and define it so that $g(i)$ is the set of all sentences in $f(i)$ which are true in $\langle x, \in \rangle$. Also, we will define $f(i)$ so that it shall also contain all formulas obtained at previous stages $f(j)$; the same will hold for g.

Now, for any formula ϕ we consider the least i such that $\phi \in f(i)$ (such an i exists by the definition of f and by the method we have constructed ϕ with). If ϕ is also in $g(i)$, then we have $\mathrm{Sat}(\phi, x)$, as required.

The following remarks are crucial:

- Due to the existence of a least positive integer i as described above, we will technically only be required to consider finite sequences.

- Clearly, a formula with free variables can be turned into a sentence if and only if its free variables are bounded by a quantifier. In order to verify that a formula is indeed a sentence, we use Sub and hence perform this particular check. Furthermore, it is clear that, by the construction of f, we can construct formulas that are not sentences only by introducing variables.

In order to translate "truth" in the structure $\langle x, \in\rangle$, we need to find LST-formulas that hold if and only if the truth in $\langle x, \in\rangle$ is determined. Naturally, following the intrinsic construction of all LST-formulas, we begin by considering the atomic formulas: for example, consider the LST-formula

$$AT(\phi, x) \leftrightarrow \phi \text{ is an atomic } \mathscr{L}_x\text{-formula and it is true in } \langle x, \in\rangle.$$

We can rewrite this so that

$$AT(\phi, x)$$

is the formula

$$\exists y_1, y_2 \in x \ (y_1 \in y_2 \wedge F_\in(\phi, \mathring{y}_1, \mathring{y}_2)) \vee \exists y \in x \ (F_=(\phi, \mathring{y}, \mathring{y})).$$

This expresses precisely what we aimed to describe: the formula $AT(\phi, x)$ is true if and only if ϕ is an atomic $\mathscr{L}_x$-formula which is true in $\langle x, \in\rangle$.

We then write out an LST-formula with free variables ϕ and x that "builds" f and g as required. In the last part of the LST-formula, we check whether the given formula ϕ is actually an element of $g(||g||)$. If it is, we have obtained truth in $\langle x, \in\rangle$. □

Remark. *Rather than writing* $\text{Sat}(\phi, x)$*, we shall write*

$$\models_x \phi$$

in the style of model theory. We can in fact interpret this as "x models ϕ", which provides us with the analogy.

The technical work is done, and we can now state the crucial equivalence we have been working towards:

Theorem 5.54 (The Correctness Theorem)**.** *Let* $\Phi(\vec{v}^n)$ *be any LST-formula, and suppose that* $\phi(\vec{v}^n)$ *is its analogue in* $\mathscr{L}$*. That is,* $\phi(\vec{v}^n)$ *has been constructed in* $\mathscr{L}$ *to have the same structure as* $\Phi(\vec{v}^n)$*. Then the following equivalence holds:*

$$\text{ZF} \vdash \forall u \ \forall \vec{x}^n \in u^n \ (\Phi^u(\vec{x}^n) \leftrightarrow \text{Sat}(\phi(\vec{\mathring{x}}^n), u))$$

Note that we can rewrite this as

$$\text{ZF} \vdash \forall u \ \forall \vec{x}^n \in u^n \ (\Phi^u(\vec{x}^n) \leftrightarrow \models_u \phi(\vec{\mathring{x}}^n)).$$

Hence we have related the truth of an LST-formula to its $\mathscr{L}$ counterpart in a 1-to-1 fashion. The result follows from the consistent recursive construction of $\mathscr{L}_u$ based on LST. The result is proven by induction on the complexity of Φ; the details are omitted.

5.5.5 Absoluteness again

Now that we have mastered the transition from LST to $\mathscr{L}$, we need to translate many of the metatheoretical notions into $\mathscr{L}$, too.

We adapt the notion of quantifier complexity to formulas of $\mathscr{L}$ as follows: we define the Lévy hierarchy for formulas in $\mathscr{L}$ in exactly the same way as we did in definition 5.17 for LST-formulas. However, due to the formal construction of formulas in $\mathscr{L}$, analysing their quantifier complexity will be much easier if we only allow single quantifiers and not blocks of quantifiers. (If one recalls the formal definition for quantifiers in $\mathscr{L}$, then it is clear why blocks of like quantifiers render the construction of formulas much more complex.)

Let M, N be structures and assume M is a substructure of N. We can now internalise the metamathematical notions we have defined in LST earlier and express them within set theory:

Definition 5.55. Let ϕ be an $\mathscr{L}$-formula. If

$$\forall \vec{x} \in M \; (\models_M \phi(\vec{\mathring{x}}) \text{ implies } \models_N \phi(\vec{\mathring{x}}))$$

then we call ϕ U-absolute for M, N. Similarly, if

$$\forall \vec{x} \in M \; (\models_N \phi(\vec{\mathring{x}}) \text{ implies } \models_M \phi(\vec{\mathring{x}}))$$

then ϕ is D-absolute for M, N.

Further, we may deduce the following analogy:

Lemma 5.56. *Let M, N be transitive $\mathscr{L}$-structures. Assume M is a substructure of N, as before. If ϕ is an $\mathscr{L}$-formula then*

- *if ϕ is Σ_0 then ϕ is absolute for M, N;*
- *if ϕ is Σ_1 then ϕ is U-absolute for M, N;*
- *if ϕ is Π_1 then ϕ is D-absolute for M, N.*

The proofs are almost identical with those given in theorem 5.22 and hence omitted.

Finally, we can state the crucial correspondence that will be used throughout:

Theorem 5.57 (The Second Correctness Theorem). *Let $\Phi(\vec{x})$ be a Σ_0 formula in LST. If $\phi(\vec{x})$ is its analogue in $\mathscr{L}$, then the following holds:*

$$\text{ZF} \vdash \text{“For any transitive set } M,\ \forall \vec{x} \in M \; (\Phi(\vec{x}) \leftrightarrow \models_M \phi(\vec{\mathring{x}}))\text{”}$$

Intuitively, we here obtain an equivalence very similar to theorem 5.54. Indeed, this result is analogous to absoluteness for transitive classes and guarantees that the truth of Σ_0-formulas in LST-formulas coincides with the truth of the associated $\mathscr{L}$-formulas.

5.5.6 Aside: Basic Set Theory

Our main goal is to find notions that are absolute for as many transitive models as possible. In order to guarantee such a strong level of absoluteness, it is vital to prove the absoluteness in a theory as basic as possible. Then we have automatically verified absoluteness in all extension of that theory.

We consider the following theory called *Basic Set Theory*:

Axioms of BS

Extensionality:	E	$\forall x\, \forall y\, (\forall z\, (z \in x \leftrightarrow z \in y) \to x = y)$
Induction:	IN	$\forall \vec{w}^n\, (\forall x\, ((\forall y \in x\, (\Phi(y, \vec{w}^n))) \to \Phi(x, \vec{w}^n)) \to \forall x\, \Phi(x, \vec{w}^n))$ for any LST-formula Φ with free variables among $x, \vec{w}^n$.
Pairing:	PA	$\forall x\, \forall y\, \exists z\, (x \in z \wedge y \in z)$
Union:	U	$\forall \mathscr{F}\, \exists A\, \forall Y\, \forall x\, (x \in Y \wedge Y \in \mathscr{F} \to x \in A)$
Cartesian Product:	CA	$\forall x\, \forall y\, \exists z\, \forall u\, (u \in z \leftrightarrow \exists a \in x\, \exists b \in y\, (u = (a, b)))$
Infinity:	I	$\exists x\, (0 \in x \wedge \forall y \in x\, (y + 1 \in x))$
Σ_0-Comprehension:	Σ_0C	$\forall \vec{w}^n\, \forall x\, \exists y\, \forall z\, (z \in y \leftrightarrow (z \in x \wedge \Phi(\vec{w}^n, z)))$ for any Σ_0 LST-formula $\Phi(\vec{w}^n, z)$.

Note that BS is a subtheory of ZF: the axioms E, PA, U, and I are exactly the same as in ZF. Similarly, we can prove CA using the definition of the Kuratowski pair as well as P. Further, IN is clearly a theorem of ZF.

Further, remark that IN and Σ_0C are both schemas with one axiom for each LST-formula. In C we have the Cartesian product expressed implicitly, i.e. not in its full LST-rendering.

We could now continue our investigation of the syntactic structure of the formulas we examined in the previous subsection. Indeed, if we achieve absoluteness in BS then we also do so in any transitive model (and hence, in particular, in any model of ZF).

If we extend BS by one additional axiom schema, we obtain *Kripke-Platek Set Theory*, or KP for short. This theory is of immense set theoretical importance as it is the weakest subtheory of ZF that allows the construction of the constructible universe (see [Dev17, p. 48] for details).

The axiom schema to add is called the Σ_0 *Collection Schema*:

$$\forall \vec{a}\, (\forall x\, \exists y\, (\Phi(y, x, \vec{a})) \to \forall u\, \exists v\, \forall x \in u\, \exists y \in v\, (\Phi(y, x, \vec{a})))$$

for any LST-formula Φ that is Σ_0.

As mentioned before, however, we will only work in ZF as it suffices for our needs of introducing the crucial concepts behind the constructible universe.

5.5.7 Definability

In order to construct the constructible universe, we need to define what "constructible" actually means. This crucial notion will be our building block for the different levels of the constructible universe later on.

In the course of the following sections, it will be imperative to consider extensions of our language $\mathscr{L}_u$. In particular, we would like to add finitely many predicate letters $\mathring{A}_i$. We go about doing so as follows: fix a positive integer k and define

$$A_1 \subset u^{n_1}, \ldots, A_k \subset u^{n_k}$$

where $n_1, \ldots, n_k$ are positive integers, respectively. In order to add these subsets A_i of u^{n_i}-tuples to our language, we describe them as predicate letters

$$\mathring{A}_1, \mathring{A}_2, \ldots, \mathring{A}_k$$

of arity n_i, respectively. Hence

$$\mathring{A}_i(a) \text{ is true if and only if } a \in A_i.$$

We denote the resulting language by

$$\mathscr{L}_u(\mathring{A}_1, \ldots, \mathring{A}_k).$$

Remark. *Previously, we used the ring notation in order to signify sets as constants of the respective language. In this case, however, we use the ring notation so as to avoid confusion between the set* A_i *and the predicate letter* $\mathring{A}_i$ *that "indicates" membership of the set* A_i*.*

In the previous section, we translated sentences from LST into our new language $\mathscr{L}$ that is made up exclusively of sets. This manifests a transition from the metatheory into the actual set theoretical realm we strived to obtain. Due to our constructions, we may easily apply our translation conventions to our extended language $\mathscr{L}_u(\mathring{A}_1, \ldots, \mathring{A}_k)$.

Similarly, we will have to formalise the ideas of absoluteness in order to guarantee consistency of absolute notions as we strived in LST. Section 10 of Chapter 1 in [Dev17, pp. 44-8] gives a detailed analysis of how this can be achieved; we resort to giving an outline here.

In the previous section, we focussed on absoluteness of formulas with respect to ZF for transitive classes. In this section, we adapt this procedure with two changes:

- we consider definability rather than absoluteness;
- classes will not only be transitive but also amenable.

All the results that can be derived will be formally defined within set theory, i.e. they are not metamathematical constructions. So far we have done the following: in LST, we have

- considered a formula Φ; and
- shown that Φ is absolute with respect to ZF for *transitive* classes.

We will now try and emulate this in $\mathscr{L}$. As we will not have access to LST-formulas in $\mathscr{L}$, it does not make sense to define definability in the same way in which we defined absoluteness with respect to a theory. Instead, we consider classes which model that respective theory. Hence

- we consider a formula ϕ; and
- show that ϕ is definable with respect to M, an *amenable* set that "models" ZF.

Why are these changes required? Within LST, we aimed to preserve the notion of our formulas when switching models of ZF. In $\mathscr{L}$, we are more concerned about definability. Further, by the remark above, we require classes in order to determine definability for $\mathscr{L}$-formulas. Such classes must model the theory of interest, ZF in our case. Amenable classes are exactly those which "model" BS, basic set theory, and are therefore sufficient for our needs.

Given a set M, it will be convenient for us to associate an $\mathscr{L}$-formula $\phi(v_1, \ldots, v_n)$ with the set of all n-tuples $\vec{a}^n$ in M^n such that $\phi(\vec{\mathring{a}}^n)$ holds. Here we make use of the language extension defined above. This gives rise to the following:

Definition 5.58. Let M be a set and suppose $N \subset M$. We say a set $R \subset M^m$ is $\underline{\Sigma_n^M(N)}$ if there exists a Σ_n formula $\phi(v_0, \ldots, v_m)$ in $\mathscr{L}_M$ such that

- if $\mathring{a}$ is a constant symbol in ϕ then $a \in N$; and
- $\forall \vec{x}^m \in M^m \; (\vec{x}^m \in R \leftrightarrow \models_M \phi(\vec{\mathring{x}}^m))$.

If $M = N$, we write $\Sigma_n(M)$, and if $N = \emptyset$ we write Σ_n^M.

We extend the definition naturally to the analogous cases of $\Pi_n^M(N)$ and $\Delta_n^M(N)$ formulas.

Note that the Σ_n-formula mentioned above is an $\mathscr{L}_M$-formula, *not* an LST-formula. A good way of remembering the latter notational conventions is the following:

- if $N = M$, we need to use constants in M in order to find a suitable formula; in some sense, we actively consider the constants in M within the $\mathscr{L}_M$-language;
- if, however $N = \emptyset$, then we do not require any constants in M in order to find a suitable Σ_n-formula. There is no need to implicitly augment the language. (Of course, no actual augmentation takes place, as $\mathscr{L}_M$ already contains all the constant symbols $\mathring{a}$ for all $a \in M$; this notation only illustrates that we do not formally need such constant symbols.)

Note that if R is Σ_n^M, then it is trivially also $\Sigma_n(M)$. This should be clear from the explanation above: assuming the existence of specific constant symbols is a weaker hypothesis than not permitting any such.

Definition 5.59. Let $R \subset M^m$. We say that $\underline{R \text{ is } M\text{-definable}}$ if R is $\Sigma_n(M)$ for some n.

Note that the previous definition makes sense. Indeed, it covers all formulas once we notice that any formula can be expressed in this form using lemma 5.20 part (vii).

Finally, the following definition wraps up our discussion of definability.

Definition 5.60. Let A be a class of m-tuples. Consider a class of structures $\{M_\alpha : \alpha < \gamma\}$ for some $\gamma \in \mathbf{ON}$. Then A is $\underline{\text{uniformly } \Sigma_n^M}$ if there exists a Σ_n^M-formula $\phi(v_1, \ldots, v_m)$ for which

$$A \cap M_\alpha^m = \{\vec{x}^m : \models_{M_\alpha} \phi(\vec{\mathring{x}}^m)\}$$

for all $\alpha \in \gamma$.

Note that the required Σ_n-formula must hold for all M_α *at the same time*. This is exactly the reason why we do not allow any constant symbols (and hence insist on Σ_n^M formulas) as any constant symbols would have to depend on the structures M_γ but hold for all of them simultaneously.

Uniform definability will allow us to prove definability for a host of classes at the same time. In particular, using this definition, we will shortly prove definability of formulas for all L_α at limit stages $\alpha > \omega$. This clearly holds as any Σ_n^M formula is in particular $\Sigma_n(M)$ as remarked above (cf. definition 5.59).

5.6 The Constructible Universe

Now that we have the tools available to work within set theory, we are able to formally define the constructible universe. Note that the entire construction is done within set theory. Due to our coding procedure, there is no need to consider LST (indeed, we do not have access to LST-formulas within the universe as it comprises sets only; the metatheory, however, may still be described by LST-formulas, as we have done before).

Definition 5.61. A set y is called x-definable if there is an $\mathscr{L}_x$-formula $\phi(v_0)$ (where v_0 is a variable in $\mathscr{L}_x$) such that

$$y = \{a \in x : \models_x \phi(\mathring{a})\}.$$

That is, y can be expressed as the set of elements $a \in x$ that satisfy an $\mathscr{L}_x$-formula ϕ so that $\phi(\mathring{x})$ is true in $(x, \in)$.

For a given set x, the set of all x-definable sets is defined by

$$\mathrm{Def}(x) = \{\{a \in x : \models_x \phi(\mathring{a})\} : \phi(v_0) \text{ is some } \mathscr{L}_x\text{-formula}\}.$$

In other words, $\mathrm{Def}(x)$ is the set of all those sets that can be identified by a formula ϕ which is modelled by x.

By iterating over this definition, we build new sets from somewhat simple sets (we understand the sets in $\mathscr{L}_x$ quite well). Further, and crucially, every set in the constructible universe (which is made up of definable sets) can be identified with an $\mathscr{L}$-formula, similar to the equivalence between classes and LST-formulas we mentioned previously.

Definition 5.62. The constructible hierarchy of sets is defined inductively as follows: set

$$L_0 = \emptyset,$$

and define

$$L_{\alpha+1} = \mathrm{Def}(L_\alpha).$$

Finally, if α is a limit ordinal, define

$$L_\alpha = \bigcup_{\beta<\alpha} L_\beta.$$

Now define the constructible universe, denoted by L, to be

$$L = \bigcup_{\alpha \in \mathbf{ON}} L_\alpha.$$

Remark. *As we shall see later, we allow sufficiently many elements to belong to each L_α so that the power set axioms still holds within L. Crucially, though, we restrict the number of elements so that the peculiarities of the unrestricted power set operation that are prevalent in V disappear. (We will use the notion of inner models to prove this in due course.)*

It is clear that $\mathrm{Def}(L_\alpha) \subset \mathcal{P}(L_\alpha)$. The following result will be used throughout:

Lemma 5.63. *Let $\Phi(\vec{v}^{n+1})$ be an LST-formula and consider a set X. Then the set*

$$Y = \left\{y \in X : \Phi^X(y, \vec{v}^n)\right\}$$

is an element of $\mathrm{Def}(X)$.

Proof. The proof follows from theorem 5.54: by definition of definability, $Y \in \mathrm{Def}(X)$ if and only if there exists an $\mathscr{L}$-formula $\phi(v)$ such that $Y = \{y \in X : \models_X \phi(\mathring{y})\}$. By theorem 5.54, we have equivalence of truth between the two languages, i.e. $\models_X \phi(\mathring{y})$ if and only if $\phi^X(y)$ for all $y \in X$, as required. □

In our attempt to grasp the constructible universe, we prove the following results:

Proposition 5.64. *Let L be the constructible universe. Then the following hold:*

(i) if $\beta \leq \alpha$ then $L_\beta \subset L_\alpha$;

(ii) for each α, the set L_α is transitive;

(iii) for each α we have $L_\alpha \subset V_\alpha$, and for $\alpha \leq \omega$, we have equality;

(iv) if $\alpha < \beta$, then α and L_α are elements of L_β;

(v) for each α, the following equality holds:

$$L \cap \alpha = L_\alpha \cap \mathbf{ON} = \alpha;$$

(vi) for each $\alpha \geq \omega$, we have $|L_\alpha| = |\alpha|$.

We will require the following very easy lemmata:

Lemma 5.65. *Let $(A_\alpha : \alpha \in \mathbf{ON})$ be a hierarchy of transitive sets. Then the union $\bigcup_{\alpha \in \mathbf{ON}} A_\alpha$ is also transitive.*

Proof. Let $x \in \bigcup_{\alpha \in \mathbf{ON}} A_\alpha$. Then, by definition, $x \in A_{\alpha'}$ for some $\alpha' \in \mathbf{ON}$. Using transitivity of $A_{\alpha'}$, we see that $x \subset A_{\alpha'}$, and as $A_{\alpha'} \subset \bigcup_{\alpha \in \mathbf{ON}} A_\alpha$, the result is proved. □

Lemma 5.66. *Let α be an infinite ordinal. Then $\mathscr{L}_{L_\alpha}$ has cardinality $|L_\alpha|$.*

Proof. By definition, the language LST is countable. Hence, by our faithful translation of LST into sets, so is $\mathscr{L}$. As we have augmented the language with constant symbols taken from L_α, we see that $\mathscr{L}_{L_\alpha}$ has cardinality $\max(\aleph_0, |L_\alpha|)$. Hence, in particular, if α is infinite, then $\mathscr{L}_{L_\alpha}$ has cardinality $|L_\alpha|$, as required. □

A short remark before we give the proof of proposition 5.64 will aid understanding: note that if $x \in V_{\alpha+1}$, then $x \in \mathcal{P}(V_\alpha)$, by the definition of the cumulative hierarchy. Hence, in particular, $x \subset V_\alpha$. We will use this fact in the proof below.

Proof of proposition 5.64. The proofs are based on Devlin's approach in [Dev17, pp. 59-60]. Further comments and explanations have been added accordingly.

(i/ii) We prove the result by induction on α for both results (i) and (ii) simultaneously. The advantages of this approach will be obvious.

- For $\alpha = 0$, we have $L_\alpha = \emptyset$, which is trivially a subset of any L_β. Similarly, the empty set is trivially transitive.
- Assume $x \in L_\alpha$. For the successor case, it suffices to show that $L_\alpha \subset L_{\alpha+1}$. If we now use the inductive hypothesis for (ii), then $x \subset L_\alpha$. Thus, we may write

$$x = \{y \in L_\alpha : \models_{L_\alpha} \text{“}\mathring{y} \in \mathring{x}\text{”}\} \in \mathrm{Def}(L_\alpha) = L_{\alpha+1},$$

as required. Note that this only works as the LST-formula $y \in x$ is Σ_0, and hence absolute for transitive classes. Applying the correctness theorems now, we obtain the result.

(This is the strength of the correctness theorems: for simple formulas, we may consider the LST-formula and deduce absoluteness without the need to go through the hassle of explicitly considering the $\mathscr{L}$-analogue.)

For (ii), we assume that L_α is transitive. We need to show that $L_{\alpha+1}$ is also transitive. Assume that $x \in y \in L_{\alpha+1}$, hence we are required to show that $x \in L_{\alpha+1}$. Recall that $\mathrm{Def}(L_\alpha) \subset \mathcal{P}(L_\alpha)$, and hence if $y \in L_{\alpha+1}$, then $y \subset L_\alpha$. Thus $x \in L_\alpha$, and by (i), we have that $L_\alpha \subset L_{\alpha+1}$, which completes the proof.

– Finally, if α is a limit ordinal and L_β is transitive for all $\beta < \alpha$, then L_α is transitive by the previous lemma. Similarly, as $L_\alpha = \bigcup_{\beta<\alpha} L_\beta$, (i) follows immediately, too.

(iii) We begin by showing that $L_\alpha \subset V_\alpha$ for every $\alpha \in \mathbf{ON}$. We will use an easy inductive argument: by definition, $L_0 = V_0$. If $L_\alpha = V_\alpha$, then

$$L_{\alpha+1} = \mathrm{Def}(L_\alpha) \subset \mathcal{P}(L_\alpha) \subset \mathcal{P}(V_\alpha) = V_{\alpha+1}.$$

At limit stages, we simply combine all the elements at the successor stages, and hence the inclusion holds.

Now assume $\alpha < \omega$. From the previous part, it suffices to show that $V_\alpha \subset L_\alpha$. We proceed by induction and suppose that $V_\alpha = L_\alpha$. We now exploit the fact that definability allows a finite number of parameters and that V_α is finite: let $x \in V_{\alpha+1}$. Note that, by its construction, $x \subset V_\alpha$, which equals L_α by assumption. It is clear that we can define any set $\{a_1, \dots, a_n\} \subset V_\alpha = L_\alpha$ by

$$\{a \in L_\alpha : \models_{L_\alpha} \text{“}\mathring{z} = \mathring{a}_1 \vee \ldots \vee \mathring{z} = \mathring{a}_n\text{”}\}$$

which is Σ_0 and hence definable (see definition 5.59). Thus

$$\{a \in L_\alpha : \models_{L_\alpha} \text{“}\mathring{z} = \mathring{a}_1 \vee \ldots \vee \mathring{z} = \mathring{a}_n\text{”}\} \in \mathrm{Def}(L_\alpha) = L_{\alpha+1}$$

and so $V_\alpha \subset L_\alpha$ which yields $V_{\alpha+1} = L_{\alpha+1}$, as required. By the construction of L, we see that

$$L_\omega = \bigcup_{\beta<\omega} L_\beta = \bigcup_{\beta<\omega} V_\beta = V_\omega$$

for all limit ordinals λ, where the second equality holds by the inductive result above.

(iv) We can use the results in (i). Hence it suffices to show that if $\alpha \in \mathbf{ON}$ then $\alpha \in L_{\alpha+1}$ and $L_\alpha \in L_{\alpha+1}$. The fact that $L_\alpha \in L_{\alpha+1}$ follows immediately from the rendering

$$L_\alpha = \{x \in L_\alpha : \models_{L_\alpha} \text{“}\mathring{x} = \mathring{x}\text{”}\},$$

which is clearly L_α-definable and hence an element of $L_{\alpha+1}$.

For the case of $\alpha \in L_{\alpha+1}$, we proceed by induction. Assume $\gamma \in L_{\gamma+1}$ for all $\gamma < \alpha$. By (i), $\gamma \in L_\alpha$ as $L_\gamma \subset L_\alpha$ for all $\gamma < \alpha$. As α is the set comprising all ordinals $\gamma < \alpha$, we hence see that $\alpha \subset L_\alpha$ (clearly, this follows as α comprises all those ordinals γ which are smaller than itself and each such $\gamma \in L_\alpha$). Using transitivity, we see that $\gamma \subset L_\alpha$. Hence we may deduce that

$$\alpha = L_\alpha \cap \mathbf{ON}\,.$$

The LST-formula $\mathbf{ON}(v_0)$ (which holds if and only if v_0 is an ordinal) is Σ_0. Therefore, by the second correctness theorem, it is absolute for transitive classes, and hence in particular for L_α (by (ii)). Thus

$$\alpha = \{z \in L_\alpha : \models_{L_\alpha} \text{“}\,\mathbf{ON}(\mathring{z})\text{”}\}$$

which is, as before, L_α-definable. Thus $\alpha \in \mathrm{Def}(L_\alpha) = L_{\alpha+1}$, as required.

(We require to explicitly use the fact that the LST-formula “x is an ordinal” is absolute as the element α we consider when constructing L_α is an element of the universe, and not internalised within L. Absoluteness as well as the second correctness theorem allow the faithful translation we require in order to prove the result.
Note that, technically, we have to use the $\mathscr{L}$-analogue of the LST-formula $\mathbf{ON}(v_0)$. However, we could also invoke lemma 5.63 and use the LST-formula directly. Hence we do not make a formal distinction above.)

(v) In the previous proof, we have shown that $L_\alpha \cap \mathbf{ON} = \alpha$. Now, as $L_\alpha \subset L$ and $\alpha \subset \mathbf{ON}$, it follows immediately that $L \cap \alpha = \alpha$, as required.

(vi) From (v) it follows immediately that $|\alpha| \leq |L_\alpha|$, hence only one inequality remains to be shown. We again proceed by induction. As we have already shown that V_α and L_α coincide for all $\alpha \leq \omega$, induction on $\alpha \geq \omega$ suffices.

- The base case $L_\omega = V_\omega = \omega$ is immediate by virtue of (iii).
- Assume γ is a limit ordinal and suppose that $|L_\alpha| = |\alpha|$ for all $\alpha < \gamma$. Then we use the properties of cardinal arithmetic of section 3.4 in order to write

$$|L_\gamma| = \left|\bigcup_{\alpha<\gamma} L_\alpha\right| \leq \sum_{\alpha<\gamma} |L_\alpha| = \sum_{\alpha<\gamma} |\alpha| = |\gamma|,$$

  where the penultimate equality holds by the inductive hypothesis.
- Now assume that $|L_\alpha| \leq |\alpha|$. Then, using lemma 5.66, we may conclude

$$|L_{\alpha+1}| = |\,\mathrm{Def}(L_\alpha)| \leq |L_\alpha| \leq |\alpha| = |\alpha + 1|$$

  as the L_α-definable sets must be representable by $\mathscr{L}_{L_\alpha}$-formulas.

Thus the proof is complete. □

Note that part (iv) in the theorem above proves that all ordinals are elements of L, or equivalently

$$\mathbf{ON} \subset L.$$

Further, part (vi) shows that there are sets that are not definable by formulas (otherwise $V = L$ were true; we will consider this special case shortly).

Remark. *As the reader will have noticed, we have expressed the required $\mathscr{L}$-formulas used above in quotation marks. This is done to visualise that the actual formula is, of course, a set, that would formally have to be written out with terms available to us in $\mathscr{L}$ only. In this case, we wold have to make repeated use of $F_\wedge$ and $F_=$, for example.*

As is done regularly in mathematics, we consider structures that give rise to interesting theories and results. The following result verifies that the constructible universe is of interest to us.

Definition 5.67. Let T be a subtheory of ZF. We call a transitive proper class M an inner model of T if T proves Φ^M for all Φ in T. In symbols, we have

$$T \vdash \Phi^M.$$

for each $\Phi \in T$.

The following theorem is fundamental to the entire theory of constructibility.

Theorem 5.68. *The constructible universe is an inner model of* ZF*.*

Unless otherwise stated, we describe the reasoning presented by Devlin in [Dev17, pp. 60-3].

Proof. We follow the natural approach and consider every axiom $\Phi \in$ ZF and show that ZF proves Φ^L.

E: As L is transitive, the fact that L is extensional follows immediately from lemma 5.23. However, in order to illustrate the reasoning, we will give a detailed proof below: as per the definition of the relativised formula, we need to show that

$$\mathrm{ZF} \vdash \mathsf{E}^L.$$

If we unwrap the relativisation, we obtain

$$\mathrm{ZF} \vdash (\forall x\, \forall y\, (\forall z\, (z \in x \leftrightarrow z \in y) \rightarrow x = y))^L$$

which is, by definition, the same as

$$\mathrm{ZF} \vdash \forall x \in L\, \forall y \in L\, (\forall z\, (z \in x \leftrightarrow z \in y) \rightarrow x = y)^L.$$

Using the definition of relativisation yet again, we obtain

$$\mathrm{ZF} \vdash \forall x \in L\, \forall y \in L\, (\forall z \in L\, (z \in x \leftrightarrow z \in y) \rightarrow x = y).$$

Now assume x and $y \in L$ are given. Then, by applying transitivity of L, we know that if $z \in x$ then $z \in L$. Thus we can rewrite the last line as

$$\mathrm{ZF} \vdash \forall x \in L\, \forall y \in L\, (\forall z\, (z \in x \leftrightarrow z \in y) \rightarrow x = y).$$

Now note that $\forall z\, (z \in x \leftrightarrow z \in y) \rightarrow x = y)$ is the axiom of extensionality itself, which is an axiom of ZF. Hence the result holds.

F: For the axiom of foundation, we need to verify that

$$\mathrm{ZF} \vdash (\forall x\, (\exists y\, (y \in x) \rightarrow \exists y\, (y \in x \wedge \neg \exists z\, (z \in x \wedge z \in y))))^L.$$

By the same reasoning as above, we need to show that

$$\mathrm{ZF} \vdash \forall x \in L\, (\exists y \in L\, (y \in x) \rightarrow \exists y \in L\, (y \in x \wedge \neg \exists z \in L\, (z \in x \wedge z \in y))).$$

Let $x \in L$ be given, hence we need to find a $y \in L$ such that $y \in x$ and every element $z \in L$ that is an element of y is not an element of x. We use transitivity and see that if $y \in x$ then $y \in L$. Similarly, if $z \in y$ then $z \in L$. By the actual axiom of foundation, we have

$$\exists y\, (y \in x \wedge \neg \exists z\, (z \in x \wedge z \in y))$$

which is exactly the element we require. Transitivity now yields the result.

PA: For the pairing axiom, we apply exactly the same reasoning as above: we are required to show

$$\mathrm{ZF} \vdash (\forall x\ \forall y\ \exists z\ (x \in z \wedge y \in z))^L$$

which is the same as

$$\mathrm{ZF} \vdash \forall x \in L\ \forall y \in L\ \exists z \in L\ (x \in z \wedge y \in z).$$

Let $x, y \in L$ be given. Define $\alpha = \max(\mathrm{rank}(x), \mathrm{rank}(y))$. Hence $x, y \in L_{\alpha+1}$ (using (i) of proposition 5.64). Now define

$$z = \{w \in L_{\alpha+1} : \models_{L_{\alpha+1}} \text{“}\mathring{w} = \mathring{x} \vee \mathring{w} = \mathring{y}\text{”}\}.$$

This set z is clearly $L_{\alpha+1}$-definable and hence an element of $L_{\alpha+2} \subset L$. Now theorem 5.54 proves the result.

U: We now omit many of the trivial steps applied earlier: assume $x \in L$. We need to find a set $y \in L$ such that an element $z \in L$ is in y if and only if there is some $u \in x$ for which $z \in u$ (this is our understanding of the union of sets). In this case, we take the element provided to us by the actual union axiom and show that it is definable: let $y = \bigcup x$. As $x \in L$, in particular, $x \in L_\alpha$ for some ordinal α. By transitivity, $x \subset L_\alpha$, and as $y = \bigcup x$, we see that y is the union of a subset of L_α and hence a subset of L_α itself. As in the PA-case above, define

$$y = \{z \in L_\alpha : \models_{L_\alpha} \text{“}\exists v \in \mathring{x}\ (\mathring{z} \in v)\text{”}\}.$$

Note that $y \in \mathrm{Def}(L_\alpha) = L_{\alpha+1}$. Again, by invoking theorem 5.54, we have found the set y, as required.

R: For the axiom of replacement, we use Kunen's approach given in [Kun80, p. 169]. We need to show that

$$\mathrm{ZF} \vdash (\forall A\ \forall \vec{w}^n\ (\forall x \in A\ \exists! y\ (\Phi(x, y, A, \vec{w}^n)) \\ \rightarrow \exists Y\ \forall x \in A\ \exists y \in Y (\Phi(x, y, A, \vec{w}^n))))^L.$$

As previously, assume $A, w_1, \dots, w_n \in L$ are given and assume that

$$\forall x \in A\ \exists! y \in L\ (\Phi^L(x, y, A, \vec{w}^n))$$

holds. Consider a function f such that $f(x) = \mathrm{rank}_L(x)$ for all $x \in L$ (note that this is a well-defined function). Applying the actual axiom of replacement, we may consider the ordinal

$$\alpha = \sup(\{f(y) + 1 : \exists x \in A\ (\Phi^L(x, y, A, \vec{w}^n))\}).$$

Finally, set $Y = L_\alpha$. Clearly, Y now satisfies the constraints above, and by part (iv) of proposition 5.64, $Y \in L$, as required.

S: For the axiom of separation, fix an LST-formula Φ and denote its $\mathscr{L}$-analogue by ϕ. We need to show that

$$\mathrm{ZF} \vdash (\forall z\ \forall \vec{w}^n\ \exists y\ \forall x\ (x \in y \leftrightarrow (x \in z \wedge \Phi(x, z, \vec{w}^n))))^L,$$

hence let $z \in L$ as well as parameters $\vec{w} \in L^n$ be given. We want to use the generalised reflection principle: assume $\alpha \in \mathbf{ON}$ such that $x, w_1, \dots, w_n \in L_\alpha$ (clearly, such an ordinal exists by definition of L). We now apply the generalised reflection theorem and hence obtain an ordinal β such that

$$\forall \vec{z} \in L_\beta \, (\Phi^{L_\beta}(\vec{z}) \leftrightarrow \Phi^L(\vec{z})).$$

The generalised reflection theorem is clearly applicable as L is transitive. Now we consider

$$y = \{z \in L_\beta : \models_{L_\beta} \text{“}(\phi(\mathring{z}, \vec{\mathring{w}}) \wedge \mathring{x} \in \mathring{z})\text{”}\}.$$

Clearly, $y \in \mathrm{Def}(L_\beta) = L_{\beta+1}$, hence we have found a level within the cumulative hierarchy which contains y. By the equivalence of LST and $\mathscr{L}$-formulas proven in lemma 5.63, we see that

$$y = \{z \in x : \Phi^{L_\beta}(z, \vec{w})\}.$$

But as β was chosen so that the formula reflects upwards, we have

$$y = \{z \in x : \Phi^L(z, \vec{w})\}$$

from which the result follows as $y \in L_{\beta+1} \subset L$.

P: For the power set axiom, we follow a similar approach to that employed for F. Let y be the set given by the actual power set axiom. We need to show that $y \in L$. Consider $y = \{z \in \mathcal{P}(x) : z \in L\}$. We want to use the axiom of replacement: consider a function f that is defined by $f(z) = \mathrm{rank}(z)$ for each $z \in y$. By replacement, the image of f is a set, and hence consider $\beta = \sup(\mathrm{img}(f)) + 1$. Now β exceeds every $f(z)$, and so $y \subset L_\beta$. But note that

$$y = \{z \in L_\beta : \models_{L_\beta} \text{“}\mathring{z} \subset \mathring{x}\text{”}\} \in \mathrm{Def}(L_\beta) = L_{\beta+1} \subset L$$

as required.

I: The axiom of infinity is proven immediately by noticing that, for example, $\omega \in L_{\omega+1}$ which proves the claim immediately.

Hence the proof is complete. □

This reasoning justifies the nomenclature: assume ZF is consistent and hence suppose a proper transitive class M is a model of ZF. If we construct L within M in the way described above, then L is in fact a proper subclass of M which also models ZF. It is hence an inner model.

Most of the proofs above follow quickly from the transitivity of L. Those which do not follow as easily can be reduced to questions in which we simply need to find a level of the constructible hierarchy which contains the needed elements. Then, we can use the fact that definability lets us define the sets we need directly using appropriate formulas, and the result follows. The proof of R is special as it requires us to apply the generalised reflection principle (see theorem 5.6) first; this, in turn, enables us to find a suitable level in the constructible hierarchy, as required.

The theory we have developed so far gives rise to one crucial question: what happens if we assume that the ground universe of our mathematical discourse is the constructible universe? That is equivalent to saying: what if every set is constructible by assumption? The Axiom of Constructibility postulates that every set is constructible, and it is therefore usually denoted by the shorthand

$$V = L.$$

Of course, the associated formula is given by

$$\forall x\ \exists \alpha\ (x \in L_\alpha).$$

In order to investigate this axiom, at first we improve our understanding of the constructible universe itself. The following definition will be immensely useful:

Definition 5.69. Let M be a transitive set. Then we say M is amenable if the following conditions are satisfied:

(i) $\omega \in M$;

(ii) $\forall x \in M\ (\bigcup x \in M)$;

(iii) $\forall x, y \in M\ (\{x, y\} \in M)$;

(iv) $\forall x, y \in M\ (x \times y \in M)$;

(v) if $R \subset M$ and R is $\Sigma_0(M)$ then $\forall x \in M\ (R \cap x \in M)$.

Amenable sets will provide us with an analogue to transitive sets introduced earlier: we will be able to prove multiple absoluteness results for amenable classes M in the same way in which we proved absoluteness for transitive classes earlier. Note that an amenable class is in fact a "model" of the theory BS. As BS is a subtheory of ZF, amenable sets are of interest to us and will take on the role transitive classes held in previous sections when we considered absoluteness.

Proposition 5.70. *Let α be a limit ordinal greater than ω. Then L_α is amenable.*

We again follow Devlin (cf. [Dev17, pp. 63-4]).

Proof. We verify the definition:

(i) This is true by virtue of the fact that $\omega \in L_{\omega+1}$.

(ii) Assume $x \in L_\alpha$. By definition, there is a least $\beta \in \mathbf{ON}$ such that $x \in L_\beta$. Note that we can write
$$\bigcup x = \{z \in L_\beta : \models_{L_\beta} \text{“}\exists u \in \mathring{x}\ (\mathring{z} \in u)\text{”}\}$$
which is clearly in $\mathrm{Def}(L_\beta) = L_{\beta+1}$.

(iii) Assume $x, y \in L_\alpha$ and denote by β the least ordinal for which $x, y \in L_\beta$. We apply the same trick as before and note that if $w = \{x, y\}$ then
$$w = \{z \in L_\beta : \models_{L_\beta} \text{“}\mathring{z} = \mathring{x} \vee \mathring{z} = \mathring{y}\text{”}\},$$
which is clearly an element of $L_{\beta+1}$, as required.

(iv) Fix any $x, y \in L_\alpha$. As before, there exists a least $\beta < \alpha$ such that $x, y \in L_\beta$. By transitivity, $x \subset L_\beta$ and $y \subset L_\beta$. We need to show that any ordered pair $(a, b) = \{\{a\}, \{a, b\}\}$ with $a \in x$ and $b \in y$ is an element of L_γ for some $\gamma < \alpha$. Then we can use a suitable $\mathscr{L}_{L_\gamma}$-formula in order to show membership of the set $x \times y = \{(a, b) : a \in x \wedge b \in y\}$ of L.

Hence fix $a \in x$ and $b \in y$. By transitivity, $a, b \in L_\beta$. From part (iii) we see that $\{a\}$ and $\{a, b\}$ are elements of $L_{\beta+1}$ (by virtue of the same formula used in the proof of (iii)). Hence we can write
$$x \times y = \{z \in L_{\beta+2} : \models_{L_{\beta+2}} \text{“}\exists a \in \mathring{x}\ \exists b \in \mathring{y}\ (z = (a, b))\text{”}\}$$
which is clearly $L_{\beta+2}$-definable and hence an element of $L_{\beta+3}$. As α is a limit ordinal, we have $\beta + 3 < \alpha$ and so $L_{\beta+3} \subset L_\alpha$. This proves the result.

(v) Let $R \subset L_\alpha$. If R is $\Sigma_0(L_\alpha)$, then there exists a Σ_0 formula $\phi(v_0, \ldots, v_n)$ such that

$$\forall x \in L_\alpha \ (x \in R \leftrightarrow \models_{L_\alpha} \phi(\mathring{x}, \vec{\mathring{a}}^n)) \tag{$*$}$$

for some constants $\vec{\mathring{a}}^n \in L_\alpha$. Fix any $u \in L_\alpha$. We now need to show that $R \cap u \in L_\alpha$. As before, we determine the least ordinal β such that $u, \vec{a} \in L_\beta$. By transitivity, if $u \in L_\beta$, then $u \subset L_\beta$. Thus we may rewrite the equation above as

$$R \cap u = \{x : x \in u \wedge x \in R\}$$

as

$$R \cap u = \{x \in L_\beta : x \in u \wedge x \in R\}.$$

By the absoluteness of Σ_0 formulas in $\mathscr{L}$ (see lemma 5.56), we see that as ϕ is also absolute for L_β, L_α. Hence, by definition

$$\models_{L_\beta} \phi(\mathring{x}, \vec{\mathring{a}}) \leftrightarrow \models_{L_\alpha} \phi(\mathring{x}, \vec{\mathring{a}})$$

holds for all $x \in L_\beta$. Thus, again, may write

$$\begin{aligned} R \cap u &= \{x \in L_\beta : x \in u \wedge x \in R\} \\ &= \{x \in L_\beta : x \in u \wedge \ \models_{L_\alpha} \phi(\mathring{x}, \vec{\mathring{a}})\} \quad \text{(by the equivalence ($*$) shown above)} \\ &= \{x \in L_\beta : x \in u \wedge \ \models_{L_\beta} \phi(\mathring{x}, \vec{\mathring{a}})\} \quad \text{(by absoluteness)} \\ &= \{x \in L_\beta : \models_{L_\beta} (\text{``}\mathring{x} \in \mathring{u}\text{''} \wedge \phi(\mathring{x}, \vec{\mathring{a}}))\} \end{aligned}$$

which is clearly L_β-definable, and hence an element of $L_{\beta+1}$. Again, as $\beta + 1 < \alpha$, the result follows.

Hence the proof is complete. □

This result will be crucial as we have can translate many of the complexity results that were based on the absoluteness of *LST*-formulas into definability results of the respective $\mathscr{L}$ counterparts.

5.6.1 Absoluteness and Definability in *LST* and $\mathscr{L}$

Our goal in this section is to prove that L is indeed an inner model of ZF + $V = L$. We provide an outline of the roadmap we take to obtain the result below.

We have defined the construction of the constructible universe in an informal manner. What we have yet to show is how we can express definability within set theory (and not, as before, in the metatheoretical sense). We do this in the same way in which we coded formulas into sets: we exhibit a suitable *LST*-formula and show that it is absolute for inner models of ZF. However, there are two more step to verify: we need to make sure that each such formula's meaning within set theory coincides with its metamathematical interpretation defined in *LST*. Further, it is integral to check that definability is also preserved when passing from one amenable class to another.

Remark. *In summary, we have to show the following: assume M is an inner model of* ZF*. The techniques used in the following are based on the absoluteness results form the previous section as well as our definition of definability within $\mathscr{L}$. The reasoning is the following: for each LST-formula $\Phi(\vec{v}^n)$ whose $\mathscr{L}$-analogue $\phi(\vec{v}^n)$ we aim to use in our model M, there are two steps to verify:*

- *Firstly, we are required to show that Φ is absolute for transitive models with respect to* ZF*. We will usually do this by proving that $\Phi(\vec{v}^n)$ is Δ_1^{ZF}.*

- *Secondly, we need to show that the class ϕ (this is the subclass of M^n containing all those n-tuples for which $\phi(\vec{\mathring{a}}^n)$ holds) is Δ_1^M.*

Then we have verified that absoluteness is guaranteed between transitive classes, and definability is preserved between amenable classes.

Let Φ be an LST-formula, and suppose that Φ is Δ_1^{ZF}. Recall that L_α is amenable (and hence "model" of BS). If we show that ϕ, the $\mathscr{L}$-analogue of Φ, satisfies

$$\Phi(\vec{v}) \leftrightarrow \models_{L_\alpha} \phi(\vec{\mathring{v}}) \tag{$*$}$$

then we may deduce that ϕ, the set of all n-tuples satisfying $\phi(\vec{v}^n)$, is uniformly $\Sigma_1^{L_\alpha}$, as required. This follows directly from definition 5.60: if $(*)$ holds, then clearly

$$\phi \cap L_\alpha^n = \{\vec{x}^n : \models_{L_\alpha} \phi(\vec{\mathring{x}}^n)\}$$

by definition.

In order to verify the result, the following steps are required in order to turn a Σ_1^{ZF} absoluteness result into a Σ_1^M definability result (where M is amenable):

- we rewrite $\Phi(y)$ as $\exists x\ \Psi(x, y)$, where Ψ is Σ_0;

- we identify the formula Φ by the class it determines. I.e. we write

  $$A_\Phi = \{y : \exists x\ \Psi(x, y)\};$$

- we consider the $\mathscr{L}$-analogue $\psi(x, y)$, and deduce from the second correctness theorem that if $x, y \in L_\alpha$ then

  $$\Psi(x, y) \leftrightarrow \models_{L_\alpha} \psi(\mathring{x}, \mathring{y}).$$

  This is applicable as Ψ is Σ_0 and L_α is amenable and hence transitive. We may deduce that if $y \in L_\alpha$ then whenever $\models_{L_\alpha} \exists x\ \psi(x, \mathring{y})$ then $\exists x\ \Psi(x, y)$, which is one direction of the equivalence we require;

- for the other direction, given $y \in L_\alpha$, we need to show that if there is an x such that $\Phi(x, y)$ holds, then this element x is in fact in L_α. We usually achieve this using the same method we have applied above: we try and define x in a way that allows us to determine at which level x lives within the constructible hierarchy.

  Note, that if we show that $x \in L$ is unique, we have also shown that A_Φ is $\Pi_1^{L_\alpha}$, as required. This holds by our remark given after the proof of proposition 5.49.

We may now use this technique to verify equivalence of formulas in LST and their $\mathscr{L}$-analogues within L. The crucial step here is to formally define definability, as it is required to construct L.

We skip the details (a detailed description of how the following formula can be obtained is presented in [Dev17, pp. 67-9]; all the proofs that are used along the way follow the pattern we have outlined above) and present the following result:

Proposition 5.71. *Consider the LST-formula $H(x,\alpha)$ given by*

$$\exists f\ (G(f,\alpha) \wedge x = f(\alpha))$$

where $G(f,\alpha)$ is a Δ_1^{ZF} formula saying "$f = (L_\gamma : \gamma \leq \alpha)$", for which A_G is uniformly $\Delta_1^{L_\alpha}$ for uncountable limit ordinals α.

Proof. The proof is omitted, details can be found in [Dev17, pp. 69]. ☐

It is clear that $H(x,\alpha)$ holds if and only if x is a level of the constructible hierarchy.
The following result will be crucial. We return to absoluteness:

Theorem 5.72. *Let M be an inner model of* ZF*. Then $L_\alpha \in M$ for any ordinal $\alpha \in \mathbf{ON}$. Further,*

$$(H(x,\alpha))^M \text{ holds only if } x = L_\alpha.$$

In future, we write

$$(L_\alpha)^M = L_\alpha$$

instead of

$$(H(x,\alpha))^M \to x = L_\alpha.$$

Further, we may also write L_α^M rather than $(L_\alpha)^M$.

The original presentation of the next proof can be found in [Dev17, p. 70].

Proof. As the function G is Δ_1^{ZF}, we can construct the sequence $(L_\alpha : \alpha \leq \gamma)$ for any $\gamma \in \mathbf{ON}$ in ZF. Now consider the particular inner model M. By said proposition 5.71, we can construct the sequence $(L_\alpha^M : \alpha < \gamma)$ and have also verified that $L_\alpha^M \in M$ for all such α (this follows since M is an inner model of ZF). But now note that $H(x,\alpha)$ is Δ_1^{ZF} and hence absolute for transitive classes (note that, by definition 5.67, every inner model of a subtheory of ZF is a transitive proper class). Thus

$$\forall x \in M\ \left(H(x,\alpha) \leftrightarrow H(x,\alpha)^M\right),$$

and so, noting that $H(x,\alpha)$ is defined to say "$x = L_\alpha$", we have

$$\forall x \in M\ \left(\text{“}x = L_\alpha\text{”} \leftrightarrow \text{“}x = L_\alpha^M\text{”}\right).$$

This completes the proof. ☐

The following two corollaries are of great importance:

Corollary 5.73. *Let M be an inner model of* ZF*. Then*

$$(L)^M = L.$$

In particular, by theorem 5.68, we have

$$(L)^L = L.$$

The result follows immediately (the original proof can be found in [Dev17, p. 70]).

Proof. By the previous theorem, we see that for any inner model M of ZF, we have $(L_\alpha)^M = L_\alpha$. Hence, as $L = \bigcup_{\alpha \in \mathbf{ON}} L_\alpha$, we have $(L)^M = L$.
The second part follows as L is an inner model of ZF. ☐

The previous theorem is a remarkable result: it states that if we consider the constructible universe and build a class according to the rules outlined at the beginning of this section *within* L, then we obtain L again.

This informal explanation motivates the following corollary:

Corollary 5.74 (The Minimal Model Property)**.** *The constructible universe L is the smallest inner model of* ZF*.*

This proof follows Devlin's approach in [Dev17, p. 77].

Proof. We have seen that L is an inner model of ZF. Assume M is an inner model of ZF. Then, the previous corollary says that

$$(L)^M = L,$$

and hence $L \subset M$. Thus every inner model of ZF is a superclass of L, and so the result follows. □

Finally, we are ready to state the main result of the section:

Theorem 5.75. *The class L is a model of the theory* $\mathrm{ZF} + V = L$*.*

See [Dev17, p. 71] for the original presentation of this proof:

Proof. We have shown that L is an inner model of ZF, hence we only need to show that

$$\mathrm{ZF} \vdash (V = L)^L.$$

From the previous theorem, we see that $(L)^L = L$. But we also see that $(V)^L = L$. Thus we have

$$(V)^L = (L)^L$$

and hence

$$(V = L)^L,$$

as required. □

5.7 The Axiom of Choice in L

As mentioned in the motivation of this section, we will now use the structure of the constructible hierarchy in order to prove that the Axiom of Choice holds in the constructible universe. Formally, we will show that

$$\mathrm{ZF} \vdash (\mathrm{AC})^L.$$

In LST, we can express the Axiom of Choice in the following way:

$$\forall x\ ((\forall y \in x\ (y \neq \emptyset) \wedge \forall y, y' \in x\ (y \neq y' \to \forall w\ (w \in y \to w \notin y'))) \\ \to \exists z\ \forall y \in x\ \exists! v \in y\ (v \in z))$$

In order to prove the Axiom of Choice in L, it clearly suffices to prove any equivalence; we shall prove the following:

Theorem 5.76 (AC in L)**.** *Any set $x \in L$ can be well-ordered.*

The following notion will be used: recall that for an element $x \in L$, we define the rank of x with respect to L, denoted by $\mathrm{rank}_L(x)$, to be the least ordinal α for which $x \in L_{\alpha+1}$.

We will require the following two orderings. Both will be of a purely technical nature. They will enable us to easily give a well-ordering of the elements of L, which, in turn, will allow us to prove that AC holds in L.

Definition 5.77. Consider ordinal sequences $\gamma = (\gamma_1, \dots, \gamma_n)$ and $\delta = (\delta_1, \dots, \delta_m)$. We denote the lexicographic ordering on the ordinal sequences by $<_*$. As a reminder: we say

$$\gamma <_* \delta$$

if and only if

$$n < m$$

OR

$$n = m \wedge \gamma(i) < \delta(i).$$

where $i < n$ is the least integer for which $\gamma(i) \neq \delta(i)$.

Definition 5.78. Let ϕ and ψ be $\mathscr{L}$-formulas. Then we say

$$\phi <_\dagger \psi \tag{(1)}$$

if and only if

$$\phi \text{ is an initial segment of } \psi \tag{(2)}$$

OR

$$m = n \text{ and } k(\phi(i)) < k(\phi(i))$$

where i is the least integer for which $\phi(i)$ and $\psi(i)$ differ and where k is defined by

$$k(x) = \begin{cases} x, & \text{if } x \in 9 \\ n + 9, & \text{if } x \text{ is an } \mathscr{L}\text{-variable and hence of the form } (2, n). \end{cases}$$

Let $x, y \in L$. In our definition of the ordering that will well-order L, one step in our classification will examine the least formulas ϕ and ψ for which $x = \{z \in L_\alpha : \models_{L_\alpha} \phi\}$ and $y = \{z \in L_\alpha : \models_{L_\alpha} \psi\}$. The use of the term "least" above indicates that we will require $<_\dagger$ to be a well-ordering.

Lemma 5.79. *The ordering $<_\dagger$ is a well-ordering.*

Proof. This follows easily from the fact that both ϕ and ψ are actually finite sequences. It is clear that $<_\dagger$ is a linear ordering. In order to show that it is a well-ordering, in view of a contradiction, consider a non-empty set $P = \{\phi_i : i \in \omega\}$ (as we construct a counterexample, we may assume that P is countable) for which

$$\phi_0 > \phi_1 > \dots$$

As all ϕ_i are finite, we cannot have an infinite decreasing chain $\phi_0 > \phi_1 > \dots$ that is solely determined by condition (1) in definition 5.78. Hence there must exist $i \in \omega$ such that $||\phi_i|| = ||\phi_j||$ for all $j > i$. Thus, the linear strict ordering must be determined by condition (2) above for all ϕ_j with $j > i$. But then we would have an infinite strictly decreasing sequence of natural numbers determined by the function k, which is clearly impossible. $\square$

We briefly explain the ordering $<_\dagger$: it reads the finite sequence of sets making up ϕ and ψ and, in case one is not an initial segment of the other, it considers the first element in which they differ and checks whether this element is a key in our language $\mathscr{L}$ (i.e. a number we use to identify brackets, membership, equality, etc. when we constructed $\mathscr{L}$) or whether the element is an ordered pair of the form $(2, n)$, i.e. a variable in $\mathscr{L}$. Note that as $\mathscr{L}$ does not have any constant symbols, this list of cases is exhaustive.

Remark. *As in the previous section in which we constructed the language $\mathscr{L}$, the actual definition of the ordering $<_\dagger$ does not matter; as is obvious from its definition, it is a purely syntactical ordering, and does not attach any particular set theoretical meaning to its elements.*

Lastly, the well-ordering of L we will give below uses one more trick. We know that we can express any set in L using an $\mathscr{L}$-formula (just as we can express any class by an LST-formula). However, as the next lemma shows, we may express such sets in a very special way:

Lemma 5.80. *If x is an element of $L_{\alpha+1}$ (that is, $\operatorname{rank}_L(x) \leq \alpha$), then there exists a formula $\phi(v_0, \dots, v_n)$ as well as ordinals $\gamma_1, \dots, \gamma_n$, where each $\gamma_i < \alpha$, such that*

$$x = \{y \in L_\alpha : \models_{L_\alpha} \phi(\mathring{y}, \mathring{L}_{\gamma_1}, \dots, \mathring{L}_{\gamma_n}).$$

Note this is well-defined as $L_\gamma \in L_\alpha$ by (iv) of proposition 5.64. The proof is based on Devlin's approach in [Dev17, p. 72]:

Proof. We prove the result by induction on α.

- The case $\alpha = 0$ is trivial, as $L_0 = \emptyset$.
- Assume the hypothesis holds for all $\beta < \alpha$. Consider $x \in L_{\alpha+1}$. By the definition of L, there exists an $\mathscr{L}$-formula $\psi(v_0, \dots, v_n)$ and parameters $\vec{p}^n$ (where each $p_i \in L_\alpha$) such that

$$x = \{y \in L_\alpha : \models_{L_\alpha} \psi(\mathring{y}, \vec{\mathring{p}})\}.$$

  Note that this is close to the statement we want to obtain; we only need to deal with the parameters p_i and turn them into appropriate L_{γ_i} applied to a suitable $\mathscr{L}$-formula ϕ. Now, each p_i is of rank γ_i. Define $\gamma = \max(\gamma_1, \dots, \gamma_n)$, then $p_i \in L_\gamma$ for each $1 \leq i \leq n$. By assumption, $\gamma < \alpha$, and thus we can use the inductive hypothesis and express each p_i in the form

$$p_i = \{y \in L_\gamma : \models_{L_\gamma} \psi_i(\mathring{y}, \mathring{L}_{\gamma_{i,1}}, \dots, \mathring{L}_{\gamma_{i,k(i)}})\}$$

  for suitable formulas ψ_i and ordinals $\gamma_{i,j}$ and some integer $k(i)$ depending on i. Our goal is to combine all these formulas into one formula ϕ using levels of the constructible hierarchy, L_γ, as parameters.

  In order to achieve this, we extend the formulas $\psi_i(v_0, \dots, v_n)$ into a formula of the form $\psi_i'(v_0, \dots, v_{n+1})$ by binding their unbounded parameters to a fixed parameter $v_{k(i)+1}$; given our definition of γ_i, we want this parameter to be $\mathring{L}_\gamma$. Then

$$p_i = \{y \in L_\alpha : \models_{L_\alpha} (\mathring{y} \in \mathring{L}_\gamma \wedge \psi_i'(\mathring{y}, \mathring{L}_{\gamma_{i,1}}, \dots, \mathring{L}_{\gamma_{i,k(i)}}, \mathring{L}_\gamma))\}.$$

  But now we may combine all these ψ_i'-formulas in order to define x as

$$\begin{aligned} x = &\{y \in L_\alpha : \models_{L_\alpha} \exists \vec{p}^n \ (\psi(\mathring{y}, \vec{\mathring{p}}) \\ &\wedge \forall i < n \ \forall v \ (v \in p_{i+1} \leftrightarrow (v \in \mathring{L}_\gamma \wedge \psi_{i+1}'(\mathring{y}, \mathring{L}_{\gamma_{i+1,1}}, \dots, \mathring{L}_{\gamma_{i+1,k(i+1)}}, \mathring{L}_\gamma))) \end{aligned}$$

which we can rewrite as

$$x = \{y \in L_\alpha : \models_{L_\alpha} \phi(\mathring{y}, \mathring{L}_\gamma, \mathring{L}_{\gamma_{1,1}}, \dots, \mathring{L}_{\gamma_{1,k(1)}}, \dots, \mathring{L}_{\gamma_{n,k(n)}})\}$$

as required.

Hence the result is proven by induction. □

We can now state the ordering required to well-order L.

Definition 5.81. Let x and y be elements of L and define an ordering $<_L$ on L by

$$x <_L y$$

if and only if

A

$$\mathrm{rank}_L(x) < \mathrm{rank}_L(y)$$

OR

B

$$x, y \in L_{\alpha+1} \setminus L_\alpha$$

for some $\alpha \in \mathbf{ON}$

AND

B.1

$$x = \left\{z \in L_\alpha : \models_{L_\alpha} \phi(\mathring{z}, \mathring{L}_{\gamma_1}, \dots, \mathring{L}_{\gamma_n})\right\}$$

and

$$y = \left\{z \in L_\alpha : \models_{L_\alpha} \psi(\mathring{z}, \mathring{L}_{\delta_1}, \dots, \mathring{L}_{\delta_m})\right\}$$

and

$$\phi <_\dagger \psi$$

where ϕ and ψ are the least $\mathscr{L}$-formulas with respect to the ordering $<_\dagger$ for which ordinals $\vec{\gamma} < \alpha$ and $\vec{\delta} < \alpha$ as required above exist

OR

B.2

the formulas ϕ and ψ from **B.1** coincide and

$$(\gamma_1, \dots, \gamma_n) <_* (\delta_1, \dots, \delta_n)$$

where $(\gamma_1, \dots, \gamma_n)$ and $(\delta, \dots, \delta_n)$ are the least sequences with respect to $<_*$ that define x and y as in **B.1**.

Of course, this well-ordering only works in L as we may identify every set in L by an $\mathscr{L}$-formula. This follows directly as every set in L is constructible.

It is clear now that the proof of theorem 5.76 will be of a very direct nature: by exhibiting a global well-ordering of L we verify that each element of L may also be well-ordered, which yields the result.

We now explain the ordering in detail. Let $x, y \in L$.

- In the first step of the comparison, we compare the rank of the elements. This is a very natural way of comparing elements in any transitive hierarchy. In particular, it gives details about the complexity of the element in terms of the parameters necessary; recall that, by definition, an element $x \in L_{\alpha+1}$ can be expressed using parameters from L_α only. Thus, the higher in the hierarchy we find x, the more parameters from higher up in the hierarchy are required.
- If this comparison is inconclusive, then we use the previous lemma in order to express x and y in normal form, i.e. in terms of a formula with parameters of levels of the hierarchy strictly below $\text{rank}_L(x)$ and $\text{rank}_L(y)$ (which are equal by assumption). Informally, this tells us how "complex" the formulas defining x and y (as opposed to x and y themselves) are; this is valid as we can associate every element of L with a formula. By choosing the least such formula, we obtain the canonical form which provides us with the additional information of uniqueness.
- Finally, if this comparison does not yield a conclusion either, then the formulas defining x and y as described above must coincide. We now consider the ordinal sequences that feature as levels of parameters of the canonical formulas defining x and y. As the formulas coincide, so does the number of parameters in both formulas. Hence, in order to make a distinction, we verify which parameters exactly are required in order to define x and y as described above. In case there are multiple such ordinal sequences for x or y, we use the $<_*$-least such ordinal sequence. Comparing these $<_*$-least sequences of x and y yields the required ordering of x and y.

  At this point it also clear that if these ordinal sequences coincide, too, then x and y must be equal as they can be expressed by the exact same formula.

Proposition 5.82. *The ordering $<_L$ well-orders L.*

Proof. Firstly, this uses the well-ordering of ordinals when the rank of two elements of L is considered. The result then follows directly from the proof that $<_\dagger$ is a well-ordering. Further, note that $<_*$ is also a well-ordering as, generalising the proof of lemma 5.79, we would otherwise obtain an infinite strictly decreasing sequence of ordinals. □

Of course, we are now required to find an equivalent $\mathscr{L}$-formula that is absolute with respect to $\text{ZF} + V = L$. We will do this in the same way as we have before: we find *LST*-formulas that translate our informally defined ordering above into a metamathematical statement. We then show absoluteness of these formulas with respect to $\text{ZF} + V = L$. This will ensure that the well-ordering will be the same in every model of $\text{ZF} + V = L$.

Finally, we will give a function F that well-orders L and prove

$$\text{ZF} \vdash \text{“}F\colon\ \mathbf{ON} \to L \text{ is a bijection”}.$$

Considering the graph of this function L will then prove the result.

We begin by translating the well-ordering into *LST*. As writing out the respective formulas is not particularly illuminating to the reader we state the following theorem and give an outline of the structure of the formulas.

Theorem 5.83. *There exists an LST-formula* $\mathrm{NF}(\alpha, x, \phi, t)$ *such that* $\mathrm{NF}(\alpha, x, \phi, t)$ *holds if and only if* ϕ *is an* $\mathscr{L}$*-formula,* t *is a finite sequence of ordinals less than* α*, the length of* ϕ *equals the domain of* t *and, crucially,* $x = \{y \in L_\alpha : \models_{L_\alpha} \phi(\mathring{y}, \mathring{L}_{t(0)}, \ldots, \mathring{L}_{t(n-1)}\}$.

Here, NF is short for *normal form*; we have mimicked the construction of the normal form of elements of L as proven in lemma 5.80.

Proof. Omitted. Details can be found in [Dev17, p. 73]. □

Eventually, we want to find an *LST*-formula which we shall call $\mathrm{WO}(x, y)$, which is true if and only if $x <_L y$. We now describe the construction of this *LST*-formula $\mathrm{WO}(x, y)$ in detail.

As used in the well-ordering $<_L$ in part **B.1**, we need to express the sentence "ϕ is the least $\mathscr{L}$-formula with respect to $<_\dagger$ such that $\mathrm{NF}(\alpha, x, \phi, t)$". This relates to our mentioning of the *canonical form* we aim to describe now.

Definition 5.84. Let x be a set, $\alpha \in \mathbf{ON}$ and let ϕ be an $\mathscr{L}$-formula. Then we denote by $\mathrm{CF}(\alpha, x, \phi)$ the *LST*-formula

$$\exists t\ (\mathrm{NF}(\alpha, x, \phi, t)) \wedge \forall \phi'\ (\exists t'\ (\mathrm{NF}(\alpha, x, \phi', t')) \rightarrow (\phi <_\dagger \phi' \vee \phi = \phi')).$$

Here, CF is short for *canonical form*. Indeed, it is clear that $\mathrm{CF}(\alpha, x, \phi)$ holds if and only if ϕ is the $<_\dagger$-least formula for which x can be represented in normal form.

Following from **B.2** above, we need to code the sentence "t is the least sequence of ordinals less than α with respect to $<_*$ such that $\mathrm{NF}(\alpha, x, \phi, t)$ holds". This will be our next step:

Definition 5.85. Let x be a set, $\alpha \in \mathbf{ON}$, let t be a sequence of ordinals less than α, and let ϕ be an $\mathscr{L}$-formula. Then we denote by $\mathrm{LOS}(\alpha, x, \phi)$ the *LST*-formula

$$\mathrm{NF}(\alpha, x, \phi, t) \wedge \forall t'\ (\mathrm{NF}(\alpha, x, \phi, t') \rightarrow t \leq_* t').$$

The shorthand LOS stands for *least ordinal sequence.*

In our definition of the well-ordering $<_L$, we assumed in step **B.2** that ϕ and ψ defining x and y respectively coincide and are the canonical such formulas for both elements. As we shall perform this check of canonical formulas in **B.1** (and hence strictly before we consider the ordering $<_*$ as outlined in the definition above), there is no need to include the formula for the canonical form CF above; when definition 5.85 is used within $\mathrm{WO}(x, y)$, the formula ϕ will be of canonical form already by assumption in **B.1**.

Remark. *The observant reader will have noticed that we have not expressed the orderings* $<_*$ *and* $<_\dagger$ *in LST. We have refrained from doing so in order to preserve an appropriate level of readability. As an addition to Devlin's presentation, we give the LST-formulas for both orderings below:*

- *The ordering* $<_*$ *can be described as follows: consider the sequences of ordinals* $\gamma = (\gamma_0, \ldots, \gamma_{m-1})$ *and* $\delta = (\delta_0, \ldots, \delta_{n-1})$*. Then*

$$\begin{aligned}\gamma <_* \delta \leftrightarrow \exists s\ \exists t\ (&\mathrm{dom}(\gamma) = s \wedge \mathrm{dom}(\delta) = t \\ &\wedge\ (s < t \\ &\quad \vee\ (s = t \wedge \exists i < t\ \forall j < t\ ((j < i \rightarrow \gamma(j) = \delta(j)) \wedge \gamma(i) < \delta(i))))).\end{aligned}$$

- *For the ordering $<_\dagger$, consider $\mathscr{L}$-formulas ϕ and ψ. Then*

$$\begin{aligned}\phi <_\dagger \psi \leftrightarrow \exists s\, \exists t\, (&\mathrm{dom}(\phi) = s \wedge \mathrm{dom}(\psi) = t \wedge s \leq t \\ &\wedge ((\forall i < s\, (\phi(i) = \psi(i))) \\ &\quad \vee \\ &\quad (s = t \wedge \exists i < t\, \forall j < t\, ((j < i \rightarrow \phi(j) = \psi(j)) \wedge k(\phi(i)) < k(\psi(i))))))\end{aligned}$$

where

$$\begin{aligned}k(a) < k(b) \leftrightarrow\ &\exists n\, \exists m\, (a = n \wedge b = m \wedge n < m) \\ &\vee (\exists n\, \exists m\, \exists r\, (a = (m, n) \wedge b = r \wedge n + 9 < r)) \\ &\vee (\exists n\, \exists m\, \exists r\, (a = r \wedge b = (m, n) \wedge r < n + 9)) \\ &\vee (\exists n\, \exists m\, \exists r\, \exists s\, (a = (m, n) \wedge b = (r, s) \wedge n + 9 < s + 9)).\end{aligned}$$

We now have all the ingredients we require in order to express the ordering $<_L$ in *LST*.

Theorem 5.86. *Let $x, y \in L$. there exists an LST-formula* $\mathrm{LL}(x, y)$ *such that* $\mathrm{LL}(x, y)$ *holds if and only if $x <_L y$.*

The proof below is based on the same patterns we have used many times before. It follows the original source [Dev17, p. 74].

Proof. We compose the formula $\mathrm{LL}(x, y)$ in the natural way and copy the build-up given in the definition of $<_L$ in definition 5.81: let $\mathrm{LL}(x, y)$ denote the formula

$$\begin{aligned}&\exists \alpha\, (x \in L_\alpha \wedge y \notin L_\alpha) && (\mathbf{A}) \\ &\vee \\ &\exists \alpha\, ((x \in L_{\alpha+1} \setminus L_\alpha \wedge y \in L_{\alpha+1} \setminus L_\alpha) && (\mathbf{B}) \\ &\quad \wedge \\ &\quad (\exists \phi\, \exists \psi\, (\mathrm{CF}(\alpha, x, \phi) \wedge \mathrm{CF}(\alpha, y, \psi) \wedge \phi <_\dagger \psi) && (\mathbf{B.1}) \\ &\qquad \vee \\ &\qquad \exists \phi\, (\mathrm{CF}(\alpha, x, \phi) \wedge \mathrm{CF}(\alpha, y, \phi)) \\ &\qquad \wedge\ \exists s\, \exists t\, (\mathrm{LOS}(\alpha, x, \phi, s) \wedge \mathrm{LOS}(\alpha, x, \phi, t) \wedge s <_* t))) && (\mathbf{B.2})\end{aligned}$$

It easily seen that this formula is as required. □

Note that in the proof above we used the shorthand $x \in L_{\alpha+1} \setminus L_\alpha$ for the longer $x \in L_{\alpha+1} \wedge x \notin L_{\alpha+1}$.

We are now required to consider the quantifier complexity of the formula given above in order to guarantee absoluteness and preserve definability between models. Notice that we have multiple unbounded existential quantifiers in the terms of $\mathrm{LL}(x, y)$ that consider cases **B**, **B.1** and **B.2**. However, we can bind all of these quantifiers to $L_{\max(\omega, \alpha+4)}$. This can be deduced by writing out the formulas for NF, CF, etc. in detail and noting where we subtly introduced ordered pairs and functions (and hence climbed up the constructible hierarchy in order to be able to define these sets as needed), for example. (The reasoning here is very similar to how we proved proposition 5.70 part (iii), for instance.)

Hence we obtain the formula $\mathrm{ll}(x, y, w)$ given by

$$\exists \alpha\ ((x \in L_{\alpha+1} \setminus L_\alpha \wedge y \in L_{\alpha+1} \setminus L_\alpha) \qquad \textbf{(B)}$$
$$\wedge$$
$$(\exists \phi \in w\ \exists \psi \in w\ (\mathrm{CF}(\alpha, x, \phi) \wedge \mathrm{CF}(\alpha, y, \psi) \wedge \phi <_\dagger \psi) \qquad \textbf{(B.1)}$$
$$\vee$$
$$\exists \phi \in w\ (\mathrm{CF}(\alpha, x, \phi) \wedge \mathrm{CF}(\alpha, y, \phi))$$
$$\wedge\ \exists s \in w\ \exists t \in w\ (\mathrm{LOS}(\alpha, x, \phi, s) \wedge \mathrm{LOS}(\alpha, x, \phi, t) \wedge s <_* t))). \qquad \textbf{(B.2)}$$

Assume $\mathrm{ll}'(x, y, \alpha, w)$ denotes the *LST*-formula $\mathrm{ll}(x, y, w)$ defined above *excluding* the leading existential quantifier $\exists \alpha$ (and hence α is free in ll'). Using this, we can finally define the required *LST*-formula.

Definition 5.87. Let $x, y \in L$. Then we define $\mathrm{WO}(x, y)$ to be the formula

$$\exists \alpha\ (x \in L_\alpha \wedge y \notin L_\alpha) \vee \exists \alpha\ \exists w\ (w = L_{\max(\omega, \alpha+4)} \wedge \mathrm{ll}'(x, y, \alpha, w)).$$

In order to prove that L can be well-ordered, we require one more crucial equivalence.

Lemma 5.88. *Let $x, y \in L_\gamma$ for some $\gamma \in$ **ON**. The LST-formula $\mathrm{WO}(x, y)$ is $\Delta_1^{\mathrm{ZF}\,+\,V=L}$. Further, denote by $\mathrm{wo}(x, y)$ the $\mathscr{L}$-counterpart of $\mathrm{WO}(x, y)$. Then the equivalence*

$$\mathrm{WO}(x, y) \leftrightarrow \models_{L_\gamma} \mathrm{wo}(\mathring{x}, \mathring{y})$$

holds with $\gamma = \max(\omega, \alpha + 5)$.

Of course, this equivalence is crucial in defining the required well-order on L. Its proof is based on ideas by Devlin (see [Dev17, pp. 66-7]).

Remark. *We recall the crucial reasoning behind the following proof: in LST, we can preserve the meaning of formulas between transitive models of our theory* ZF *by ensuring the formulas are absolute. This does not suffice for $\mathscr{L}$: the meaning of formulas (again, we consider the class of all n-tuples that satisfy that particular formula) between amenable classes, i.e. those that "model"* BS*, is invariant only if it is definable in all models. Hence, once we have shown that a particular formula is uniformly $\Delta_1^{L_\alpha}$ for limit ordinals $\alpha > \omega$, then we may deduce that the formula is invariant when passing from one amenable class to another.*

Proof. The proof is quite technical and hence omitted. It is based on the ideas presented in section 5.6.1. In short, it reduces to an exercise in finding where the required set lives within the constructible hierarchy. Details can be found in [Dev17, pp. 74-5]. □

The reasoning in section 5.6.1 also allows us to deduce the following corollary (Devlin does so in [Dev17, pp. 74-5]):

Corollary 5.89. *The formula $\mathrm{wo}(x, y)$ is uniformly $\Delta_1^{L_\alpha}$ for all limit ordinals $\alpha > \omega$.*

Proof. The result follows from the fact that if $x, y \in L_\gamma$ then, as we have proven above, $\mathrm{WO}(x, y) \leftrightarrow \models_{L_\gamma} \mathrm{wo}(\mathring{x}, \mathring{y})$. Then the argument in section 5.6.1 yields the result. □

From now, we are rather more flexible with our notation. Due to the lemma and corollary above, we may freely alternate between $\mathrm{WO}(x, y)$, $\mathrm{wo}(x, y)$, and the much more intuitive $x <_L y$.

Definition 5.90. Let $x \in L$. We define the predecessor function $\mathrm{pr}(x)$ by

$$\mathrm{pr}(x) = \{y : y <_L x\}.$$

This function will be of vital importance when we construct F, the function that will well-order L.

Lemma 5.91. *The predecessor function satisfies the following:*

(i) if α is an uncountable limit ordinal, then $x \in L_\alpha$ implies that $\mathrm{pr}(x) \in L_\alpha$*;*

(ii) if $x \in L$ then $\mathrm{pr}(x) \in L$*; and*

(iii) the class pr *is uniformly $\Delta_1^{L_\alpha}$ for limit ordinals $\alpha > \omega$.*

Proof. The proofs are based on Devlin's approaches, which can be found in [Dev17, pp. 75-6] for reference.

(i) Fix any $x \in L$. We need to show that $\mathrm{pr}(x) \in L$. In order to take advantage of the fact that every element of L can be expressed by a formula, once we have determined that formula, we are done.

Choose an ordinal $\beta < \alpha$ so that $x \in L_\beta$. Now, if $y <_L x$ then $y \in L_\beta$ by the fact that $<_L$ orders L and by transitivity of L_β. Furthermore, using the equivalence of $\mathrm{WO}(x, y)$ and $\mathrm{wo}(\mathring{x}, \mathring{y})$ for all $x, y \in L_\beta$, we may deduce that

$$\mathrm{pr}(x) = \{y : \models_{L_\gamma} \mathrm{wo}(\mathring{y}, \mathring{x})\}$$

for $\gamma = \max(\omega, \beta + 5)$ (as shown in lemma 5.88). As before, note that $\mathrm{pr}(x) \in L_{\gamma+1}$, and hence $\mathrm{pr}(x) \in L$, as required.

(ii) This result follows immediately from (i).

(iii) Omitted. The proof is similar to previous definability proofs: we exhibit a suitable formula and show that it is $\Sigma_1^{L_\alpha}$. Using a result given in Devlin's original presentation (see [Dev17, p. 47, 10.4 Corollary]), we can interpret pr as a function from L_α to L_α; then the result follows. (Again, this result is necessary in order to prove that definability is invariant between amenable sets.)

Thus the proof is complete. □

We now give the following important lemma:

Lemma 5.92. *There is a Σ_1 formula* $\mathrm{Enum}(\alpha, x)$ *in LST such that*

$$\mathrm{ZF} \vdash \text{“If } F = \{(\alpha, x) | \mathrm{Enum}(x, \alpha)\} \text{ then } F \text{ is a bijection between } \mathbf{ON} \text{ and } L\text{”.}$$

See [Dev17, p. 76] for the original proof, we give it in full detail below:

Proof. In order to make sense of this statement, we need to agree on what particular meaning we assign to the formula Enum. As we want to use the well-ordering $<_L$, it makes sense to enumerate (hence the nomenclature) the elements of L. This will provide us with the required well-ordering. Indeed, if we now consider the formula Enum and range over all $\alpha \in \mathbf{ON}$, then $\mathrm{Enum}(x, \alpha)$ holds if and only if x is the element at index α under the well-ordering $<_L$.

We prove the result by writing down a suitable formula: consider the formula

$$\begin{aligned}\exists f\ (&f \text{ is a function} \wedge \operatorname{dom}(f) = \alpha + 1 \\ &\wedge\ \forall \beta \in \alpha + 1\, \forall \gamma \in \alpha + 1\ (\beta < \gamma \to f(\beta) <_L f(\gamma)) \\ &\wedge\ \exists z\ (z = \operatorname{pr}(x) \wedge \forall y \in z\, \exists \beta \in \alpha\ (y = f(\beta)) \wedge f(\alpha) = x)).\end{aligned}$$

This formula clearly yields exactly what we need: it holds if and only if there is a function f that preserves ordering with respect to $<_L$ and it orders all elements $<_L$-below α in its correct place according to $<_L$. Observe that f orders the entire class L.

Further, note that this formula is absolute for L, as the function f whose existence is postulated by the LST-formula above only exists in L. Note that the relativised formula only ranges over the elements of L, and since $(L)^L = L$, absoluteness follows. □

All in all, note that as ZF proves the existence of such a function above, and by the absoluteness shown, we see that whenever M is an inner model of ZF, then such a function f as described above exists for $(L)^M = L$.

The main result, theorem 5.76, now follows easily. Devlin provides the following line of reasoning in [Dev17, p. 76]; we extend the proof slightly.

Proof of theorem 5.76: We need to find a function F that well-orders L. But the previous lemma provides us with exactly this: if we interpret the set F as the graph of a function, then F is a bijection from the ordinals to L that preserves the ordering $<_L$ which we proved above to be a well-ordering. □

Remark. *The heart of the proof that* ZF *proves* $(\mathrm{AC})^L$ *is the definition of the well-ordering* $<_L$. *The fact that we may associate every constructible set with a formula renders the result possible. As seen above, we can use this characteristic of the constructible universe to establish our well-ordering, which, in turn, allows us to find a function manifesting the well-ordering. The fact that we are required to denote sets in terms of formulas is exactly the reason why we cannot adapt this technique to the case of* $V \neq L$.

Further, what we have shown above is actually stronger than AC. *The fact that we may well-order the whole class L (note that L is a proper class) and not just its elements (which are sets, of course) is called "Axiom of Global Choice". It is obvious that* AGC *implies* AC.

This is a remarkable result. By restricting the sets permitted to belong to our class L, we have managed to augment L with additional properties that are independent of ZF.

We may now use this fact in order to deduce that AC is relatively consistent with ZF. As we cannot show proper consistency (this would require to prove that ZF is consistent first, which is impossible by Gödel's second incompleteness theorem), we consider a special case called *relative consistency*: assuming ZF is consistent, what can we say about $\mathrm{ZF} + \Phi$ for some LST-sentence Φ? As it turns out, the previous result of finding a proper class that is an inner model of $\mathrm{ZF} + \Phi$ allows us to deduce relative consistency.

The following theorem will provide this crucial link. The proof is based on explanations by Devlin in [Dev17, p. 77].

Theorem 5.93. *Let* Φ *be an LST-sentence and suppose that* M *is an inner model of* $\mathrm{ZF} + \Phi$. *If* ZF *is consistent, then so is* $\mathrm{ZF} + \Phi$.

Proof. We argue by contradiction. Suppose that ZF + Φ is inconsistent. Hence

$$\mathrm{ZF} + \Phi \vdash \Psi$$

where Ψ is a contradiction. By definition, we can find a formal proof (i.e. a sequence of theorems of ZF + Φ) which we denote by $\Psi_1, \ldots, \Psi_n, \Psi$. Clearly, each such Ψ_i is either an axiom of ZF + Φ or can be deduced from some of the Ψ_j with $j < i$ by applying rules of logic (i.e. carrying out a formal deduction). By assumption, ZF is consistent, and hence if Ψ_i is an axiom of ZF + Φ, then Ψ_i^M holds as it is a theorem of ZF by assumption on M. But then each sentence of the sequence Ψ_i leading up to the contradiction Ψ is also a theorem of ZF once it is relativised to M (the logical deductions that construct Ψ_i can equally be carried out in M as M is an inner model of ZF + Φ). Thus, in particular, Ψ^M is a theorem of ZF. Using the fact that M is an inner model of ZF + Ψ, we see that Ψ^M is, of course, also an inconsistency. Hence ZF proves a contradiction and is hence inconsistent. Contradiction. □

Note we did not appeal to the Axiom of Choice in the proof of theorem 5.76. Thus we may deduce the following crucial result as a corollary.

Corollary 5.94. *If* ZF *is consistent, then so is* ZFC.

In order to conclude this section, we state the following corollary.

Corollary 5.95. *If* ZF *is consistent, then so is* ZF + $V = L$.

Finally, we compose the corollaries above and give Devlin's result from [Dev17, p. 77].

Corollary 5.96. *If* ZF *is consistent, then so is* ZFC + $V = L$.

Proof. This follows immediately from the fact that L is an inner model of ZFC and of ZF + $V = L$ as shown above. □

The previous corollary in combination with theorem 5.93 tell us that whenever we have a sentence Φ in LST for which ZFC + $V = L \vdash \Phi$, then we automatically see that Φ is consistent with ZFC.

5.8 What next?

As we have shown above, the constructible universe is special. Not only does it satisfy the Axiom of Choice, using the so-called *condensation lemma*, we may also show that GCH holds in L as well. This is one step along the way to Paul Cohen's famous result proving the independence of CH in ZFC, as presented in his original papers [Coh63] and [Coh64]. Cohen proved the independence by exhibiting a model in which CH fails. It is clear that then neither CH nor $\neg$ CH can be deduced from ZF.

As it turns out, the constructible universe solves further problems of combinatorial set theory. The following short digression will allow us to tie this section to the previous sections on combinatorial set theory.

Definition 5.97. The $\Diamond$-principle states that there are sets $A_\alpha \subset \alpha$ for $\alpha < \omega_1$ for which

$$\forall A \subset \omega_1 \ (\{\alpha < \omega_1 : A \cap \alpha = A_\alpha\} \text{ is } \omega_1\text{-stationary}).$$

Such a sequence $(A_\alpha)_{\alpha<\omega_1}$ is called a $\Diamond$-sequence.

Further, one can easily show that the $\Diamond$-principle implies the Continuum Hypothesis.

Theorem 5.98. *If there is a $\Diamond$-sequence, then* $2^{\aleph_0} = \aleph_1$.

The proof is very straightforward. We extend Kunen's proof from [Kun80, p. 80] by additional explanation below.

Proof. Assume there is a $\Diamond$-sequence $(A_\alpha)_{\alpha<\omega_1}$. Suppose A is a subset of ω. We show that $A = A_\alpha$ for some $\omega < \alpha < \omega_1$. By assumption, the set

$$S_A := \{\alpha < \omega_1 : A \cap \alpha = A_\alpha\}$$

is ω_1-stationary. As ω_1 is regular with cardinality $\aleph_1$, and as S_A is unbounded by proposition 4.25, we must have $|S_A| = \aleph_1$. Hence, in particular, there exists $\alpha \in S_A$ which is greater than ω. Hence, by definition of S_A,

$$A \cap \alpha = A_\alpha.$$

But note that A is a subset of ω and α is uncountable. Hence

$$A \cap \alpha = A$$

which yields

$$A = A_\alpha$$

as required.

Note that A was chosen arbitrarily. Thus we see that every $A \subset \omega$ is an element of the $\Diamond$-sequence $(A_\alpha)_{\alpha<\omega_1}$. Finally, note that $|\{A_\alpha : \alpha < \omega_1\}| = \aleph_1$. By Cantor's theorem, $|\mathcal{P}(\omega)| > \aleph_0$, and hence we may deduce that $|\mathcal{P}(\omega)| = |\{A_\alpha : \alpha < \omega_1\}| = \aleph_1$, which proves the result. $\square$

As Jensen showed in [Jen72, p. 292], the Axiom of Constructibility implies the $\Diamond$-principle (via showing that the Suslin hypothesis fails if $V = L$). Hence it follows directly that the $\Diamond$-principle is independent of ZF since so is the Axiom of Constructibility.

The constructible universe is of interest as it solves many of the combinatorial problems that cannot be solved in ZFC alone. However, at the same time, it is also quite restrictive. Assuming the axiom $V = L$ leaves us with far fewer sets to consider which, in some sense, trivialises many combinatorial problems as seen above.

All in all, the constructible universe is a strong tool to prove relative consistency results for ZF and ZFC and to improve our understanding of results in combinatorial set theory. However, restricting ourselves to $\mathrm{ZF} + V = L$ simplifies set theory to an extent in which many of the classical problems considered in the field are trivially solved (in [Jen72, p. 229], Jensen calls this restriction *micro set theory*). Hence $\mathrm{ZF} + V = L$ is not generally considered an alternative to ZF, but merely an interesting extension in its own right.

At the beginning of section 5.6 we remarked that the constructible universe steers clear of the peculiarities that are caused by the unrestricted power set operator. But one might want to argue that precisely these quirks and "anomalies" that give rise to problems such as GCH, for example, render set theory so fascinating. Hence, a strong assumption such as $V = L$ strips set theory of part of its attractiveness that is at its core: the ability to pose far-reaching questions.

6 Closing Words

This report was designed to give an overview of some topics in combinatorial set theory as well as provide an introduction to the constructible universe. Of course, we have only scratched the surface; set theory is an incredibly deep topic with connections to many different branches of mathematics. With applications ranging from topology (via descriptive set theory, for example) to algebra (the Whitehead problem springs to mind, for instance), set theory can provide (quite surprising) solutions to problems that appear unassociated at first glance.

We hope that reading the preceding sections has given the reader an idea of the unique and utterly fascinating nature of some of the countless facets of set theory.

I would like to take the opportunity to thank my supervisor Gareth Jones for his invaluable feedback, countless suggestions, and his willingness to help. Working on this project has been a great pleasure.

References

[BP77] J. E. Baumgartner and K. Prikry. Singular Cardinals and the Generalized Continuum Hypothesis. *The American Mathematical Monthly*, 84(2):108–113, 1977.

[CK62] C. C. Chang and H. Jerome Keisler. An Improved Prenex Normal Form. *The Journal of Symbolic Logic*, 27(3):317–326, 1962.

[Coh63] Paul J. Cohen. The Independence of the Continuum Hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 50(6):1143–1148, 12 1963.

[Coh64] Paul J. Cohen. The Independence of the Continuum Hypothesis, II. *Proceedings of the National Academy of Sciences of the United States of America*, 51(1):105–110, 01 1964.

[Dev93] Keith J. Devlin. *The Joy of Sets: Fundamentals of Contemporary Set Theory.* Undergraduate texts in mathematics. Springer-Verlag, New York ; London, 2nd edition, 1993.

[Dev17] Keith J. Devlin. *Constructibility.* Perspectives in Logic. Cambridge University Press, 2017.

[Dil50] R. P. Dilworth. A Decomposition Theorem for Partially Ordered Sets. *Annals of Mathematics*, 51(1):161–166, 1950.

[Eas70] William B. Easton. Powers of regular cardinals. *Annals of Mathematical Logic*, 1(2):139 – 178, 1970.

[FH08] Sy-David Friedman and Radek Honzik. Easton's theorem and large cardinals. *Annals of Pure and Applied Logic*, 154(3):191–208, 2008.

[Gal94] Fred Galvin. A proof of Dilworth's Chain Decomposition Theorem. *The American Mathematical Monthly*, 101(4):352–353, 1994.

[Göd38] Kurt Gödel. The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 24(12):556–557, 1938.

[HH99] András Hajnal and Peter Hamburger. *Set Theory.* London Mathematical Society Student Texts: 48. Cambridge University Press, 1st english edition, 1 1999.

[Jec95] Thomas Jech. Singular Cardinals and the pcf Theory. *The Bulletin of Symbolic Logic*, 1(4):408–424, 1995.

[Jen72] R. Björn Jensen. The fine structure of the constructible hierarchy. *Annals of Mathematical Logic*, 4(3):229–308, 1972.

[Kan03] Akihiro Kanamori. *The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings.* Springer-Verlag Berlin Heidelberg, Berlin, 1st edition, 2003.

[Kun80] Kenneth Kunen. *Set theory: An Introduction to Independence Proofs*, volume 102. North-Holland Pub. Co. Sole distributors for the U.S.A. and Canada, Elsevier North-Holland, Amsterdam, New York, 1st edition, 1980.

[Mag77] Menachem Magidor. On the Singular Cardinals Problem II. *Annals of Mathematics*, 106(3):517–547, 1977.

[Put63] Hilary Putnam. A note on constructible sets of integers. *Notre Dame J. Formal Logic*, 4(4):270–273, 1963.

[Res66] Michael David Resnik. On Skolem's Paradox. *The Journal of Philosophy*, 63(15):425–438, 1966.

[Sil75] Jack Silver. On the singular cardinals problem. *Proceedings of the International Congress of Mathematicians, Vancouver 1974*, 1(1):265–268, 1975.

[WM14] Mitsuko Wate-Mizuno. Mathematical recreations of Dénes Kőnig and his work on graph theory. *Historia Mathematica*, 41(4):377–399, November 2014.