Axiomatic Set Theory

P.D. Welch

September 24, 2023

Contents

Page

Chapter 1

Axioms and Formal Systems

1.1 Introduction

The great German mathematician David Hilbert (1862-1943) in his address to the second International Congress of Mathematicians in Paris 1900 placed before the audience a list of the 23 mathematical problems he considered the most relevant, the most urgent, for the new century to solve. Hilbert had been a defender of Cantor's seminal work on on infinite sets, and listed the *Continuum Hypothesis* as one of the great unsolved questions of the day. He accordingly placed this question at the head of his list. The hypothesis is easy to state, and understandable to anyone with the most modest of mathematical education: *if A is an infinite subset of the real line continuum* $\mathbb{R}$*, then there is a bijection of A either with* $\mathbb{N}$ *the set of natural numbers, or with all of* $\mathbb{R}$. Phrased in the terminology that Cantor introduced following his discovery of the uncountability of the reals and his subsequent work on *cardinality* the hypothesis becomes: *for any such A, if A is not* countable *then it has the cardinality of* $\mathbb{R}$ *itself.* CH thus asserts that there is no cardinality that is intermediate between that of $\mathbb{N}$ and that of $\mathbb{R}$. If the cardinality of $\mathbb{N}$ is designated ω_0(or $\aleph_0$) and that of the first uncountable cardinal as ω_1(or $\aleph_1$) then CH is often written as "$2^{\omega_0} = \omega_1$" (or $2^{\aleph_0} = \aleph_1$) the point being here is that the real continuum can be identified with the class of infinite binary sequences ${}^{\mathbb{N}}2$ and the latter's cardinality is 2^{ω_0}.

Sometimes called the *Continuum Problem*, Cantor wrestled with this question for the rest of his career, without finding a solution. However, in this quest he also founded the subject of *Descriptive Set Theory* that seeks to prove results about sets of reals, or functions thereon, according to the complexity of their *description*. Such hierarchical bodies of sets were to become very influential in the Russian school of analysts (Suslin, Lusin, Novikoff) and the French (Lebesgue, Borel, Baire). The notion of a hierarchy built up by considering complexity of definition of course also invites methods of mathematical logic. Descriptive Set Theory has figured greatly in modern set theory, and there is a substantial body of results on the *definable continuum* where one tries to establish CH type results not for the whole continuum but just for "definable parts" thereof. Cantor was able to show that *closed* subsets of $\mathbb{R}$ satisfied CH: they were either countable or could be seen to contain a subset which was of cardinality the continuum. This allows one to say that then countable unions of closed sets also satisfied the CH. Cantor hoped to be able to prove CH for increasingly complicated sets of real numbers, and somehow exhaust all subsets in this way. The analysts listed above made great strides in this new field and were able to show that any *analytic* set satisfied CH. (At the same time they were producing results indicating that such sets were very "regular": they were all Lebesgue measurable, had a categorical property defined by Baire and many other such properties. Borel in particular defined a hierarchy of sets now named after him, which gave very real substance to Cantor's efforts to build up a hiearchy of increasing complexity from simple sets.)

However it was clear that although the study of such sets was rewarded with a regular picture of their properties, this was far from proving anything about all sets. We now know that Cantor was trying to prove the impossible: the mathematical tools available to him at his day would later be seen to be formalisable in Zermelo-Fraenkel set theory with the Axiom of Choice , AC (an axiom system abbreviated as ZFC). Within this theory it was shown (by (COHEN (1934-2007)) that CH is strictly unprovable. If he had taken the opposite tack and had thought the CH false, and had attempted to produce a set Aneither of cardinality that of $\mathbb{N}$ nor of $\mathbb{R}$ he would have been equally stuck: by a result of Gödel within ZFC it turns out that $\neg$ CH is strictly unprovable.

It is the aim of this course to give a proof of this latter result of Gödel. The method he used was to look at the cumulative hierarchy V (which we may take to be the universe of sets of mathematical discourse) in which all the ZF axioms are seen to be true, and to carve out a special transitive subclass - the class of *constructible sets*, abbreviated by the letter L. This L was a proper transitive class of sets (it contains all ordinals) and it was shown by Gödel (i) That any axiom σ from ZF was seen to hold in L; moreover (ii) Both AC and CH held in L. This establishes the unprovability of $\neg$ CH from the ZF axioms: L is a structure in which any axioms of ZF used in a purported proof of $\neg$ CH were true, and in which CH was true. However a proof of $\neg$ CH from that axiom set would contradict the fact that rules of first order logic are sound, that is truth preserving.

In modern terms we should say that Gödel constructed the first *inner model* of set theory: that is, a transitive class W containing all ordinals, and in which each axiom of ZFC can be shown to be true. Such models generalising Gödel's construction are much studied by contemporary set theorists, so we are in fact as interested in the construction as much as (or even more so now) than the actual result.

It is a perhaps a curious fact that such inner models invariably validate the CH but most set theorists do not see that fact alone as giving much evidence for a solution to the problem: the inner model L and those generalising it are built very carefully with much attention to detail as to how sets appear in their construction. Set theorists on the whole tend to feel that there is no reason that these procedures exhaust all the sets of mathematical discourse: we are building a very smooth, detailed object, but why should that imply that V is L? Or indeed any other of the later generations of models generalising it?

However it is one of facts we shall have to show about L that in one sense it is "self-constructing": the construction of L is a mathematical one; it therefore is done within the axiom system of ZF; but (we shall assert) L itself satisfies all such axioms; *ergo* we may run the construction of the constructible hierarchy *within* the model L itself (after all it is a universe satisfying all ZF axioms). It will be seen that this process activated in L picks up all of L itself: in short, the statement "$V = L$" is valid in L. The conclusion to be drawn from this is that from the axioms of ZF we cannot prove that there are sets that lie outside L. It is thus consistent with ZF that $V = L$ is true! If $V = L$ is true, then there are many consequences for mathematics: the study of L is now highly developed and many consequences for analysis, algebra,... have been shown to hold in L whose proof either remains elusive, or else is downright unprovable without assuming some additional axioms. It is a corollary to the consistency of $V = L$ with ZF, that we cannot use this method of constructing an inner model to find one in which $\neg$ CH holds: if it is consistent that

$V = L$ then it is consistent that L is the only inner model there is, so no construction using the axioms of ZF alone can possibly produce an inner model of $\neg$ CH.

We are thus left still in the state of ignorance that Hilbert protested was not the lot of mathematicians as regards the CH.[1] Cohen's proof that CH is not provable from the ZFC axioms does not proceed by using inner models (we have mentioned reasons why it cannot) but by constructing models of the axioms in a boolean valued logic: statements there do not have straightforward true/false truth values. In Cohen's models, when constructed aright, all axioms of ZF (and sentences provable from them in first order logic) receive the topmost truth value "1", and contradictions $\neg\sigma \wedge \sigma$, receive the bottom value "0". Cohen constructed such a model in which $\neg$ CH received a "non-0" truth value in the Boolean algebra, value p say. Consequently CH is not provable from ZF, else the Boolean model would have to assign the non-zero p to the inconsistent statement CH $\wedge \neg$ CH and such is not possible in these models. This literally taken, says absolutely nothing about sets in the universe V since the model is a sub-universe of V with a non-classical interpretation. It speaks only about what can or cannot be proven in first order logic from the axioms of ZFC. [2]

There are many results in set theory, in particular in axiomatic systems that enhance ZFC with some "strong axiom of infinity" that indicate that the CH is actually false (that $2^{\aleph_0} = \aleph_2$ often occurs in such cases). At present this can only be taken as some kind of quasi-empirical evidence and so is a source of much discussion.

Prerequisites: Cohen's proof is beyond the scope of this course, but we shall do Gödel's construction of L in detail. This will involve extending the basic results on *ordinal and cardinal numbers* and their *arithmetic*; we shall have recourse to *schemes of ordinal and* $\in$*-recursion*. The reader is assumed familiar with a development of these topics, as well as with the notion and basic properties of *transitive sets*. Although Gödel gave a presentation of the constructible hierarchy using a functional hierarchy, with almost all logic eliminated, (mainly as a way of presenting his results to "straight" mathematicians) we shall be going the traditional route of defining a "Definability" operator using all the syntactic resources of a formal language $\mathcal{L}$ and the methods of modern logic. Formal derivability $T \vdash \sigma$ will always mean that σ is derivable from the axioms T in one, or any, system of classical first order calculus familiar to the reader.

Acknowledgements: these notes are heavily indebted to a number of sources: in particular to Ronald B. Jensen : *Modelle der Mengenlehre* (Springer Lecture Notes in Maths, vol 37,1967), and his subsequent lecture notes.

1.2 Preliminaries: axioms and formal systems.

We introduce the formal first order language $\mathcal{L}$, and see how we can use *class terms* expressed in it. We then give a formulation of the Zermelo-Fraenkel axioms themselves.

---

[1] Or any mathematical problem "*You can find [the solution to any mathematical problem] by pure reason, for in mathematics there is no* ignorabimus" Hilbert, Lecture delivered to the 2nd International Congress of Mathematicians, Paris 1900.

[2] This is only one way of interpreting Cohen's forcing technique. See Kunen [4] *Set Theory: An Introduction to Independence Proofs*

Hilbert in 1900

1.2.1 The formal language of ZF set theory; terms

ZF set theory is formulated in a formal first order language of predicate logic with axioms for equality. The components of that language $\mathcal{L} = \mathcal{L}_{\dot\in}$ are:

(i) set variables; $v_0, v_1, \ldots, v_n, \ldots$ (for $n \in \mathbb{N}$)

(ii) two binary predicate symbols: $\dot=, \dot\in$

(iii) logical connectives: $\vee, \neg$

(iv) brackets: $($, $)$

(v) an existential quantifier: $\exists$.

The *formulae* of $\mathcal{L}$ are defined inductively in a way familiar for any first order language. We assume the reader has seen this done for his or herself and do not repeat this here. We assume also that the notion of *free variable* ($\mathrm{FVbl}(\varphi)$) and *subformula* of a formula φ as inductively defined over the collection of all formulae is also familiar. We shall use the notation $\phi(y/x)$ for the formulae ϕ with the free variable occurrences of the variable x replaced by the variable y. A formula with no free variables is called a *closed* formula or a *sentence*. It is sometimes convenient to augment the language $\mathcal{L}$ with other predicate symbols $\vec{A} = A_0, A_1, \ldots$; if this is done we denote the appropriate language by $\mathcal{L}_{\vec{A}}$.

We use the binary predicate symbol $\in$ as a relation to be interpreted as membership: "$v_0 \in v_1$" will be interpreted as "v_0 is a member of v_1" *etc.* We often use other letters also to stand in for variables v_k: typically x, y, z, and recalling the convention from ST: α, β, for ordinals *etc. etc.* [3] . It is so convenient to adopt these conventions that we do so immediately even when we write out our basic axioms. Note that in our statement of the Extensionality Axiom Ax 1 we also abbreviate "$\neg\exists v_k \neg\psi$" as usual by "$\forall v_k \psi$".

Ax0 ***(Extensionality)***

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

[3]Formally speaking the symbols $x, y, \alpha, \beta, \ldots$ are not part of $\mathcal{L}$: they have the status of *metavariables* in the *metalanguage*; the latter is the language we use to *talk about* $\mathcal{L}$. The metalanguage consists of English with a liberal admixture of such metavariables and other symbols as and when we require them. Some of our metatheoretical arguments require some simple arithmetic, as when we prove something about formulae or terms by induction. These arguments can all be done with *primitive recursive* arithmetic.

This single axiom expresses the fact that identity of sets is based solely on membership questions about the two sets.

We have seen that collections, or "classes" based on unguarded specifications within the language $\mathcal{L}$ can lead to trouble; recall Russell's Paradox: $R =_{\text{df}} \{x \mid x \notin x\}$ was a class that could not be considered to be a set. Likewise $V =_{\text{df}} \{x \mid x = x\}$ is not a set. Such collections we called "proper classes". It might be thought that this mode of introducing collections or classes is fraught with potential danger, and although we successfully used these ideas in ST perhaps it would be safer to do without them? In fact such methods of specifying collections is so useful that instead of being wary of them, we shall embrace them full heartedly whilst keeping them at a safe distance from our formal language $\mathcal{L}$.

Definition 1.1 *(i) A* class term *is a symbol string of the form* $\{x \mid \phi\}$ *where* x *is one of the variables* v_k *and* ϕ *is a formula of our language.*

(ii) A term t *is either a variable or a class term.*

(iii) The free variables *of a term* t *are given by:*

$\text{FVbl}(t) =_{\text{df}} \text{FVbl}(\phi)\backslash\{x\}$ *if* $t = \{x \mid \phi\}$; $\text{FVbl}(x) = \{x\}$ *if* x *is a variable.*

We allow terms to be substituted for variables in atomic formulae $x = y$ and $x \in y$, and for free variables in general in formulae of $\mathcal{L}$. We thus may write $\phi(t/x)$ for the formula ϕ with instances of x replaced by t. Just as for substitutions of variables in ordinary formulae in first order predicate logic, we only allow substitutions of terms t into formulae ψ where free variables of t do not become unintentionally bound by quantifiers of ψ. Substitutions can always be effected after a suitable change of the bound variables of ψ. A term t with $FVbl(t) = \varnothing$ is called a *closed* term.

A term of the form $\{x \mid \phi\}$ is *not* part of our language $\mathcal{L}$: it is to be understood purely as an abbreviation. Likewise $\phi(t/x)$ is not part of our language if t is a class term. We understand these abbreviations as follows:

$$\begin{array}{rcl}
y \in \{x \mid \phi\} & \text{is} & \phi(y/x); \\
\{x \mid \phi\} = \{z \mid \psi\} & \text{is} & \forall y(\phi(y/x) \longleftrightarrow \psi(y/z)) \\
z = \{x \mid \phi\} & \text{is} & \forall y(y \in z \longleftrightarrow y \in \{x \mid \phi\}) \\
\{x \mid \phi\} \in \{z \mid \psi\} & \text{is} & \exists y(y = \{x \mid \phi\} \wedge \psi(y/z)) \\
\{x \mid \phi\} \in z & \text{is} & \exists y(y \in z \wedge y = \{x \mid \phi\})
\end{array}$$

Although class terms appear on both sides of the above, this in fact gives a precise recursive way of translating a "*generalised formula*" containing class terms into one that does not. Note that a simple consequence of the above is that for any x we have $x = \{y \mid y \in x\}$. Note in particular that the fourth line ensures that if we write "$s \in t$" for terms s, t then s must be a set.

We now name certain terms and define some operations on terms. Again these are metatheoretical operations: we are talking *about* our language $\mathcal{L}$, and talking about, or *manipulating terms*, is part of that meta-talk.

Definition 1.2 *(i)* $V =_{\text{df}} \{x \mid x = x\}$; $\varnothing =_{\text{df}} \{x \mid x \neq x\}$;

(ii) $s \subseteq t =_{\text{df}} \forall x(x \in s \longrightarrow x \in t)$

(iii) $s \cup t =_{\text{df}} \{x \mid x \in s \vee x \in t\}$; $s \cap t =_{\text{df}} \{x \mid x \in s \wedge x \in t\}$;

$\neg s =_{\text{df}} \{x \mid x \notin s\}$; $s\backslash t =_{\text{df}} \{x \mid x \in s \wedge x \notin t\}$

(iv) $\bigcup s =_{\mathrm{df}} \{x \mid \exists y(y \in s \wedge x \in y)\}$; $\bigcap s =_{\mathrm{df}} \{x \mid \forall y(y \in s \longrightarrow x \in y)\}$
(v) $\{t_1, \dots, t_n\} =_{\mathrm{df}} \{x \mid x = t_1 \vee x = t_2 \vee \cdots \vee x = t_n\}$
(vi) $\langle x, y\rangle =_{\mathrm{df}} \{\{x\}, \{x, y\}\}$ *(the ordered pair set)*
(vii) $\langle x_1, x_2, \dots, x_n\rangle =_{\mathrm{df}} \langle\langle x_1, \dots, x_{n-1}\rangle, x_n\rangle$ *(the ordered n-tuple)*
(viii) $x \times z =_{\mathrm{df}} \{\langle u, v\rangle \mid u \in x \wedge v \in z\}$ *(the Cartesian product of x, z)*
$t^2 =_{\mathrm{df}} t \times t$; $t^{n+1} = t^n \times t$;
(ix) $\mathcal{P}(x) =_{\mathrm{df}} \{y \mid y \subseteq x\}$ *(the "power class" of x.)*

At the moment the above objects just have the status of syntactic names of certain terms, but we are going to adopt axioms that will assert that the classes defined are in fact sets. Indeed we shall say "x is a set"$\Longleftrightarrow_{\mathrm{df}}$"$x \in V$". In (viii) we have introduced a useful syntactic device: instead of writing

$$x \times z = \{y \mid \exists u \exists v(u \in x \wedge v \in z \wedge y = \langle u, v\rangle)\}$$

we have placed the constructed term $\langle u, v\rangle$ to the left of the $|$. In general we introduce this abbreviation: we let $\{t \mid \varphi\} =_{\mathrm{df}} \{z \mid \exists \vec{u}(z = t \wedge \varphi)\}$ (whose notation is probably more easily understood through the example above, here $\vec{u}$ is a list of variables containing all those free in t and φ).

Ax1 ***(Empty Set Axiom)*** $\varnothing \in V$.
Ax2 ***(Pairing Axiom)*** $\{x, y\} \in V$.
Ax3 ***(Union Axiom)*** $\bigcup x \in V$.

LEMMA 1.3 $t \in V \longleftrightarrow \exists y(y = t)$.

PROOF: (Actually 1.3 is a *theorem scheme*: for each term t there is a lemma corresponding to the definition of the term t.) By our rules on translation 1.2, $t \in V \leftrightarrow \exists y(y = t \wedge (x = x)(y/x)) \leftrightarrow \exists y(y = t \wedge (y = y)) \leftrightarrow \exists y(y = t)$. Q.E.D.

LEMMA 1.4 ***Ax0-3*** *prove:* $x \cup y \in V$; $\{x_1, \dots, x_n\} \in V$.

PROOF: By Ax2 $\{x, y\} \in V$ and then by Ax3 $\bigcup\{x, y\} \in V$. And $\bigcup\{x, y\} = x \cup y$ (by Ax0). Repeated application of Ax0-3 shows $\{x_1, \dots, x_n\} \in V$ (Exercise). Q.E.D.

There now follow a sequence of definitions of basic notions which we have already seen in ST.

DEFINITION 1.5 *Let r be a term. (i) r is a* relation $\Longleftrightarrow_{\mathrm{df}} r \subseteq V \times V$
(ii) r is an n-ary relation $\Longleftrightarrow_{\mathrm{df}} r \subseteq V^n$.

We write in (i) xry or rxy instead of $\langle x, y\rangle \in r$ and in (ii) $rx_1 \cdots x_n$ instead of $\langle x_1, \dots, x_n\rangle \in r$.

DEFINITION 1.6 *If r, s are relations and u a term we set:*
(i) $\mathrm{dom}(r) =_{\mathrm{df}} \{x \mid \exists y(xry)\}$; $\mathrm{ran}(r) =_{\mathrm{df}} \{y \mid \exists x(xry)\}$; $\mathrm{field}(r) =_{\mathrm{df}} \mathrm{dom}(r) \cup \mathrm{ran}(r)$.
(ii) $r \restriction u =_{\mathrm{df}} \{\langle x, y\rangle \mid xry \wedge x \in u\}$.
(iii) $r``u =_{\mathrm{df}} \{y \mid \exists x(x \in u \wedge xry\}$.
(iv) $r^{-1} =_{\mathrm{df}} \{\langle y, x\rangle \mid xry\}$.
(v) $r \circ s =_{\mathrm{df}} \{\langle x, z\rangle \mid \exists y(xry \wedge ysz)\}$.

We may define the unicity quantifier:

Definition 1.7 $\exists! x\varphi \Longleftrightarrow_{\text{df}} \exists z(\{z\} = \{x \mid \varphi\})$.

We now define familiar functional concepts.

Definition 1.8 *Let f be a relation.*
(i) f is a function $(\text{Fun}(f)) \Longleftrightarrow_{\text{df}} \forall x, y, z(fxy \wedge fxz \longrightarrow y = z)$ *(we write f(x)=y).*
(ii) f is an n-ary function$\Longleftrightarrow_{\text{df}}$ *f is a function* $\wedge\, \text{dom}(f) \subseteq V^n$
(we write $f(x_1, \ldots, x_n) = y$ instead of $f(\langle x_1, \ldots, x_n\rangle) = y$).
(iii) $f : a \longrightarrow b \Longleftrightarrow_{\text{df}} \text{Fun}(f) \wedge \text{dom}(f) = a \wedge \text{ran}(f) \subseteq b$.
(iv) $f : a \longrightarrow_{(1-1)} b \Longleftrightarrow_{\text{df}} f : a \longrightarrow b \wedge \text{Fun}(f^{-1})$ *("f is an injection or (1-1)").*
(v) $f : a \longrightarrow_{\text{onto}} b \Longleftrightarrow_{\text{df}} f : a \longrightarrow b \wedge \text{ran}(f) = b$ *("f is onto").*
(vi) $f : a \longleftrightarrow b \Longleftrightarrow_{\text{df}} f : a \longrightarrow_{(1-1)} b \wedge f : a \longrightarrow_{\text{onto}} b$ *("f is a bijection").*

Definition 1.9 *(i)* ${}^a b =_{\text{df}} \{f \mid f : a \longrightarrow b\}$ *the class of all functions from a to b.*
(ii) Let f be a function such that $\varnothing \notin \text{ran}(f)$. Then the generalised cartesian product *is*

$$\prod f =_{\text{df}} \{h \mid \text{Fun}(h) \wedge \text{dom}(h) = \text{dom}(f) \wedge \forall x \in \text{dom}(f)(h(x) \in f(x))\}.$$

Note that $\prod f$ consists of *choice functions* for $\text{ran}(f)$: each h "chooses" an element from each appropriate set.

1.2.2 The Zermelo-Fraenkel Axioms

The axioms of ZFC (Zermelo-Fraenkel with Choice) then are the following:

Ax0 ***(Extensionality)*** $\forall x \forall y(\forall z(z \in x \leftrightarrow z \in y) \leftrightarrow x = y)$
Ax1 ***(Empty Set)*** $\varnothing \in V$
Ax2 ***(Pairing Axiom)*** $\{x, y\} \in V$
Ax3 ***(Union Axiom)*** $\bigcup x \in V$
Ax4 ***(Foundation Scheme)*** For every term a: $a \neq \varnothing \longrightarrow \exists x(x \in a \wedge x \cap a = \varnothing)$
Ax5 ***(Separation Scheme)*** For every term a: $x \cap a \in V$
Ax6 ***(Replacement Scheme)*** For every term f: $\text{Fun}(f) \longrightarrow f``x \in V$.
Ax7 ***(Infinity Axiom)*** $\exists x(\varnothing \in x \wedge \forall y(y \in x \longrightarrow y \cup \{y\} \in x))$
Ax8 ***(PowerSet Axiom)*** $\mathcal{P}(x) \in V$
Ax9 ***(Axiom of Choice)*** $\text{Fun}(f) \wedge \text{dom}(f) \in V \wedge \varnothing \notin \text{ran}(f) \longrightarrow \prod f \neq \varnothing$.

Note 1.10 (i) ZF comprises Ax0-8; Sometimes Ax6 is replaced by:
Ax6* ***(Collection Scheme)*** For every term r: $\forall x r``x \neq \varnothing \longrightarrow \forall w \exists t(\forall u \in w \exists v \in t(\langle u, v\rangle \in r)$.
The Axiom of Choice is equivalent over ZF to the Wellordering Principle:
Ax9* ***(Wellordering Principle)*** $\forall x \exists r(\text{Rel}(r) \wedge \langle x, r\rangle$ *is a wellordering*$)$.
There are two useful subsystems. ZF with Replacement dropped is called Z for Zermelo. ZF^- is Ax0-5,6*,7; ZFC^- is ZF^- with Ax9*.

(ii) ZF is an infinite list of axioms: Ax4,5,6, (and 6*) are *schemes*: there is one axiom for each formula defining the mentioned terms a and f (or r in 6*). We shall later prove that it cannot be replaced by a finite list with the same consequences.

(iii) The statements differ in their formulation from ST: Foundation was there stated just for sets, and was a single axiom; Separation was given its synonym "Comprehension": and was stated as follows:

(*The set of elements of a set z satisfying some formula, form a set.*)

For each formula $\varphi(v_{0,} \ldots v_{n+2})$ *(with free variables amongst those shown),*

$$\forall z \forall w_1 \ldots \forall w_n \exists y \forall x (x \in y \leftrightarrow x \in z \wedge \varphi[z, x, w_1, \ldots, w_n]).$$

The formulation above shows how powerful and succinct a formulation we have if we allow ourselves to use terms. Likewise Replacement there had a much longer (but equivalent) formulation.

(iv) The axioms are of different kinds: one group asserts that simple operations on sets leads to further sets (such as Union, Pairing). Another group consists of set existence axioms (Empty Set, Separation). Others are of "delimiting size" nature: the power class $\mathcal{P}(x)$ may be thought to be a large incoherent collection of *all* subsets of x. The power set axiom claims that this is not a large collection but merely another set. The Replacement Scheme assures us that functions however defined cannot create a non-set from a set. It thus also in effect delimits size. This axiom is due to Fraenkel. The term 'Replacement' comes from the idea that if one has a set, and a method (or function) for replacing each member of that set by a different set, then the resulting object is also a set. Zermelo's achievement was to recognize (a) the utlility, if not the necessity, of formulating a formal set of axioms for the new subject of set theory - which he then enunciated; (b) that the Separation scheme was a method to avoid paradoxes of the Russell/Burali-Forti kind. Zermelo essentially wrote down the system Z although Separation was given a second order formulation. Later Skolem gave the familiar first order formulation equivalent to the above. Again the Axiom of Choice asserts the existence of a rather specialised set: a choice function for a collection of sets. In ST we adopted the axiom that "Every set can be wellordered" for AC (on pedagogical grounds). We saw there that this principle was equivalent to the existence of choice functions.

(v) One may ask simply: Are these right axioms? There are indeed other formulations of set theory, some involving class terms more directly as further objects. Our point of view is that the V hierarchy comprises all that is needed for mathematics, further we have a somewhat less developed intuition about what such "objects" these free-standing class terms could possible be: if they are attempts to continue the V-hierarchy even further, by using the power class operation "just one more time" this would seem to miss the point. Since we have no need for classes as some other kind of separate entities of a different sort, we avoid them.

One formulation of set theory (which Gödel used - and is named von Neumann-Gödel-Bernays) does however include *class variables* in the object language but disallows quantification over classes: it can be shown that this system is *conservative over* ZFC: that is, it proves no more theorems about sets than ZFC itself, and so is treated by set theorists virtually as a harmless variant of ZFC.

We use a *first order* formulation of set theory (meaning that quantifiers $\exists x$,$\forall x$ quantify only over our objects of interest, namely sets. A second order formulation ZF^2 is possible, where, as in any second order language, we are allowed quantifiers such as $\exists P$, $\forall P$ that range over predicates P of sets. There are two points that could be made here. Firstly, as a predicate P is extensionally a collection of sets itself, even to understand the *meaning* of a second order sentence involving say a quantifier $\forall P$ is to already claim an understanding about the universe V. And it is V itself that we are trying to understand in the first place. As in all areas of mathematics, first order formulations of theories are the most successful: we may not know of a first order sentence σ whether it is true or not, but we do know precisely what it means for it to be true. Secondly the tools of mathematical logic are the most useful in the setting of first order logic. The deductive system associated to ZF^2 lacks a Completeness Theorem, and hence Compactness and Löwenheim-Skolem Theorems fail. In ZF^2 it is possible to argue that since the only possible models of ZF^2 are V itself and possible initial segments of V of the form V_κ (as Zermelo demonstrated), then ZF^2 shows that, *e.g.* CH has a definite truth value: namely that obtained by inspecting that level of the V-hierarchy where all subsets of $\mathbb{N}$ live: $V_{\omega+1}$. However as to what that truth value is, we have no idea. Hence we are no further forward! Indeed second order logic and ZF^2 seems not to give us any tangible information about the universe of sets that we do not obtain from the first order formulations of ZFC and its extensions.

(vi) Some formulations or viewpoints concerning the mathematical hierarchy of sets take as the base of that hierarchy not the empty set ("$V_0 = \varnothing$") but rather a collection of "atoms" or base objects: thus instead we take $V_0[U] = U$ where U is this collection of *Urelemente* and we build our hierarchy by iterating the power set operation from this point onwards. This may be of presentational benefit, but, at least if U is a set (meaning that it has a cardinality), then this is of limited foundational interest to the pure set theorist.[4] The reason being, that, if $|U| = \kappa$ say, then we may build an isomorphic copy of $V[U]$ *inside* V, by starting with some κ sized set or structure which is an appropriate copy of U. Hence again to study V is to study all such universes $V[U]$, and we may limit our discussion to universes of "pure sets" without additional atomic elements.

1.3 Transfinite Recursion

We recall the definitions of *transitive set*.

Definition 1.11 *x is* transitive (Trans(x)) *if* $\forall z \in x(z \subseteq x)$.

We have the following scheme of $\in$-induction:

Lemma 1.12 **(scheme of $\in$-induction)** *For any formula φ:*

$$\forall x[\forall y \in x \varphi(y) \to \varphi(x)] \to \forall x \varphi(x).$$

This principle was used in the proof of:

Theorem 1.13 **(Transfinite Recursion along $\in$)**
If G is a term and $G : V \times V \to V$ then there is a term F giving $F : V \to V$

$$(*) \qquad \forall x F(x) = G(x, F \restriction x).$$

Moreover the term defines a unique function, in that if F' is any other term satisfying $(*)$ then, $\forall x F(x) = F'(x)$.

[4] This is not to say that models with atoms are without utility: formulations of ZFwith atoms, "ZFA", are of great use for studying universes in which the Axiom of Choice fails. The point being made is that we cannot get any additional knowledge about *foundational* questions by using them.

Note: (i) Usually one speaks instead of G, F being defined by formulae φ_G, φ_F etc., but we have replaced that with talk about terms. In the proof of Theorem 1.13 we, in effect, saw how to build up from the formula φ_G the formula φ_F. This is in essence a *Theorem Scheme*: it is one theorem for each term G. The 'canonical' procedure for building the formula φ_F given φ_G now becomes a method for building a canonical term defining F from one defining G.

(ii) Often one first proves a transfinite recursion theorem along On: as the ordering relation amongst ordinals *is* the $\in$-relation, we can view the latter theorem as simply a special case of Theorem 1.13. From these we proved the existence of functions giving the arithmetical operations on ordinals, and their basic properties. It is often useful to have the notion of a wellfounded relation in general:

DEFINITION 1.14 *If R is relation on a class A then we say R is wellfounded iff for any z, if $z \cap A \neq \varnothing$ then there is $y \in z \cap A$ which is R-minimal (that is $\forall x \in z \cap A(\neg xRy)$).*

An important example of a definition by transfinite recursion along $\in$ is that of the *transitive closure operation* TC.

DEFINITION 1.15 TC *is that class term given by Theorem 1.13 satisfying*

$$\forall x\, \mathrm{TC}(x) = x \cup \bigcup\{\mathrm{TC}(y) \mid y \in x\}$$

EXERCISE 1.1 In ST TC was given an alternative (but equivalent) definition, and was shown to satisfy the definition of TC above. Rework this by showing, using the above definition, that: (i) $x \in y \longrightarrow \mathrm{TC}(x) \subseteq \mathrm{TC}(y)$. (ii) Show that $\mathrm{TC}(x)$ is the smallest transitive set t satisfying $x \subseteq t$. [Hint: Use $\in$-recursion.] (Thus if $\mathrm{Trans}(t) \wedge x \subseteq t \longrightarrow \mathrm{TC}(x) \subseteq t$.) Moreover $\mathrm{Trans}(x) \longleftrightarrow \mathrm{TC}(x) = x$. (iii) Define by recursion on ω: $\bigcup^0 x = x; \bigcup^{n+1} x = \bigcup(\bigcup^n x)$; $\mathrm{tc}(x) = \bigcup\{\bigcup^n(x) \mid n < \omega\}$. Show that $\mathrm{tc}(x) = \mathrm{TC}(x)$.

DEFINITION 1.16 *For $x \subseteq On, x \in V$, $\sup(x) =_{df}$ the least ordinal γ so that $\beta \in x \to \beta \leq \gamma$.*

In particular if x has no largest element, then $\sup(x) = \bigcup x$.

DEFINITION 1.17 **(The rank function ρ)** *The rank function is defined by transfinite recursion on $\in$:*

$$\rho(x) = \sup\{\rho(y) + 1 \mid y \in x\}.$$

EXERCISE 1.2 Show that: (i) the relation $xRy \longleftrightarrow x \in \mathrm{TC}(y)$ is wellfounded; (ii) $\forall x(\rho(x) = \rho(\mathrm{TC}(x)))$; (iii) $\mathrm{Trans}(x) \longrightarrow \rho``x \in \mathrm{On}$.

DEFINITION 1.18 **(The Cumulative Hierarchy)** *The V_α function is defined by transfinite recursion on* On *as : $V_\alpha = \{x \mid \rho(x) < \alpha\}$.*

In ST we defined the V_α hierarchy by iterating the power set operation. The previous definition does not use AxPower and together with the next exercise shows that we can define the latter hierarchy without it.

EXERCISE 1.3 Define by recursion $R_0 = \varnothing, R_{\alpha+1} = \mathcal{P}(R_\alpha)$ and for $\mathrm{Lim}(\lambda)$, $R_\lambda = \bigcup_{\alpha<\lambda} R_\alpha$. Show by transfinite induction that for any $\alpha \in \mathrm{On}$ that $R_\alpha = V_\alpha$.

1.4 Relativisation of terms and formulae

We may classify concepts according to the syntactic complexity of their definitions. Accordingly we then first classify formulae of our language $\mathcal{L}$ as follows.

Bounded quantifiers: $\forall v_i \in v_j\psi$, $\exists v_i \in v_j\psi$ abbreviate: $\forall v_i(v_i \in v_j \to \psi)$ and $\exists v_i(v_i \in v_j \wedge \psi)$ respectively.

The Levy hierarchy: We stratify formulae according to their complexity by counting alternations of quantifiers. We first define the Δ_0-formulae of $\mathcal{L}$ inductively:

(i) *$v_i \in v_j$ and $v_i = v_j$ are Δ_0.*)

(ii) *If φ, ψ are Δ_0, then so are $\neg\varphi$ and $(\varphi \vee \psi)$.*

(iii) *If φ is Δ_0 so is $\exists v_i \in v_j\varphi$.*

Having defined Δ_0 as those without unbounded quantifiers, we then proceed, first setting $\Sigma_0 = \Pi_0 = \Delta_0$:

(i) *If φ is Π_{n-1} then $\exists v_{i_1} \cdots \exists v_{i_n}\varphi$ is Σ_n.*

(ii) *If φ is Σ_n then $\neg\varphi$ is Π_n.*

One should note that if a formula is classified as Σ_n then it is logically equivalent to a Σ_m formula (or to a Π_m-formula) for any $m \geq n$, by the trivial process of adding dummy quantifiers at the front. Of particular interest are *existential* formulae: those that are Σ_1: $\exists x\varphi$ with φ Δ_0. Such assert a simple set existence statement, and *universal* formulae: these are Π_1: $\forall x\varphi$ whose truth requires, *prima facie*, an inspection over all sets (although in practice we shall see that by the Downward Lowenheim Skolem Theorem, we may sometimes restrict that apparent unbounded search). Occasionally, for T a finite set of formulae, we write $\bigwedge\!\!\!\bigwedge T$ for the single formula that is their conjunction.

Some terms will be seen to be *definite* in that they define the same object in whatever world the definition takes place. This may sound obscure at the moment, but one can perhaps see that the definition of the empty set provides a definite object $\varnothing$ which is "constant" across possible universes where it might be defined; likewise given any structure U with sets x, y as members and in which the Pairing Axiom holds, then the term $t = \{u \mid u = x \vee u = y\}$ defines the *same* object in U as in any other structure satisfying these conditions. This is in contradistinction to a term such as $t = \{y \mid y \subseteq x\}$ which defines the power set of a set x: although the defining formula "$v_0 \subseteq v_1$" is extremely simple, which subsets of x get picked up when we apply the definition, depends on which *structure U* we apply the definition in. It is thus not a *definite term*. We shall need to investigate this and give a criterion for when terms are definite. This leads on to the important notion of *absoluteness*.

We shall be interested in looking at models $\langle M, E\rangle$ of ZFC + Φ for various statements Φ. For this to be really meaningful we shall want that certain terms and notions defined by certain formulae that are interpreted in the model $\langle M, E\rangle$ mean the same thing as when that term or formula is applied in $\langle V, \in\rangle$: this is the notion of "absoluteness". Certain (simple) objects, such as $\varnothing, \omega$ and the like, are defined by the same syntactic terms evaluated in V or in M. It is possible to think about models where the interpretation of the $\in$symbol is something other than the usual set membership relation. Such models are called *non-standard* models, and do not feature highly in this course (or in the wider development of set theory). We shall be most interested in *transitive* sets or classes W and where E is taken to be the genuine set membership relation $\in$. Such an $\langle W, \in\rangle$ is called a *transitive $\in$-model*. However terms can have different interpretations even when considered in V and in a standard transitive model $\langle W, \in\rangle$. We first have to say what it means for an axiom or sentence φ to "hold" or "to be interpreted" in such a structure. We build up a definition by recursion on the structure of φ by straightforwardly restricting quantifiers to the new putative "universe" W.

DEFINITION 1.19 *Let W be a term. We define by recursion on complexity of formulae φ of $\mathcal{L}$ the* relativisation *of φ to W, φ^W:*
(i) $(x \in y)^W =_{\mathrm{df}} (x \in y)$; $(x = y)^W =_{\mathrm{df}} (x = y)$;
(ii) $(\neg\psi)^W =_{\mathrm{df}} \neg(\psi^W)$;
(iii) $(\psi \vee \chi)^W =_{\mathrm{df}} (\psi^W \vee \chi^W)$;
(iv) $(\exists x\psi)^W =_{\mathrm{df}} \exists x \in W\psi^W$ *if x is not in* $\mathrm{FVbl}(W)$ *; otherwise this is undefined.*

Notice that we can always ensure that $(\varphi)^W$ is defined by replacing the bound variables in φ by others different from those of W. We tacitly that this has always been done when discussing relativising formulae. It is immediate that, *e.g.* $(\forall x\psi)^W \longleftrightarrow \forall x \in W\psi^W$. We shall be thinking of class terms W as being potential $\in$- structures - meaning that we shall be thinking of them potentially as models $\langle W, \in\rangle$. We shall read $(\varphi)^W$ as "φ holds in W" or "φ holds relativised to W". The following theorem (with $\Gamma = \varnothing$) says the theorems of predicate calculus in $\mathcal{L}$ are valid in non-empty $\in$-structures $\langle W, \in\rangle$. We use the shorthand that if Γ is a finite set of formulae, then $\bigwedge\!\!\!\bigwedge \Gamma$ is the single formula that is the conjunction of those in Γ.

THEOREM 1.20 *Let $\Gamma \cup \{\sigma\}$ be a finite set of sentences in $\mathcal{L}$ and W a transitive non-empty term; assume that if $\vec{x}$ is a list of all the variables occurring in $\Gamma \cup \{\sigma\}$ then $\vec{x} \cap \mathrm{FVbl}(W) = \varnothing$.*
If $\Gamma \vdash \sigma$ then $(\bigwedge\!\!\!\bigwedge \Gamma)^W \longrightarrow \sigma^W$.

PROOF: By induction on the length of the derivation of σ from Γ. Q.E.D.

This is just as it should be: roughly, it is a form of *Soundness*: if we can prove that σ is derivable from a set of axioms true in a structure, then σ should be true in that structure.

LEMMA 1.21 *Let W be a transitive class term, Then* $(\mathrm{AxExt})^W$.

PROOF: The Axiom of Extensionality relativised to W is:
$(\forall x\forall y(\forall z(z \in x \leftrightarrow z \in y) \to x = y))^W$
$\leftrightarrow \forall x \in W\forall y \in W(\forall z(z \in x \leftrightarrow z \in y) \to x = y)^W$
$\leftrightarrow \forall x \in W\forall y \in W(\forall z \in W(z \in x \leftrightarrow z \in y)^W \to (x = y)^W)$
$\leftrightarrow \forall x \in W\forall y \in W(\forall z \in W(z \in x \leftrightarrow z \in y) \to x = y)$
Since W is transitive, if $x, y \in W$ then $x, y \subseteq W$. Hence if $\exists z(z \in x\backslash y \cup y\backslash x)$ then $\exists z \in W(z \in x\backslash y \cup y\backslash x)$. Hence the $\to$ of the last equivalence is true! Q.E.D.

The next concern is how to relativise a formula that contains class terms. It should turn out that if we have such a formula we should be able to first relativise the terms it contains to W (Def.1.22) and then substitute the results into the relativised formula of $\mathcal{L}$.

DEFINITION 1.22 *Let $t = \{x \mid \varphi\}$ be a class term; the* relativisation of t to W, *is:* $t^W =_{\mathrm{df}} \{x \in W \mid \varphi^W\}$.

Example (i) $V^W = \{x \mid x = x\}^W = \{x \in W \mid (x = x)^W\}$. Since $(x = x)^W$ is just $x = x$, this renders $V^W = V \cap W = W$.

Example (ii) $(\bigcup x)^W = (\{z \mid \exists y(z \in y \in x)\})^W = \{z \in W \mid (\exists y(z \in y \in x))^W\} = \{z \in W \mid \exists y \in W(z \in y \in x)^W\} = \{z \in W \mid \exists y \in W(z \in y \in x)\}$.

Notice that if additionally W is a transitive term, (*i.e.* defines a transitive class) then $x \in W \longrightarrow x \subseteq W$; moreover $\forall y(y \in x \longrightarrow y \subseteq W)$. Hence $\{z \in W \mid \exists y \in W(z \in y \in x)\} = \{z \mid \exists y(z \in y \in x)\}$ and so in this case $(\bigcup x)^W = \bigcup x$. This demonstrates that $\bigcup$ is an *absolute operation for transitive classes* and the process of relativisation yields the same set. We shall be particularly interested in such absolute operators and similarly absolute properties for transitive classes.

Lemma 1.23 *Let* $t_0, \ldots, t_n$ *and* W *be terms, with* W *transitive, let* $\varphi(x_0, \ldots, x_n)$ *be in* $\mathcal{L}$; *then assuming* $\vec{y} \supseteq \mathrm{FVbl}(\varphi(t_0, \ldots, t_n))$:

$$\forall \vec{y} \in W(\varphi(t_0, \ldots, t_n)^W \longleftrightarrow \varphi^W(t_0^W, \ldots, t_n^W)).$$

Remark: The lemma is about syntax, formulae and terms. The x_i's are (meta-)variables ("meta" because they are standing in for some official variables $v_{i_0}, \ldots v_{i_n}$). In this context the notation is supposed to mean that each of the terms $t_0, \ldots, t_n$ is then substituted for the corresponding variable $x_0, \ldots x_n$. Above we said that we should more properly indicate this by: "$\varphi(t_0/x_0, \ldots, t_n/x_n)$" but this becomes too cumbersome, and too tedious to do all the time, so we just leave it for the reader to do depending on the context.

Proof: By induction on the complexity of φ. Q.E.D.

Exercise 1.4 Convince yourself of the truth of the last lemma. [Hint: At least set out the base cases of the induction: suppose φ is $v_0 \in v_1$ and let $t_0 = x$, $t_1 = \{z \mid \psi\}$. Then $(x \in t_1)^W \leftrightarrow (x \in \{z \mid \psi\})^W \leftrightarrow \psi(x/z)^W \leftrightarrow x \in \{z \mid z \in W \wedge \psi^W\} \leftrightarrow x^W \in (t_1)^W$. The other base cases are relatively straightforward, but a little lengthy to write out. The inductive step for non-atomic formulae is easy by comparison.]

Lemma 1.24 *Let* W *be a transitive term and suppose for any* $x, y \in W$, $\{x, y\} \in W$, *then* $(\mathrm{AxPair})^W$.

Proof: We need to show $(\{x, y\} \in V)^W$. First just note that by Def. 1.22:

$$\{x, y\}^W = \{z \in W \mid (z = x \vee z = y)^W\} = \{z \in W \mid z = x \vee z = y\} = \{x, y\}.$$

By supposition we have that: $\forall x, y \in W(\{x, y\} \in W) \leftrightarrow$

$$\leftrightarrow \forall x, y \in W(\{x, y\}^W \in V^W) \leftrightarrow (\forall x, y(\{x, y\} \in V)^W.$$

(The last $\leftrightarrow$ uses implicitly an atomic formula clause from **??**) Q.E.D.

Lemma 1.25 *Let* W *be a transitive term.*

(i) If for any $x \in W$, $\bigcup x \in W$ *then* $(\mathrm{AxUnion})^W$;

(ii) If $\omega \in W$ *then* $(\mathrm{Ax.Infinity})^W$.

(iii) If for any $x \in W$ *and any term* a $x \cap a^W \in W$ *then* $(\mathrm{AxSeparation})^W$;

(iv) If for any $x \in W$, *and term* f *with* f^W *being a function,* f^W"$x \in W$ *holds, then* $(\mathrm{AxReplacement})^W$.

(v) If for any term r *with* r^W *being a relation with* $\forall x \in W r^W$"$x \neq \varnothing$, *and if for any* $z \in W$ *there is* $w \in W$ *so that* $(\forall u \in z \exists v \in w(r(\langle u, v\rangle))^W$ *holds, then* $(\mathrm{AxCollection})^W$.

PROOF: (i) By *Example (ii)* above, because W is assumed transitive, $(\bigcup x)^W = \bigcup x$. Moreover $V^W = \{z \in W \mid z = z\} = W$. By assumption $\forall x \in W \bigcup x \in W$. Hence

$\forall x \in W(\bigcup x)^W \in W \leftrightarrow \forall x \in W(\bigcup x \in V)^W \leftrightarrow (\forall x \bigcup x \in V)^W$; the latter is $(\text{AxUnion})^W$.

(iii) We need to show $(\forall x \;\; a \cap x \in V)^W$. Suppose $\vec{y} = \text{FVbl}(a)$. This is equivalent (by Lemma **??**) to, $\forall \vec{y} \in W$:

$$\forall x \in W((a \cap x)^W \in V^W)^W \leftrightarrow \forall x \in W((a \cap x)^W \in W).$$

But, for any $x \in W$, :

$$(a \cap x)^W = \{z \in W \mid (z \in a \wedge z \in x)^W\} = \{z \in W \mid z \in a^W \wedge z \in x\}.$$

As $\text{Trans}(W)$, $x \subseteq W$, so this is $a^W \cap x$. By assumption this is indeed in W. Q.E.D.

EXERCISE 1.5 Show (ii), (iv) and (v) of the last Lemma.

LEMMA 1.26 *Let W be a non-empty transitive term satisfying all the hypotheses of Lemmata 1.24, 1.25. Then* $(\text{ZF}^-)^W$ *that is, each axiom of* ZF^- *holds in* W.

PROOF: We are only left with the Axioms of the Empty Set and Foundation. But $\varnothing^W = \varnothing$ (Check!), and $\varnothing$ is a member of any non-empty transitive class (why?). For Foundation let a be a term, and suppose that $(a \neq \varnothing)^W$. Suppose $x \in a^W \cap W$. Now, by Axiom of Foundation (applied in V) as $a^W \neq \varnothing$, let x_0 be an element of a^W with $x_0 \cap a^W = \varnothing$. Hence $(a \neq \varnothing \rightarrow \exists z(z \cap a = \varnothing))^W$. Q.E.D.

Lemma 1.26 is again a theorem scheme: given a class term for which we can prove the assumptions hold for it, (which itself is an infinite list of proofs in ZF if *all* the assumptions of Lemma 1.25 are verified) *then* the lemma states that for any axiom φ of ZF^- then $\text{ZF} \vdash \varphi^W$. (This can be trivially extended to a finite list of axioms $\vec{\varphi}$ by taking a simple conjunction - but it cannot be extended to an infinite list!) The next lemma gives a sufficient (but not necessary) condition for AxPower to hold in a transitive class term. The proof is similar to those above.

LEMMA 1.27 *Let W be a transitive term satisfying for any $x \in W$, that $\mathcal{P}(x) \in W$; then* $(\text{AxPower})^W$. *Consequently if W satisfies this in addition to the hypothesis of the last lemma then* $(\text{ZF})^W$, *that is all of* ZF *holds in W.*

We shall see later that we can prove the existence of transitive $\in$-models $\langle W, \in \rangle$, with W a set, for which $(\text{ZF}^-)^W$, by establishing the existence of transitive sets satisfying precisely the above closure conditions. We thus shall show for such a W that, assuming ZF, we can show $(\text{ZF}^-)^W$. However in ZF we cannot prove the existence of sets (transitive or otherwise) W for which $(\text{ZF})^W$. (We shall see that this leads to a contradiction with Gödel's Second Incompleteness Theorem.)

EXERCISE 1.6 Let $\varphi(v_0, \ldots, v_n)$ be any formula. Let $g_\varphi(\vec{y})$ = the least β such that $\exists x \varphi(x, \vec{y}) \rightarrow \exists x \in V_\beta \varphi(x, \vec{y})$ if such an x exists; let it be 0 otherwise. Show that $\forall \xi \; g_\varphi``V_\xi \in V$. Deduce that $f_\varphi(\xi) =_{df} \sup(g_\varphi``V_\xi)$ is a welldefined function.

EXERCISE 1.7 Let W be the class term $\{\varnothing\}$. Which axioms of ZFC hold in $\langle W, \in \rangle$? Consider the class term On. Which axioms of ZFC hold in $\langle On, \in \rangle$? (NB For the latter $\langle \text{On}, \in \rangle$ just is $\langle \text{On}, < \rangle$.)

EXERCISE 1.8 Which axioms of ZFC hold in V_ω?

EXERCISE 1.9 Check, or recheck, the following basic properties of the V_α using the Definitions 1.17, 1.18 of ρ and V_α: (i) $\mathrm{Trans}(V_\alpha)$; *in particular show if* $x \in V_\alpha$ *then* $\forall y \in x(y \in V_\alpha \wedge \rho(y) < \rho(x))$;
(ii) $\alpha < \beta \longrightarrow V_\alpha \subseteq V_\beta$;
(iii) $V_{\alpha+1} = \mathcal{P}(V_\alpha)$;
(iv) *If* $x \in V$, *then* $\rho(x)$ = *least* α *so that* $x \subseteq V_\alpha$ = *least* α *such that* $x \in V_{\alpha+1}$.
(v) $\rho(\alpha) = \alpha$;
(vi) $\mathrm{On} \cap V_\alpha = \alpha$.

EXERCISE 1.10 There are a number of definable wellorders on ${}^nOn$: here is one: for $\vec{\alpha} = \langle \alpha_0, \ldots, \alpha_{n-1} \rangle, \vec{\beta} = \langle \beta_0, \ldots, \beta_{n-1} \rangle$ set $\vec{\alpha} <^n \vec{\beta}$ iff $\max(\vec{\alpha}) < \max(\vec{\beta})$ or $(\max(\vec{\alpha}) = \max(\vec{\beta})) \wedge ($ if i is least so that $\alpha_i \neq \beta_i$ then $\alpha_i < \beta_i)$. $<^n$ is then Δ_0 expressible. Check that this is a wellorder.

EXERCISE 1.11 Prove that the following is a wellorder of $On^{<\omega}$ where the latter is the class of finite sets of ordinals: for $p, q \in [On]^{<\omega}$ define $p <^* q$ iff $\max\{p \triangle q\} \in q$. That is, $p <^* q$ iff the largest element of $p \backslash q \cup q \backslash p$ is in q. [Hint: it is perhaps to easier to observe first that this ordering is just the lexicographic ordering on the sequences $\vec{p}, \vec{q} \in {}^{<\omega}On$ of the sets p, q when written out as sequences in *descending* order.]

Chapter 2

Initial segments of the Universe

In this chapter we look at some properties of initial segments of the universe V: typically local properties of singular and regular cardinals, and the classes of sets *hereditarily of cardinality* less than some κ. These do not depend on the whole universe of sets. We shall see that when studying wellfounded models of our theory, it suffices to concentrate our efforts on models $\langle M, \in\rangle$ where M is a *transitive* set, rather than more general $\langle N, E\rangle$. An important application of the Axiom of Replacement is the *Montague-Levy Reflection Theorem*: this says that for any given finite set of formulae, we can prove in our theory that there are arbitrarily large V_α that correctly 'reflect the truth' as regards what those formulae say about the sets in V_α. Cardinals κ that are simultaneously both fixed points of certain functions and regular are called *strongly inaccessible*. If such exist then we can find models, indeed of the form $\langle V_\alpha, \in\rangle$, of all the ZFC axioms. We discuss these in the last section.

2.1 Singular ordinals: cofinality

We first do some basic work on notions of *regularity*, *singularity* and *cofinality*. This then leads into the concepts of *normal functions* and *closed and unbounded sets*, and *stationary* sets. From these further large cardinals can be defined, and although we give the briefest of illustrative examples, it is not the intention of the course to go down this route, rich as it is.

2.1.1 Cofinality

Definition 2.1 *If $A \subseteq \mu$ is a set of ordinals, then we say that A* is unbounded below (or in) μ *iff* $\forall\alpha < \mu\exists\beta \in A(\beta > \alpha)$.

Note that implicitly in the above, we have that $Lim(\mu)$, *i.e.* that μ is automatically a limit ordinal.

Definition 2.2 *A function $f : \alpha \longrightarrow \beta$ is a* cofinal *map, if* $\sup(\mathrm{ran}(f)) = \beta$.

In other words the range of f is *unbounded* in β. This definition also then implicitly implies that $Lim(\beta)$.

Example (i) $f : \omega \longrightarrow \omega + \omega$ given by $f(n) = \omega + n$;

(ii) $f : \omega \longrightarrow \omega_\omega$ given by $f(n) = \omega_n$;

(iii) $g : \omega_1 \longrightarrow \omega_{\omega_1}$ given by $g(\alpha) = \omega_\alpha$ are all cofinal maps.

(iv) Define the sequence $f(0) = \omega_0; f(n+1) = \omega_{f(n)}$. Let $\kappa = \sup(f``\omega)$. Then $f : \omega \longrightarrow \kappa$ is cofinal - by construction. Note also here that $\kappa = \omega_\kappa$. (Check!) Such a κ is a *fixed point* in the enumeration of the infinite cardinals.

(v) Let $E \subseteq \beta$ be any subset. Suppose its order type is τ. We use the notation f_E for the (1-1) function that *enumerates E* in strictly increasing order. Thus $\mathrm{dom}(f_E)$ will be τ which will necessarily be no greater than β. If E is now *unbounded* in β, that is $\forall \gamma < \beta \exists \delta \in (\gamma, \beta)(\delta \in E)$, then $f_E : \tau \to \beta$ will be a cofinal map into β, that is moreover (1-1) and strictly increasing. Then any of the functions in *(i)-(iv)* can be regarded as enumerating maps of their ranges.

Definition 2.3 *The* cofinality *of a limit ordinal β is the least α so that there is a cofinal map $f : \alpha \longrightarrow \beta$. It is denoted* $\mathrm{cf}(\beta)$.

Taking f as the identity map, shows immediately that $\mathrm{cf}(\beta) \leq \beta$.

Definition 2.4 *(i) A limit ordinal β is* singular $\Longleftrightarrow_{df}$ $\mathrm{cf}(\beta) < \beta$. *Otherwise it is called* regular.

(ii) We set:

$\mathrm{Reg} =_{\mathrm{df}} \{\kappa \mid \kappa \textit{ is regular}\}$; $\mathrm{Card} =_{\mathrm{df}} \{\kappa \mid \kappa \textit{ a cardinal}\}$;

$\mathrm{SingCard} =_{\mathrm{df}} \{\kappa \in \mathrm{Card} \mid \kappa \textit{ singular}\}$; $\mathrm{LimCard} =_{\mathrm{df}} \{\alpha \in \mathrm{Card} \mid \alpha \textit{ a limit cardinal}\}$;

$\mathrm{SuccCard} =_{\mathrm{df}} \{\alpha \in \mathrm{Card} \mid \alpha \textit{ a successor cardinal}\} = \{\tau^+ \mid \tau \in On\}$.

Example (i) $\mathrm{cf}(\omega + \omega) = \omega$. The above example shows that $\mathrm{cf}(\omega + \omega) \leq \omega$; but it cannot be strictly less since no function with finite domain can have unbounded range in $\omega + \omega$. The same holds for *(ii)* above $\mathrm{cf}(\omega_\omega) = \omega$. and $\aleph_\omega = \omega_\omega$ is an example of a cardinal with a smaller cofinality. It will follow from below that $\mathrm{cf}(\omega_{\omega_1}) = \omega_1$.

The following is immediate from our definition of cardinality and cofinality.

Lemma 2.5 $\mathrm{cf}(\beta) \leq |\beta| \leq \beta$. *Thus, a regular ordinal must be a cardinal; to rephrase:*

$$\mathrm{cf}(\beta) = \beta \Longleftrightarrow \beta \text{ is regular} \Longleftrightarrow \beta \text{ is regular and a cardinal.}$$

Examples: $\omega = \omega_0 = \aleph_0 \in \mathrm{Reg}$ (Hausdorff 1908); $\omega_1 = \aleph_1 \in \mathrm{Reg}$, indeed:

Lemma 2.6 (Hausdorff 1914) *Any* $\lambda^+ \in \mathrm{Reg}$.

Proof: Suppose this failed then note that if $f : \alpha \longrightarrow \lambda^+$ with $\mathrm{ran}(f)$ unbounded in λ^+, but $\alpha < \lambda^+$, we would have that $\lambda^+ = \bigcup_{\beta<\alpha} f(\beta)$ - in other words, taking $\lambda \in \mathit{Card}$, the union of $|\alpha| \leq \lambda$ many sets of size $\leq \lambda$. Assuming AC this is impossible: as $\lambda \otimes \lambda = \lambda$, this union could have size at most λ! Q.E.D.

Thus any $\aleph_{\alpha+1} = \aleph_\alpha^+$ is regular. These are called *successor cardinals* (being indexed by successor ordinals). The first singular cardinal is $\aleph_\omega$, the next is $\aleph_{\omega+\omega}$; also $\aleph_{\omega_1}, \aleph_{\omega_\omega} \in \mathrm{Sing}$. By Hausdorff's observation above, a singular cardinal is always a *limit cardinal*: it occurs as a limit point of the cardinal enumeration function: $\alpha \rightarrowtail \aleph_\alpha$. Later we shall consider the question of whether the converse fails, that is whether there are cardinals that are simultaneously limit cardinals and regular.

Lemma 2.7 *For any limit ordinal β:*
(i) $\mathrm{cf}(\beta)$ is the least ordinal α so that there is a (1-1) strictly increasing cofinal map $f : \alpha \longrightarrow \beta$;
(ii) $\mathrm{cf}(\mathrm{cf}(\beta)) = \mathrm{cf}(\beta)$; hence (Hausdorff 1908) $\mathrm{cf}(\beta)$ is regular;
(iii) If $f : \alpha \longrightarrow \beta$ is cofinal and strictly increasing, then $\mathrm{cf}(\alpha) = \mathrm{cf}(\beta)$.

Proof: (i) Let $f : \mathrm{cf}(\beta) \longrightarrow \beta$ be any cofinal map. We define a $g : \mathrm{cf}(\beta) \longrightarrow \beta$ of the desired kind from f by recursion on $\delta < \mathrm{cf}(\beta)$:

$$g(0) = f(0);\ g(\delta+1) = \max\{g(\delta)+1, f(\delta)\} \text{ and } \mathrm{Lim}(\lambda) \to g(\lambda) = \sup\{g(\delta) \mid \delta < \lambda\}.$$

Note that a) $g(\delta) < \beta$ implies $g(\delta+1) < \beta$ and b) for any $\mathrm{Lim}(\eta)$ if $\eta < \mathrm{cf}(\beta)$, then $g(\eta)$ is properly defined, and thus is less than β. Thus we have $\mathrm{dom}(g) = \mathrm{cf}(\beta)$. By definition g is strictly increasing (and moreover is *continuous* at limit ordinals λ - see Def. 2.11(ii) below). As it dominates f it is cofinal into β.

(ii) Let $\gamma = \mathrm{cf}(\mathrm{cf}(\beta))$. Then $\gamma \leq \mathrm{cf}(\beta)$. However if $\gamma < \mathrm{cf}(\beta)$ and f, g are chosen so that $f : \gamma \longrightarrow \mathrm{cf}(\beta)$, $g : \mathrm{cf}(\beta) \longrightarrow \beta$ are both strictly increasing and cofinal, then their composition $g \circ f : \gamma \longrightarrow \beta$ cofinally, contradicting the definition of $\mathrm{cf}(\beta)$. Hence $\gamma = \mathrm{cf}(\beta)$.

(iii) Exercise. Q.E.D.

Corollary 2.8 *If $\mathrm{Lim}(\lambda)$ then $\mathrm{cf}(\omega_\lambda) = \mathrm{cf}(\lambda)$.*

Exercise 2.1 Prove (iii) of Lemma 2.7 and the corollary following.

Exercise 2.2 If $\mathrm{Lim}(\beta)$ show that for any $\alpha > 0$, $\mathrm{cf}(\alpha \cdot \beta) = \mathrm{cf}(\alpha + \beta) = \mathrm{cf}(\beta)$.

The following gives an alternative characterisation of cofinality for cardinals.

Lemma 2.9 *For any infinite cardinal β $\mathrm{cf}(\beta)$ is the least ordinal γ so that there is a sequence $\langle X_\tau \mid \tau < \gamma \rangle$ with each $X_\tau \subseteq \beta \wedge |X_\tau| < \beta$ and $\bigcup_{\tau<\gamma} X_\tau = \beta$.*

Proof: Let γ be the least such ordinal defined in the lemma. Then for some cofinal function $h : \mathrm{cf}(\beta) \to \beta$, we have $\beta = \bigcup_{\tau<\mathrm{cf}(\beta)} h(\tau)$. So $\gamma \leq \mathrm{cf}(\beta)$. So suppose for a contradiction that $\gamma < \mathrm{cf}(\beta)$, and we have $\bigcup_{\tau<\gamma} X_\tau = \beta$, with each $X_\tau \subseteq \beta \wedge |X_\tau| < \beta$. Define $f(\tau) = |X_\tau| < \beta$. As $\mathrm{dom}(f) = \gamma < \mathrm{cf}(\beta)$ we have $\mathrm{ran}(f)$ is bounded by some $\delta < \beta$. Let $g_\tau : X_\tau \leftrightarrow |X_\tau|$ be a bijection. Define $G(\xi) = \langle \tau, g_\tau(\xi) \rangle$ where τ is least so that $\xi \in X_\tau$. Then $G : \beta \to \gamma \times \delta$ is (1-1). But then $|\beta| \leq |\gamma \times \delta| = \max\{|\gamma|, |\delta|\} < \beta$. Contradiction! Q.E.D.

Exercise 2.3 (This exercise uses the definition of $h(\kappa)$ from Exercise 2.40.) Suppose κ is a singular cardinal. Show that $|h(\kappa)| = |\mathcal{P}(\kappa)|$. Calculate $\rho(h(\kappa))$.

2.1.2 Normal Functions and closed and unbounded classes

For the rest of this section we let Ω denote a regular, uncountable cardinal.

Definition 2.10 *Let A be a term and suppose $A \subseteq \Omega$.*
(i) Then A is closed *if $\forall \mu < \Omega\, (A \cap \mu$ is unbounded in $\mu \longrightarrow \mu \in A)$.*
(ii) We say A is c.u.b. in Ω *if it is both closed and unbounded in Ω.*

Note: In clause (i) we deliberately do not require Ω to be in A if the latter is unbounded in Ω. Closure is equivalent to requiring that (ii)$'$: *for any* $x \in V$ *if* $x \subseteq A$ *then* $\sup x \in A \cup \{\Omega\}$. (Exercise: Check this equivalence.)

Examples (i) The cofinal maps from the *Examples* of the last subsection are all closed and cofinal, although the first three which were maps just from ω cofinally into their range are rather trivially closed. The function g in the proof of (iii) of Lemma 2.7 was deliberately constructed to have range closed and unbounded in β - closure was obtained by taking for limit ordinals λ, $g(\lambda)$ to be the supremum of $g``\lambda$.

(ii) The class terms $\mathrm{Lim} =_{\mathrm{df}} \{\alpha \in \mathrm{On} \mid \alpha \text{ a limit ordinal}\}$, Card, LimCard, are all c.u.b. in On.

Definition 2.11 *(Normal Function). Let* $f : \Omega \longrightarrow \Omega$ *. Then* f *is* normal *if*

(i) $\alpha < \beta \longrightarrow f(\alpha) < f(\beta)$ *;*

(ii) (continuity) $\mathrm{Lim}(\lambda) \longrightarrow f(\lambda) = \sup\{f(\alpha) \mid \alpha < \lambda\}$.

Property (ii) says that f is *continuous*. Normal functions are quite common: all the ordinal arithmetic operations yield normal functions: $A_\alpha(\xi) = \alpha + \xi; M_\alpha(\xi) = \alpha \cdot \xi,\ E_\alpha(\xi) = \alpha^\xi$ are all normal functions. The $\aleph$-function which enumerates the cardinals is normal by design.

Exercise 2.4 Let $\omega \leq \kappa \in \mathrm{Reg}$. Define by induction on $\alpha < \kappa$ a function $f : \kappa \longrightarrow \kappa$, by $f(0) = 0; f(\beta + 1) = f(\beta) + \beta$ and $Lim(\lambda) \to f(\lambda) = \sup\{f(\beta) \mid \beta < \lambda\}$. Then check that f is indeed defined for all $\alpha < \kappa$ and that f is normal. Use f to define a *partition* of κ into κ many disjoint sets of cardinality κ by setting $D_\gamma = \{f(\beta) + \gamma \mid \beta > \gamma\}$. Check that $D_\gamma \cap D_{\gamma'} = \varnothing$ for $\gamma \neq \gamma' < \kappa$; and that $\bigcup_{\gamma<\kappa} D_\gamma = \kappa$.

Lemma 2.12 (Veblen 1908)

(i) Let $A \subseteq \Omega$*. Then* A *is c.u.b. in* Ω *iff the enumerating function for* A*,* f_A*, is normal with* $\mathrm{dom}(f_A) = \Omega$*;*

(ii) let $f : \Omega \longrightarrow \Omega$ *be strictly increasing. Then* f *is normal iff* $\mathrm{ran}(f)$ *is c.u.b. in* Ω.

Proof: (i) Let $f = f_A$. ($\Leftarrow$) As $\mathrm{dom}(f) = \Omega$, and f is (1-1), $ran(f)$ cannot be bounded in the cardinal Ω. So A is unbounded in Ω. The continuity of f translates directly into the closure of A: suppose $\mu < \Omega$ and $A \cap \mu$ is unbounded in μ. Let $\delta < \Omega$ be such that $f \restriction \delta$ enumerates $A \cap \mu$; then we have that $\mathrm{Lim}(\delta)$ (as $\mathrm{Lim}(\mu)$) and by continuity of f, $f(\delta) = \sup f``\delta = \mu$ and so μ must be in A.

($\Rightarrow$) Clearly f is a monotone increasing function: $\alpha < \beta < \Omega \longrightarrow f(\alpha) < f(\beta)$. As A is closed, then f_A will be also *continuous*: if $\lambda \in \Omega$ is a limit then $A \cap f``\lambda$ is unbounded in $\sup f``\lambda$. So by closure the latter is in A and is then $f(\lambda)$. Note now that $\mathrm{dom}(f)$ must be Ω since otherwise it is some $\beta < \Omega$ and f would witness that $\mathrm{cf}(\Omega) \leq \beta$. However Ω was assumed regular.

(ii) This just follows from (i), as f clearly is the enumerating function of $A = ran(f)$. See the Exercise below. Q.E.D.

Exercise 2.5 Let $f : \Omega \longrightarrow \Omega$ be strictly increasing. Then f is normal iff $\mathrm{ran}(f)$ is c.u.b. in Ω.

Lemma 2.13 *Let* $C \subseteq \Omega$ *be c.u.b. in* Ω*. Let* f_C *be the enumerating function of* C*. Then the class of* fixed points *of* $f_C : D =_{\mathrm{df}} \{\alpha < \Omega \mid f_C(\alpha) = \alpha\}$ *is c.u.b. in* Ω*. Hence for any normal function* $f : \Omega \to \Omega$ *there is a c.u.b. class of points* $\alpha < \Omega$ *that are fixed points for* f*:* $f(\alpha) = \alpha$.

Proof: Again let $f = f_C$. Let $\gamma \in \Omega$ be arbitrary. We find a member of D above γ (this shows that D is unbounded in Ω). Define: $\gamma_0 = \gamma; \gamma_{n+1} = f(\gamma_n); \gamma_\omega = \sup(\{\gamma_n \mid n < \omega\})$. Note that $\gamma_\omega \neq \Omega$. This is clear by the assumption of Ω's regularity. We claim that $\gamma_\omega \in D$. Let $\eta < \gamma_\omega$. Then for some n $\eta < \gamma_n < \gamma_\omega$. Hence $f(\eta) < f(\gamma_n) = \gamma_{n+1} < \gamma_\omega$. Hence $f``\gamma_\omega \subseteq \gamma_\omega$. As f is continuous, $f(\gamma_\omega) = \gamma_\omega \in D$. We are left only with showing that D is closed in Ω. Let $\mu < \Omega$ with $D \cap \mu$ unbounded in μ. Similar to showing the closure of γ_ω under f above, we have that $f``\mu \subseteq \mu$ (as any $\eta < \mu$ is less than some fixed point $\gamma < \mu$), and again by continuity $f(\mu) = \mu$. The last sentence is immediate as $\operatorname{ran}(f)$ is c.u.b. in Ω. Q.E.D.

Definition 2.14 *For any $E \subseteq$ On define E^* to be the class of* limit points *of E: namely those limit ordinals β such that $\beta \cap E$ is unbounded in β.*

Exercise 2.6 For any $E \subseteq$ On, show that E^* is a closed class, and if $E \in V$ with $\operatorname{cf}(\sup(E)) > \omega$, then E^* is c.u.b. below $\sup(E)$.

Exercise 2.7 Suppose $\Omega \in$ Reg. (i) Let $C, D \subseteq \Omega$ be c.u.b.in Ω. Show that $C \cap D$ is c.u.b. in Ω. (ii) Now generalise this argument: let $\gamma < \Omega$. Let $\langle C_\xi \mid \xi < \gamma \rangle$ be a sequence of c.u.b.in Ω classes. Show that $\bigcap_{\xi<\gamma} C_\xi$ is c.u.b. in Ω.

Remark 2.15 We used the letter Ω in this subsection rather than a generic mid-alphabet letter such as κ for a cardinal (our usual convention) since it is possible to construe the results here as also holding when Ω is interpreted as the class term On. To this extent On behaves like a 'regular cardinal', and we can interpret many results here as holding about terms $a \subseteq On$ which are not necessarily sets. One should be a little more careful than we have, when talking about sequences of classes if we allow $\Omega =$ On. In this case to define a sequence of classes $\langle C_\xi \mid \xi < \gamma \rangle$ with $C_\xi \subseteq$ On, we should speak about a single class term c of ordered pairs $\langle \xi, \zeta \rangle$ with $C_\xi \subseteq$ On being defined as the class $\{\zeta \mid \langle \xi, \zeta \rangle \in c\}$. With care this is unambiguous and proper, and can be done with $\gamma = On$ also. We could do the same in the following exercise, but have chosen not to, and have returned to our assumption that Ω as a regular cardinal.

Exercise 2.8 (Diagonal Intersections) Let $\Omega \in$ Reg. Let $\langle E_\xi \mid \xi < \Omega \rangle$ be a sequence of subsets of Ω. Define the *diagonal intersection* of the sequence to be the set $D = \Delta_{\xi<\Omega}\langle E_\xi \mid \xi < \Omega \rangle =_{\mathrm{df}} \{\alpha < \Omega \mid \forall \beta < \alpha(\alpha \in E_\beta)\}$. Now suppose that the E_ξ are all c.u.b. in Ω. (i) Show that the diagonal intersection D is c.u.b. in Ω. (ii) Show that $D = \bigcap_{\alpha<\Omega}(E_\alpha \cup (\alpha + 1))$.

Definition 2.16 *The $\beth$ (beth) function is defined by:*

$$\beth_0 = \omega_0; \qquad \beth_{\alpha+1} = 2^{\beth_\alpha}; \qquad \beth_\lambda = \sup\{\beth_\alpha \mid \alpha < \lambda\} \text{ if } \operatorname{Lim}(\lambda).$$

Thus $\alpha \to \beth_\alpha$ is a normal function, and has a range which, as always, is c.u.b. in On. By the last lemma it has a c.u.b. in Ω class of *fixed points* α so that $\alpha = \beth_\alpha$.

Exercise 2.9 Show that $\forall \alpha(|V_{\omega+\alpha}| = \beth_\alpha)$.

Exercise 2.10 (i) Check that the GCH (Generalised Continuum Hypothesis: that $\forall \alpha(2^{\aleph_\alpha} = \aleph_{\alpha+1})$) implies that $\forall \alpha(\aleph_\alpha = \beth_\alpha)$. (ii) Show that the first fixed point of the $\beth$ function has cofinality ω. (iii) Show that for any regular cardinal κ there is α, a fixed point of the $\beth$ function, with $\operatorname{cf}(\alpha) = \kappa$.

Definition 2.17 *(The c.u.b. filter on κ, F_κ) Let $\kappa > \omega$ be regular; let*

$$X \in F_\kappa \longleftrightarrow \exists C \subseteq \kappa(C \text{ is c.u.b.} \wedge C \subseteq X).$$

Exercise 2.11 Show that F_κ has the following properties:
(i) $X \in F \wedge Y \supseteq X \longrightarrow Y \in F$
(ii) $X, Y \in F \longrightarrow X \cap Y \in F$
(iii) $\forall \xi < \kappa \ \{\xi\} \notin F$
(iv) $\forall \xi < \kappa \forall \langle X_\zeta \mid \zeta < \xi \rangle [\forall \zeta (X_\zeta \in F) \longrightarrow \bigcap_{\zeta<\xi} X_\zeta \in F]$
(v) $\forall \langle X_\zeta \mid \zeta < \kappa \rangle [\forall \zeta (X_\zeta \in F) \longrightarrow \Delta_{\zeta<\kappa} X_\zeta \in F]$.

A non-empty collection F of subsets of κ satisfying (i) and (ii) is called a *filter* on κ; property (iii) states that the filter is *non-principal*; (iv) states that the filter is κ*-complete*; a filter closed under taking diagonal intersections (see Exercise 2.8) in (v) is called *normal*. Not listed is the obvious fact about F_κ that it is *non-trivial:* $\varnothing \notin F$. A filter is called an *ultrafilter* if for every $X \subseteq \kappa$ either X or $\kappa \backslash X$ is in F. The existence of ultrafilters on subsets of a $\kappa > \omega$ satisfying additionally (iii) and (iv) cannot be proven in ZFC, (they can for $\kappa = \omega$) but is crucial for studying many *consistency results* in forcing theory, and for considering *elementary embeddings* of the universe V to transitive subclasses of V. A class of subsets of κ on which there is an ultrafilter satisfying (i)-(iv) is often said, in an equivalent terminology, to have a 2-*valued measure*, in which case property (iv) is called "κ-additivity". Sets have value 0/1 depending on whether they are out/in the ultrafilter. (iii) then translates as "*points have measure* 0".

2.1.3 Stationary Sets

Let Ω denote either an uncountable regular cardinal κ, or else On the class of ordinals.

Definition 2.18 *Let $E \subseteq \Omega$. Then E is called* stationary *in Ω if for every $C \subseteq \Omega$ which is c.u.b. in Ω, then $E \cap C \neq \varnothing$.*

If we were to talk about a class term $S \subseteq \Omega$ being stationary where $\Omega = On$, we should declare more precisely what this means: it means that for any class term c which is a closed and unbounded class of ordinals, then we can also prove that $c \cap S \neq \varnothing$.

Stationary subsets of regular cardinals (or subclasses of On) exist: any c.u.b. subset of κ with κ regular is stationary, by Exercise 2.7 (i). (Similarly for subclasses of On). But there are other stationary subsets of regular cardinals.

Exercise 2.12 Let $S \subseteq \Omega$ be stationary and $C \subseteq \Omega$ be c.u.b. Then $S \cap C$ is stationary.

Exercise 2.13 Let $S \subseteq \Omega$ be stationary. Show that $S \cap S^*$ is stationary.

Example 1 Let $\Omega = \omega_2$. Then $S_\omega =_{\mathrm{df}} \{\alpha < \omega_2 \mid \mathrm{cf}(\alpha) = \omega\}$ and $S_{\omega_1} =_{\mathrm{df}} \{\alpha < \omega_2 \mid \mathrm{cf}(\alpha) = \omega_1\}$ are two disjoint stationary subsets of ω_2: let $C \subseteq \omega_2$ be any c.u.b. subset. Let $f : \omega_2 \longrightarrow C$ be its strictly increasing enumerating function. Then $f(\omega) \in C \cap S_\omega$ and $f(\omega_1) \in C \cap S_{\omega_1}$.

Exercise 2.14 Can you generalise this example to larger regular cardinals, *e.g.* ω_n for $n < \omega$, or any regular $\kappa > \omega_2$?

Exercise 2.15 Find $S_n \subseteq \aleph_{\omega+1}$ stationary, for $n < \omega$, with $S_{n+1} \subseteq S_n$ but with $\bigcap_n S_n = \varnothing$.

The reason for the nomenclature comes from (ii) of the following Lemma.

Lemma 2.19 (Fodor's Lemma 1956) *Let $\kappa > \omega$ be a regular cardinal. The following are equivalent.*

(i) S is stationary in κ ;

(ii) For every function $f : S \longrightarrow$ On *which is* regressive*, that is $\forall \alpha \in S(\alpha > 0 \longrightarrow f(\alpha) < \alpha)$, there is a stationary set $S_0 \subseteq S$ and a fixed α_0 so that $\forall \xi \in S_0(f(\xi) = \alpha_0)$.*

Proof: Assume (i). If (ii) failed for some regressive function f then we should be able to define for every $\alpha < \kappa$ a c.u.b. $C_\alpha \subseteq \kappa$ with $\xi \in C_\alpha \cap S \longrightarrow f(\xi) \neq \alpha$. Let $D = \{\alpha \mid \forall \beta < \alpha(\alpha \in C_\beta)\}$ be the diagonal intersection of $\langle C_\alpha \mid \alpha < \kappa \rangle$. Then D is c.u.b. in κ and for any $\xi \in D \cap S$, $f(\xi) \not< \xi$. But if $\xi \in D \cap S$ we must have $f(\xi) < \xi$, which is a contradiction. (ii) implies (i) is trivial. Q.E.D.

Remark: AC was used heavily in picking the C_α in the above; if one attempts the proof without using AC one obtains in (ii) only the conclusion that for some $\alpha_0 < \kappa$ that $f^{-1}``\alpha_0$ is unbounded in κ. Because one cannot in general pick class terms, if one attempts to prove the Lemma for stationary classes and regressive functions on all of On, rather than just κ, one again weakens the conclusion (see the next Exercise).

Exercise 2.16 (E) Let f be a function class term with $\mathrm{dom}(f)$ = On and f regressive. Show that for some α_0 $f^{-1}``\{\alpha_0\}$ is unbounded in On. [Hint: Suppose the conclusion fails; then define $g(\xi) = \sup f^{-1}``\{\xi\}$; now find α_0 closed under g: $g``\alpha_0 \subseteq \alpha_0$.]

We could have defined stationary subsets of ordinals β with $\mathrm{cf}(\beta) > \omega$. This is possible, but notice that it would make no sense to define the notion of a stationary subset β if $\mathrm{cf}(\beta) = \omega$. For, if $f : \omega \longrightarrow \beta$ is a strictly increasing function cofinal in β then $\mathrm{ran}(f)$ is c.u.b. in β; but it is easy to define another c.u.b. in β set C (of order type ω) with $\mathrm{ran}(f) \cap C = \varnothing$ so it makes little sense to even try to define stationary in this way.

We saw above that ω_2 contained two disjoint stationary subsets. In fact far more is true. (The proof of this theorem is omitted.) Any stationary set in a regular κ can be split into κ many disjoint sets which are still stationary.

Theorem 2.20 *(Bloch (1953), Fodor (1966), Solovay (1971)) Let $\kappa > \omega$ be regular, and let $S \subseteq \kappa$ be stationary. Then there is a sequence of κ many* disjoint *stationary sets $S_\xi \subseteq S$ for $\xi < \kappa$ (i.e. for $\zeta < \xi < \kappa$ $S_\xi \cap S_\zeta = \varnothing$) with $S = \bigcup_{\xi<\kappa} S_\xi$.*

Exercise 2.17 $(*)$(E) (H.Friedman) Let $S \subseteq \omega_1$ be stationary. Then for any $\alpha < \omega_1$ there is a closed subset $C_\alpha \subseteq S$ with $\mathrm{ot}(C_\alpha) = \alpha + 1$. [Hint: Do this by induction on α for any stationary S. This is trivial for $\alpha = \beta + 1$ assuming it is true for β (just add one more point $\tau \in S$ above $\sup(C_\beta)$ to C_β to get $C_{\beta+1}$ of order type $\alpha+1$). Assume $\mathrm{Lim}(\alpha)$ and for $\beta < \alpha$ we can find such C_β. Note that for any δ we can find such C_β with $\min(C_\beta) \geq \delta$ - by considering the stationary $S \backslash \delta$. Let $\langle \alpha_n \mid n < \omega \rangle$ be chosen with $\sup_n \alpha_n = \alpha$; for any δ then pick closed subsets $C_{\alpha_n} \subseteq S$ of order type $\alpha_n + 1$ and with $\min(C_{\alpha_{n+1}}) > \sup(C_{\alpha_n})$. Then $\bigcup_n C_{\alpha_n} \subseteq S$ and is closed in S with the exception of the point $\sup(\bigcup_n C_{\alpha_n})$. Call a point arrived at as a sup of such a sequence of sets C_{α_n} an "exceptional" point. We have just shown that the exceptional points are unbounded in ω_1. But now just note that a limit of exceptional points is also exceptional. That is, they form closed subset of ω_1. As S is stationary there is an exceptional point $\sigma \in S$. This σ can be added to the top of the sequence of points from the sets C'_{α_n} witnessing the exceptionality of σ; this sequence then has order type $\alpha + 1$ and is contained in S.

Remark: This is not the case at higher cardinals, *e.g.* ω_2. Let $(\star)$ be the statement "*for any $X \subseteq \omega_2$ and any $\alpha < \omega_2$ either X or $\omega_2 \backslash X$ contains a closed set C with* $\mathrm{ot}(C) = \alpha$". Then ZFC $\nvdash (\star)$.

2.2 Some further cardinal arithmetic

We give some further results on cardinal arithmetic.

Definition 2.21 *Let $\langle \kappa_\alpha \mid \alpha < \tau\rangle$ be a sequence of cardinal numbers. Let $\langle X_\alpha \mid \alpha < \tau\rangle$ be a sequence of disjoint sets, with $\kappa_\alpha = |X_\alpha|$. (i) Then we define the* cardinal sum*:*

$$\sum_{\alpha<\tau} \kappa_\alpha = |\bigcup_{\alpha<\tau} X_\alpha|.$$

(ii) The cardinal product *is defined as* $\prod_{\alpha<\tau}\kappa_\alpha = |\prod_{\alpha<\tau} X_\alpha|$.

Note: (i) as usual these values are independent of the choices of the X_α with the stipulated cardinalities. For the product the requirement that the sets X_α be disjoint may be dropped. Here this is an accord with Definition 1.9 where f is the function so that $f(\alpha) = X_\alpha$ for $\alpha < \tau$.

(ii) If all the $\kappa_\alpha = \lambda \geq \omega$ for some fixed λ, and $\tau \in$ Card, then $\sum_{\alpha<\tau}\kappa_\alpha = \tau \otimes \lambda$ and $\prod_{\alpha<\tau}\kappa_\alpha = \lambda^\tau$.

Exercise 2.18 Show that If $\omega \leq \tau \in$ *Card* and every $\kappa_\alpha \neq 0$, then $\Sigma_{\alpha<\tau}\kappa_\alpha = \tau \otimes \sup_{\alpha<\tau}\kappa_\alpha$.

Exercise 2.19 Show that $\prod_{\alpha<\tau}\kappa_\alpha^\lambda = (\prod_{\alpha<\tau}\kappa_\alpha)^\lambda$ and $\prod_{\alpha<\tau}\kappa^{\lambda_\alpha} = \kappa^{\Sigma_{\alpha<\tau}\lambda_\alpha}$.

Exercise 2.20 Show that if $\kappa_\alpha \geq 2$ for $\alpha < \tau$, then $\sum_{\alpha<\tau}\kappa_\alpha \leq \prod_{\alpha<\tau}\kappa_\alpha$.

Exercise 2.21 Show that $\prod$ distributes over $\sum$, *i.e.* that $\prod_{\alpha<\tau}\sum_{\beta<\mu}\kappa_{\alpha,\beta} = \sum_{f\in{}^\tau\mu}\prod\kappa_{\alpha,f(\alpha)}$.

Lemma 2.22 *If $\omega \leq \tau \in$ Card and $\langle \kappa_\alpha \mid \alpha < \tau\rangle$ is a non-decreasing sequence of non-zero cardinals, then* $\prod_{\alpha<\tau}\kappa_\alpha = (\sup_{\alpha<\tau}\kappa_\alpha)^\tau$.

Proof: We partition τ into τ many disjoint pieces each of size τ (by using some bijection $\pi : \tau\times\tau \leftrightarrow \tau$). Let us say then that $\tau = \bigcup_{\beta<\tau} X_\beta$. Because the sequence of the κ_α is non-decreasing, and each X_β is unbounded in τ, we still have $\sup_{\alpha\in X_\beta}\kappa_\alpha = \sup_{\alpha<\tau}\kappa_\alpha = \kappa$ say, for each $\beta < \tau$. Now note that we may reorganise the product

$$\prod_{\alpha<\tau}\kappa_\alpha \text{ as } \prod_{\beta<\tau}\left(\prod_{\alpha\in X_\beta}\kappa_\alpha\right).$$

But $\prod_{\alpha\in X_\beta}\kappa_\alpha \geq \sup_{\alpha\in X_\alpha}\kappa_\alpha = \kappa$, hence we have that $\prod_{\alpha<\tau}\kappa_\alpha \geq \prod_{\beta<\tau}\kappa = \kappa^\tau$.

Conversely $\prod_{\alpha<\tau}\kappa_\alpha \leq \prod_{\alpha<\tau}\kappa = \kappa^\tau$. Hence we have equality as desired. Q.E.D.

Exercise 2.22 $\prod_{n<\omega} n = \prod_{n<\omega} n^{\omega_0} = \omega_0^{\omega_0} = 2^{\omega_0}$; $\prod_{n<\omega}\omega_n^{\omega_0} = (\omega_{\omega_0})^{\omega_0}$;

Theorem 2.23 (**König's Theorem**) *If $\kappa_\alpha < \lambda_\alpha$ for $\alpha < \tau$ then*

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

Proof: Pick X_α for $\alpha < \tau$ with $|X_\alpha| = \lambda_\alpha$. We shall show that if $Y_\alpha \subseteq \prod_{\alpha<\tau} X_\alpha$ for $\alpha < \tau$ are such that $|Y_\alpha| \leq \kappa_\alpha$, that then $\bigcup_{\alpha<\tau} Y_\alpha \neq \prod_{\alpha<\tau} X_\alpha$. Hence we cannot have $\sum_{\alpha<\tau}\kappa_\alpha \geq \prod_{\alpha<\tau}\lambda_\alpha$. Let $P_\alpha = \{f(\alpha) \mid f \in Y_\alpha\}$ be the projection of Y_α on to the α'th coordinate. As $|Y_\alpha| < |X_\alpha|$, $|P_\alpha| < |X_\alpha|$ but $P_\alpha \subset X_\alpha$. So let $f \in \prod_{\alpha<\tau} X_\alpha$ be any function so that for any $\alpha < \tau$ $f(\alpha) \notin P_\alpha$. Then f cannot be in any Y_α. Thus $\bigcup_{\alpha<\tau} Y_\alpha \neq \prod_{\alpha<\tau} X_\alpha$ as we sought. Q.E.D.

Exercise 2.23 Deduce Cantor's Theorem that $\kappa < 2^\kappa$ from König's Theorem.

Corollary 2.24 *For all β, and for all α $\mathrm{cf}(\omega_\beta^{\omega_\alpha}) > \omega_\alpha$. Hence in particular $\mathrm{cf}(2^\kappa) > \kappa$ for any cardinal κ.*

Proof: Let κ_τ be a sequence of cardinals for $\tau < \omega_\alpha$ with $\kappa_\tau < \omega_\beta^{\omega_\alpha}$. It suffices to show that $\sum_{\tau<\omega_\alpha} \kappa_\tau < \omega_\beta^{\omega_\alpha}$. Let λ_τ be the fixed sequence with all $\lambda_\tau = \omega_\beta^{\omega_\alpha}$, for $\tau < \omega_\alpha$. By König's Lemma then

$$\sum_{\tau<\omega_\alpha} \kappa_\tau < \prod_{\tau<\omega_\alpha} \lambda_\alpha = (\omega_\beta^{\omega_\alpha})^{\omega_\alpha} = \omega_\beta^{\omega_\alpha}.$$

Q.E.D.

Corollary 2.25 $\kappa^{\mathrm{cf}(\kappa)} > \kappa$ *for any cardinal $\kappa \geq \omega$.*

Proof: For $\alpha < \mathrm{cf}(\kappa)$ let κ_α be less than κ so that $\kappa = \sum_{\alpha<\mathrm{cf}(\kappa)} \kappa_\alpha$. Then
$\kappa = \sum_{\alpha<\mathrm{cf}(\kappa)} \kappa_\alpha < \prod_{\alpha<\mathrm{cf}(\kappa)} \kappa = \kappa^{\mathrm{cf}(\kappa)}$. Q.E.D.

We may put some of these facts together to get some more information about the exponentiation function under GCH. First:

Exercise 2.24 If $\lambda < \mathrm{cf}(\kappa)$ then ${}^\lambda\kappa = \bigcup_{\alpha<\kappa} {}^\lambda\alpha = \bigcup_{\mathrm{cf}(\kappa)<\alpha<\kappa} {}^\lambda\alpha$.

Theorem 2.26 *Suppose* GCH *holds and $\kappa, \lambda \geq \omega$. Then κ^λ takes the following values:*
(i) λ^+ *if $\kappa \leq \lambda$;*
(ii) κ^+ *if $\mathrm{cf}(\kappa) \leq \lambda < \kappa$;*
(iii) κ *if $\lambda < \mathrm{cf}(\kappa)$.*

Proof: (i) follows from $\kappa^\lambda = 2^\lambda = \lambda^+$. (ii) $\kappa < \kappa^{\mathrm{cf}(\kappa)} \leq \kappa^\lambda \leq \kappa^\kappa = 2^\kappa = \kappa^+$; (iii) We use Ex.2.24. $\kappa^\lambda = |\bigcup_{\alpha<\kappa} {}^\lambda\alpha|$. But for $\alpha < \kappa$, $|{}^\lambda\alpha| \leq |{}^\alpha\alpha| = 2^{|\alpha|} = |\alpha|^+ < \kappa$. So $\kappa \leq \kappa^\lambda \leq \kappa \otimes \sup_{\alpha<\kappa} |\alpha|^+ = \kappa$. Q.E.D.

Without GCH the only known constraints on the exponentiation function for regular cardinals κ are (a) $\kappa < 2^\kappa$ and (b) $\kappa < \lambda \rightarrow 2^\kappa \leq 2^\lambda$. For singular κ the situation is more subtle and a discussion of this involves large cardinals.

Exercise 2.25 Prove that $\beth_\omega^{\aleph_0} = \Pi_n \beth_n = \beth_{\omega+1}$. [Hint: Every subset of $\beth_\omega$ can be coded as a function $\omega \rightarrow \beth_\omega$.]

Exercise 2.26 Assume *CH* but not *GCH*. Show that $(\aleph_n)^{\aleph_0} = \aleph_n$ for $1 \leq n < \omega$.

2.2.1 The Singular Cardinals Hypothesis

Without the assumption of the *GCH*, the behaviour of the exponention function at regular κ, or more simply put, the value 2^κ, is more or less independent of the values of 2^λ for regular $\lambda < \kappa$ apart from the monotonicity requirement that $\lambda < \kappa \rightarrow 2^\lambda \leq 2^\kappa$ and the additional basic constraint following on from Cantor's theorem, that $2^\kappa > \kappa$. However for singular κ this is not the case, at least for those κ with uncountable cofinality. One can show that the value of 2^κ is dependent on the value of 2^λ for a stationary set of cardinals $\lambda < \kappa$. To quote an example: if on a stationary set of $\lambda < \kappa$, we have $2^\lambda = \lambda^{++}$ then 2^κ must be κ^{++}. The value κ^{++} was just an example here: we could have written λ^+ or $\lambda^{+\cdots+}$ for a fixed row of n +'s. Then the value of 2^κ would be κ^+ or $\kappa^{+\cdots+}$ respectively. However this picture is entirely dependent on the assumption that $\mathrm{cf}(\kappa) > \omega$. For κ of cofinality ω the picture is more subtle.

The Singular Cardinals Hypothesis, SCH asserts that for all singular cardinals κ $\kappa^{\mathrm{cf}(\kappa)} = 2^{\mathrm{cf}(\kappa)} \otimes \kappa^+$.

Notice that this latter equality is always true for regular $\kappa > \omega$, as $\kappa^{\mathrm{cf}(\kappa)} = \kappa^\kappa = 2^\kappa = 2^\kappa \oplus \kappa^+ = 2^{\mathrm{cf}(\kappa)} \oplus \kappa^+$. But for any κ we have $\kappa^{\mathrm{cf}(\kappa)} \geq \kappa^+$ by Corollary 2.25 and trivially $\kappa^{\mathrm{cf}(\kappa)} \geq 2^{\mathrm{cf}(\kappa)}$. Hence the SCH is asserting that the value of $\kappa^{\mathrm{cf}(\kappa)}$ is the minimum it could be.

Lemma 2.27 *The GCH implies the SCH.*

Proof: Note that we can identify any function $f \in \kappa^{\mathrm{cf}(\kappa)}$ *via* a bijective pairing function $\pi : \kappa \times \kappa \leftrightarrow \kappa$ as itself a subset of κ, hence $\kappa^{\mathrm{cf}(\kappa)} \leq 2^\kappa$. Now let κ be a singular limit cardinal. Now assume GCH, then $2^{<\kappa} = \kappa$. But if we fix a cofinal function $f : \mathrm{cf}(\kappa) \to \kappa$ then for any $X \subseteq \kappa$, we have $X = \bigcup_{\alpha < \mathrm{cf}(\kappa)} X \cap f(\alpha)$. However for each such α, $|\mathcal{P}(f(\alpha))| < \kappa$ and so bijections between such $\mathcal{P}(f(\alpha))$ and ordinals less than κ. So we have a (1-1) map $g : \mathcal{P}(\kappa) \to \kappa^{\mathrm{cf}(\kappa)}$. Hence $2^\kappa \leq \kappa^{\mathrm{cf}(\kappa)}$. The SCH then follows as the above shows $\kappa^{\mathrm{cf}(\kappa)} = 2^\kappa = \kappa^+ = 2^{\mathrm{cf}(\kappa)} \oplus \kappa^+$. Q.E.D.

We've noted that the SCH implies that $\kappa^{\mathrm{cf}(\kappa)}$ is the least possible value. The following summarises exponentiation under this assumption.

Lemma 2.28 *Assume the SCH. Then:*
(1) if $\kappa \in SingCard$ then:
(a) if the exponentiation function 2^λ is eventually constant for $\lambda < \kappa$ then $2^\kappa = 2^{<\kappa}$;
(b) otherwise $2^\kappa = (2^{<\kappa})^+$;
(2) for $\omega \leq \kappa, \lambda \in Card$ then:
(a) if $\kappa \leq 2^\lambda$ then $\kappa^\lambda = 2^\lambda$;
(b) if $2^\lambda < \kappa$ and $\lambda < \mathrm{cf}(\kappa)$ then $\kappa^\lambda = \kappa$;
(c) if $2^\lambda < \kappa$ and $\lambda \geq \mathrm{cf}(\kappa)$ then $\kappa^\lambda = \kappa^+$.

2.3 Transitive Models

We have seen how certain assumptions about a transitive set or class term allows us to conclude that a number of the ZF axioms hold, by relativisation to that set or term. When thinking of a term W as a structure, which we more properly write $\langle W, \in \rangle$, we say that $\langle W, \in \rangle$ is a *transitive model*, or *transitive $\in$ model* if we wish to emphasise the standard interpretation. We saw that in 1.24 and 1.25 that closure under those lists of conditions ensured that $(\mathrm{ZF}^-)^W$. The following Lemma allows us to create transitive isomorphic copies $\langle M, \in \rangle$ of possibly non-transitive structures $\langle H, \in \rangle$. It is known as the "Collapsing Lemma" since it collapses any "$\in$-holes" out of the structure $\langle H, \in \rangle$. The Lemma is much more general and in fact a structure $\langle H, R \rangle$ will be isomorphic to a transitive model $\langle M, \in \rangle$ provided that R satisfies two necessary conditions: that it be wellfounded, and that it be "extensional". The latter simply requires it to be $\in$-like. Clearly these conditions are necessary, since $\in$ is itself wellfounded, and for transitive M we always have that $(\mathrm{AxExt})^M$.

Definition 2.29 *Given a term t and a relation R on t we say that R is* extensional on t *iff for any $u, v \in t, u \neq v$ there is $z \in t$ with $zRu \longleftrightarrow \neg zRv$ (i.e. $\{z \in t \mid zRu\} \neq \{z \in t \mid zRv\}$).*

Note that $\in$ is extensional on x if $\mathrm{Trans}(x)$ but need not be in general.

Lemma 2.30 (Mostowski (1949)-Shepherdson (1951) **The Collapsing Lemma)**
Let $H \in V$.

(i) Suppose that R is wellfounded and extensional on H. Then there is a unique transitive term M and a unique collapsing isomorphism $\pi : \langle H, R\rangle \longrightarrow \langle M, \in\rangle$.

(ii) Additionally if $R \restriction x^2 = \in \restriction x^2$, $x \subseteq H$, and $\mathrm{Trans}(x)$, *then* $\pi \restriction x = \mathrm{id} \restriction x$.

Proof: (i) (1) *If π exists, then it is is unique.*

Proof: Suppose π, $M = \mathrm{ran}(\pi)$ are as supposed. Let $u, v \in H$. Note if uRv then $\pi(u) \in \pi(v)$ as π preserves the order relations. Thus for $v \in H$: $\{\pi(u) \mid u \in H \wedge uRv\} \subseteq \pi(v)$.

However if $z \in \pi(v)$, then $z \in M$, as M is transitive. Hence $z = \pi(u)$ for some $u \in H$ with uRv. Hence $\{\pi(u) \mid u \in H \wedge uRv\} \supseteq \pi(v)$. Thus $\pi(v) = \{\pi(u) \mid u \in H \wedge uRv\}$. Thus the isomorphism, if it exists *must* take this form.

(2) π *exists.*

We thus define by R-recursion: $\pi(v) = \{\pi(u) \mid u \in H \wedge uRv\} \quad (*)$
and take $M = \mathrm{ran}(\pi)$. Trivially $Trans(M)$ by $(*)$. (3)-(5) will show that π is an isomorphism.

(3) π *is* (1-1).

Proof: If not pick t $\in$-minimal in M so that there exist $u \neq v$ with $t = \pi(u) = \pi(v)$. As $u \neq v$, and R is extensional, there is some w with $wRu \leftrightarrow \neg wRv$. Without loss of generality we assume $wRu \wedge \neg wRv$. Then $\pi(w) \in \pi(u) = t = \pi(v)$. So we must have that for some xRv: $\pi(x) = \pi(w)$ (as $\pi(v)$ is the set of all such $\pi(x)$'s). But now if we set $s = \pi(x)$, we have $s \in t$ and $\pi(x) = \pi(w) = s$ and, as $\neg wRv$, $x \neq w$. However this s contradicts the $\in$-minimality in the choice of t.

(4) π *is onto.*

This is trivial as M is defined to be $\mathrm{ran}(\pi)$.

(5) π *is an order preserving isomorphism.*

We have already that π is a bijection. This then follows from the definition at $(*)$: $uRv \leftrightarrow \pi(u) \in \pi(v)$.

This finishes (i). For (ii) we now assume that $R \restriction x^2 = \in \restriction x^2, \mathrm{Trans}(x)$ and $x \subseteq H$.

(6) $\pi \restriction x = \mathrm{id} \restriction x$.

Then for $v \in x$ we have $v \subseteq x \subseteq H$. Thus $(*)$ becomes, for $v \in x$: $\pi(v) = \{\pi(u) \mid u \in v\}$. Now, by $\in$-induction on $\in \restriction x \times x$ we have $\forall v \in x[(\forall u \in v \to \pi(u) = u) \to \pi(v) = v] \to \forall v \in x(\pi(v) = v)$.

Q.E.D.

The resulting structure M is called the 'collapse', or better, the 'transitive collapse' of $\langle H, R\rangle$. To illustrate how the Collapsing Lemma works note the following exercise:

Exercise 2.27 Let $\langle H, R\rangle \in WO$. Apply the Collapsing Lemma. What is the outcome?

Note the use in the above proof of a recursion along the wellfounded relation R rather than $\in$. More generalised forms of this argument are possible. We may take any class term t in place of the set H and provided the wellfounded extensional relation R is *set-like* - meaning for any $u \in t$ $\{v \mid vRu\} \in V$, then the same argument may be used, and a class term M defined in the same way.

Lemma 2.31 (General Mostowski-Shepherdson Collapsing Lemma) *Let A be a class term.*

(i) Let $R \subseteq A \times A$ be a wellfounded extensional relation which is set-like in the above sense. Then there is a unique term M, and unique collapsing isomorphism $\pi : \langle A, R\rangle \longrightarrow \langle M, \in\rangle$.

(ii) If $R = \in$ then if s is a transitive term with $s \subseteq A$, then $\pi \upharpoonright s = \mathrm{id} \upharpoonright s$.

Exercise 2.28 Show that V_ω can be 'coded' as a subset of ω: that is there is $E \subseteq \omega$ so that $\langle \omega, E\rangle \cong \langle V_\omega, \in\rangle$. [Hint: Define $nEm \longleftrightarrow_{\mathrm{df}}$ the "2^n" column in the binary expansion of m contains a 1; (thus $\{n \mid nE11\} = \{0,1,3\}$); check there is u satisfying $\langle \omega, E\rangle \cong \langle u, \in\rangle$ with $\mathrm{Trans}(u)$. Show $u = V_\omega$.]

Exercise 2.29 Show if $\langle A, \in\rangle, \langle B, \in\rangle$ are transitive sets, and $f : \langle A, \in\rangle \cong \langle B, \in\rangle$ is an isomorphism, then $f = \mathrm{id} \upharpoonright A$.

Exercise 2.30 Suppose $\mathrm{Trans}(x)$ and $f : \kappa \leftrightarrow x$ is a bijection. Define $E \subseteq \kappa \times \kappa$ by: $\langle \alpha, \beta\rangle \in E \longleftrightarrow f(\alpha) \in f(\beta)$. Show that $\langle \kappa, E\rangle \cong \langle x, \in\rangle$ and that the isomorphism is the Mostowski-Shepherdson collapse map. Let $g : \kappa \times \kappa \leftrightarrow \kappa$ be a further bijection. Then if $\widetilde{E} = g``E$, we can then think of x as coded by a subset of κ, namely by $\widetilde{E}$. Note that x will have 2^κ-many such different codes depending on the function f.

Exercise 2.31 Find an example of an $\langle x, \in\rangle$ which is not extensional. If we nevertheless apply the Mostowski-Shepherdson Collapse function π to it, what happens?

2.4 The H_κ sets

The following collects together sets whose transitive closure is of a certain maximal size. The phrase "hereditarily of [property φ]" means that not only must an x have property φ, but so must all its members, and their members, and ... and so on. In other words all of $TC(x)$ must have property φ.

Definition 2.32 *Let κ be an infinite cardinal. Then $H_\kappa =_{\mathrm{df}} \{x \mid |\mathrm{TC}(x)| < \kappa\}$ is the class of sets* hereditarily of cardinality less than κ.

We summarise some properties of these classes.

Lemma 2.33 *Let κ be an infinite cardinal.*

(i) $\mathrm{On} \cap H_\kappa = \kappa$; $\mathrm{Trans}(H_\kappa)$;

(ii) $H_\kappa \subseteq V_\kappa$ and hence $H_\kappa \in V_{\kappa+1}$, $\rho(H_\kappa) = \kappa$;

(iii) $y \in H_\kappa \wedge x \subseteq y \longrightarrow x \in H_\kappa$;

(iv) $x, y \in H_\kappa \longrightarrow \bigcup x, \{x, y\} \in H_\kappa$;

(v) (AC) κ regular $\longrightarrow \forall x(x \in H_\kappa \leftrightarrow x \subseteq H_\kappa \wedge |x| < \kappa)$.

Andrzej Mostowski 1913-1975

PROOF: (i) Exercise; (ii): We use Ex.1.2: let $\theta = \rho``\mathrm{TC}(x)$, and if $x \in H_\kappa$ we have $|\mathrm{TC}(x)| < \kappa$; hence $\theta < \kappa$. and thus $x \in V_{\theta+1}$. Thus $H_\kappa \subseteq V_\kappa$, thence $H_\kappa \in V_{\kappa+1}$ and $\rho(H_\kappa) \leq \kappa$; as $\kappa \subseteq H_\kappa$, we have $\rho(H_\kappa) \geq \kappa$. (ii) is completed.
(v) $(\rightarrow)$ Assume $x \in H_\kappa$. As $\mathrm{Trans}(H_\kappa) \wedge x \subseteq \mathrm{TC}(x)$ this follows from the definition of H_κ. $(\leftarrow)$ As $\mathrm{TC}(x) = x \cup \bigcup\{\mathrm{TC}(y) \mid y \in x\}$, it is the union of less than κ many sets all of cardinality less than κ. By AC such a union has itself cardinality less than κ so we are done. Q.E.D.

EXERCISE 2.32 Prove (i), (iii)-(iv) here. Give an example to show that the conclusion of (v) fails if κ is singular.

LEMMA 2.34 *(AC)* $\kappa > \omega \wedge \kappa$ *regular* $\longrightarrow (\mathrm{ZFC}^-)^{H_\kappa}$. *More formally: let* $\vec{\varphi}$ *be a finite list of axioms from* ZFC^-. *Then* $\mathrm{ZFC} \vdash$ "$\kappa > \omega \wedge \kappa$ *regular* $\longrightarrow (\bigwedge\!\!\!\bigwedge \vec{\varphi})^{H_\kappa}$."

PROOF: We appeal to Lemma 1.26 once we have observed that Separation, Collection, and the Wellordering Principle axioms hold relativised to H_κ, the others follow from Lemma 2.33. $(\mathrm{AxSeparation})^{H_\kappa}$ holds since if a^{H_κ} is any term, and $x \in H_\kappa$ then y= $a^{H_\kappa} \cap x$ is a subset of x and hence it satisfies $|\mathrm{TC}(y)| < \kappa$ also. Similarly, for the Axiom of Collection, if $(r$ *is a relation* $\wedge \forall x r``x \neq \varnothing)^{H_\kappa}$ then let s be the *function* (defined in V) given by $sx = y \leftrightarrow (r(x,y) \wedge x, y \in H_\kappa \wedge \forall z \prec y \neg r(x,z)) \vee (x \notin H_\kappa \wedge y = \varnothing)$ where $\langle H_\kappa, \prec\rangle \in \mathrm{WO}$ for some wellorder $\prec$. Then letting $w \in H_\kappa$ be arbitrary, and applying Replacement (again in V) we deduce that $s``w \in V$. However $s``w \subseteq H_\kappa$ and has at most $|w| < \kappa$ many elements. Hence setting $t = s``w$ we have $t \in H_\kappa$ as required in the statement of **Ax6'**. For $(\mathrm{WP})^{H_\kappa}$ let $x \in H_\kappa$ and $\langle x, \prec\rangle \in WO$. Just check as $x \in H_\kappa$ that $\prec \restriction x \times x \in H_\kappa$. Q.E.D.

We remark also that the last lemma is false for singular cardinals κ.

2.4.1 H_ω - THE HEREDITARILY FINITE SETS

For $\kappa = \omega$ then H_κ is known as the class of the *hereditarily finite* sets - and is so also more usually abbreviated as HF.

EXERCISE 2.33 Show that $V_\omega = \mathrm{HF}$. [Hint: For ($\subseteq$) use induction on n to show $V_n \in \mathrm{HF}$. For ($\supseteq$) use $\in$-induction].

LEMMA 2.35 $(\mathrm{ZFC} - \mathrm{Ax.Inf} + \neg\,\mathrm{Ax.Inf})^{\mathrm{HF}}$

PROOF: See Exercise. Q.E.D.

EXERCISE 2.34 Check that HF is closed under all the assumptions of Lemmata 1.24 and 1.25 (except 1.24 (ii)) and even the power set operation. Hence $(\mathrm{ZFC} - \mathrm{Ax.Inf})^{\mathrm{HF}}$.

EXERCISE 2.35 (Ackermann 1937) Investigate the following function $f : \mathrm{HF} \to \omega$: $f(x) = \Sigma_{y \in x} 2^{f(y)}$.

2.4.2 H_{ω_1}- THE HEREDITARILY COUNTABLE SETS

The class H_{ω_1} is also known as the class of sets *hereditarily of countable cardinality*, and so also is given the abbreviation of HC. $\mathcal{P}(\omega) \subseteq \mathrm{HC}$ and hence we regard the real continuum as a subclass of HC. At least in one crude sense, HC "is" $\mathcal{P}(\omega)$, see the following Exercise.

EXERCISE 2.36 If $x \in \mathrm{HC}$ then we have $|\mathrm{TC}(x)| \leq \omega$. Define a wellfounded extensional relation E on ω so that $\langle \omega, E\rangle \cong \langle \mathrm{TC}(x), \in\rangle$. [Hint: We have a bijection $f : N \longleftrightarrow \mathrm{TC}(x)$ for some $N \leq \omega$; define $nEm \leftrightarrow f(n) \in f(m)$. If we use a recursive pairing bijection $p : \omega \longleftrightarrow \omega \times \omega$ (for example $p^{-1}(\langle k, l\rangle) = 2^k.(2l+1) - 1$) we may further code E as a subset $\overline{E} \subseteq \omega$. We thus have effectively coded up $\mathrm{TC}(x)$ as a subset of ω.] (By using further such coding devices we may take any countable structure with domain in HC and code it up as a subset of ω. In this sense to study all countable structures is to study all of $\mathcal{P}(\omega)$.)

However unlike the case of ω and HF, we cannot identify HC with any V_α: $V_{\omega+1} \supseteq \mathcal{P}(\omega)$ but $V_{\omega+1}$ does not contain any countable ordinal $\alpha > \omega + 1$. But $\omega_1 \subseteq \mathrm{HC}$ as can be easily determined from its definition. On the other hand $|V_{\omega+2}| = |\mathcal{P}\mathcal{P}(\omega)| = 2^{2^\omega} > 2^\omega = |\mathrm{HC}|$ so $V_{\omega+2} \nsubseteq \mathrm{HC}$. Clearly then HC is not closed under the power set operation but we do have that all other ZF axioms hold there:

LEMMA 2.36 $(\mathrm{ZF}^-)^{\mathrm{HC}}$.

EXERCISE 2.37 Which axioms of ZF hold in V_α if $\mathrm{Lim}(\alpha)$? Find a wellordering $\langle A, R\rangle \in V_{\omega+\omega}$ but for which there is no ordinal $\beta \in V_{\omega+\omega}$ with $\langle A, R\rangle \cong \langle \beta, <\rangle$; hence find an instance of the Ax.Replacement that fails in $V_{\omega+\omega}$. [The latter is a model of Z, the axiom system of Zermelo which is ZF with Replacement removed. For almost all regions of mathematical discourse, $V_{\omega+\omega}$ is a sufficiently large "universe" - mathematicians never, or rarely, need sets outside of this set.]

How large is H_κ? This depends again on the power set operation on sets of ordinals. Every element of H_{κ^+} can be coded as a subset of κ. See the next exercise which just mirrors the argument of Ex.2.36.

EXERCISE 2.38 $*$[1] Extend Ex.2.36 to any H_{κ^+}. [Hint: let p now be any pairing bijection $p : \kappa \longleftrightarrow \kappa \times \kappa$. Assume $f : \kappa \longleftrightarrow \mathrm{TC}(x)$ and put $\alpha E_0 \beta$ if $f(\alpha) \in f(\beta)$. Then by the Collapsing Lemma $\langle \kappa, E_0 \rangle \cong \langle \mathrm{TC}(x), \in \rangle$. Let $E = p^{-1}``E_0$. Then any structure with domain in H_{κ^+} can be coded by a subset of $E \subseteq \kappa$.] Deduce that $|H_{\kappa^+}| = |\mathcal{P}(\kappa)|$.

We adopt the notation: For $\kappa, \lambda \in \mathrm{Card}$, $\kappa^{<\lambda} =_{\mathrm{df}} \sup\{\kappa^{\mu} \mid \mu \in \mathrm{Card} \wedge \mu < \lambda\}$.

EXERCISE 2.39 Let $\kappa \in \mathrm{Card}$. Show that $|H_\kappa| = 2^{<\kappa}$. [Hint: for κ a successor cardinal, this is the last Exercise.]

EXERCISE 2.40 (Levy) Let $h(\kappa)$ be the class of sets x with (i) $\forall y \in \mathrm{TC}(x)(|y| < \kappa)$, (ii) $|x| < \kappa$. Show that if $\kappa \in \mathrm{Reg}$, then $H_\kappa = h(\kappa)$; find an example where this fails if κ is singular.

2.5 The Montague-Levy Reflection theorem

This section proves a *Reflection Theorem*, so called because it shows that in ZF we can prove that the fact of any sentence φ holding in V is reflected by an initial portion of the universe: we shall see that $\varphi \leftrightarrow \varphi^{V_\alpha}$ for some α, indeed for unboundedly many $\alpha \in On$. However these arguments are of more interest than just as a means to solving this problem.

We shall be able to prove from this theorem that any *finite* collection S of the ZF (or ZFC) axioms can be shown to hold in a transitive set; indeed we shall see that we can always find a level of the cumulative hierarchy, a V_α, in which S is true: $\mathrm{ZF} \vdash \exists \alpha (S)^{V_\alpha}$. Of course we have just seen that all of ZF^- is true in any H_{κ^+}. If our finite list contains the Ax.Power then Reflection arguments provide a solution. From this we shall be able to see later that ZFC is not *finitely axiomatisable*: there is no finite set of axioms S that have the same deductive consequences as those of ZFC.

2.5.1 Absoluteness

DEFINITION 2.37 *Let $W \subseteq Z$ be class terms. Let $\varphi \in \mathcal{L}_\in$ with* $\mathrm{FVbl}\{\varphi\} \subseteq \{\vec{x}\}$.

(i) φ is upward absolute *for W, Z iff* $\forall \vec{x} \in W(\varphi^W \longrightarrow \varphi^Z)$;

(ii) φ is downward absolute *for W, Z iff* $\forall \vec{x} \in W(\varphi^W \longleftarrow \varphi^Z)$;

(iii) φ is absolute *for W, Z if both (i) and (ii) hold:* $\forall \vec{x} \in W(\varphi^W \longleftrightarrow \varphi^Z)$

If $Z = V$ then we omit it, and simply say "φ is upward absolute for W" etc. If $\vec{\varphi} = \varphi_1, \ldots, \varphi_n$ is a finite list of formulae then we say that $\vec{\varphi} = \varphi_1, \ldots, \varphi_n$ are upward absolute (etc.) *if their conjunction $\bigwedge\!\!\!\bigwedge \varphi \equiv \varphi_1 \wedge \cdots \wedge \varphi_n$ is.*

DEFINITION 2.38 *Given classes $W \subseteq Z$ and a term t we say t is* absolute *for W, Z iff*

$$\forall \vec{x} \in W(t(\vec{x})^W \in W \leftrightarrow t(\vec{x})^Z \in Z \wedge t(\vec{x})^Z = t(\vec{x})^W)$$

(Recall that asserting $t(\vec{x})^Z \in Z$ is to assert that $t(\vec{x})^Z$ is a set of Z. Note we could have defined 'upwards' and 'downwards' absoluteness for terms t as well.) A standard example of a term that is not absolute is given by "*the first uncountable cardinal*" ($t = \{\alpha \in \mathrm{On} \mid \alpha \text{ is countable }\}$). Suppose $W \subseteq V$. Certainly $t^V = t$ is defined: it is ω_1. It may be that t^W is defined, and is a cardinal in W. But V may simply have more onto functions f with $\mathrm{dom}(f) = \omega$ and $\mathrm{ran}(f) \subseteq \mathrm{On}$, than W has. We may thus have $t^W < t^V$. Another example is given by $t = \mathcal{P}(\omega)$.

[1] An Exercise annotated with a $*$ indicates that is perhaps harder than usual. An (E) indicates that it is *Extra* to the course.

Definition 2.39 *A list of formulae* $\overrightarrow{\varphi} = \varphi_1, \ldots, \varphi_n$ *is* subformula closed *iff every subformula of a formula is on the list.*

The following establishes a criterion for when a formula's truth value is identical when interpreted in different in different class terms.

Lemma 2.40 *Let* $\overrightarrow{\varphi}$ *be a subformula closed list. Let* $W \subseteq Z$ *be terms. The following are equivalent:*

(i) $\overrightarrow{\varphi}$ *are absolute for* $W, Z.$ *;*

(ii) whenever φ_i *is of the form* $\exists x \varphi_j(x, \vec{y})$ *(with* $\mathrm{FVbl}(\varphi_i) \subseteq \{\vec{y}\}$ *) it satisfies* the Tarski-Vaught criterion *between* W *and* Z*:*

$$\forall \vec{y} \in W[\exists x \in Z \varphi_j(x, \vec{y})^Z \longrightarrow \exists x \in W \varphi_j(x, \vec{y})^Z].$$

Proof: *(i)* $\Rightarrow$*(ii)*: Fix $\vec{y} \in W$ and assume $\varphi_i(\vec{y})^Z \equiv \exists x \in Z \varphi_j(x, \vec{y})^Z$. By absoluteness of φ_i, $\varphi_i(\vec{y})^W$, so $\exists x \in W \varphi_j(x, \vec{y})^W$ and by absoluteness φ_j, $\varphi_j(x, \vec{y})^Z$, so $\exists x \in W \varphi_j(x, \vec{y})^Z$.

(ii)$\Rightarrow$ *(i)*: By induction on the length of φ_i: we thus assume absoluteness checked for all φ_j on the list for shorter length, in particular for any subformula of φ_i.

φ_i *atomic*: absolute by definition.

$\varphi_i \equiv \varphi_j \vee \varphi_k$: then φ_i is absolute since both φ_j and φ_k are by the inductive hypothesis.

$\varphi_i \equiv \neg\varphi_j$: similar;

$\varphi_i \equiv \exists x \varphi_j(x, \vec{y})$. So fix $\vec{y} \in W$.

$$\varphi_i(\vec{y})^W \leftrightarrow \exists x \in W \varphi_j(x, \vec{y})^W \leftrightarrow \exists x \in Z \varphi_j(x, \vec{y})^Z \leftrightarrow \varphi_i(\vec{y})^Z$$

Where : the first and last equivalence is just the definition of relativisation; the second equivalence - from left to right uses the absoluteness of φ_j from the Ind.Hyp., and the fact that $W \subseteq Z$; and from right to left uses Assumption *(ii)* and again the absoluteness of φ_j from the Ind. Hyp. Q.E.D.

Lemma 2.41 *Let* W *be a transitive class term. Then any* Δ_0*-formula* φ *is absolute for* W*.*

Proof: Let φ be Δ_0 and apply the last argument (with $\overrightarrow{\varphi}$ the list of φ together with all it subformulae). The point here is that $\mathrm{Trans}(W)$ so W knows the full $\in$-relationship on its members. As any Δ_0-formula only contains bounded quantifiers, this is enough to satisfy the criterion of 2.40 when one comes to the induction step $\varphi \equiv \exists x \in y \psi$ where ψ is Δ_0 itself, in the induction step at the end of the last proof.

Exercise 2.41 Fill in the details. [Hint: by what has just been said, only the $\varphi \equiv \exists x \in y\psi$ step and the last chain of equivalences needs to be argued.]

Exercise 2.42 Let W be a transitive class term. Then (i) any Σ_1-formula φ is upwards absolute for W; (ii) any Π_1-formula φ is downwards absolute for W.

2.5.2 Reflection Theorems

We use the last criterion of absoluteness in our Reflection Theorems. The first lemma really contains the essence of the argument.

Lemma 2.42 *Let Z be a class term, and suppose we have a function F_Z with $F_Z(\alpha) = Z_\alpha$ so that $\forall\alpha(Z_\alpha \in V)$. Assume*

(i) $\alpha < \beta \longrightarrow Z_\alpha \subseteq Z_\beta$;

(ii) $\mathrm{Lim}(\lambda) \longrightarrow Z_\lambda = \bigcup_{\alpha\in\lambda} Z_\alpha$; *Then for any* $\overrightarrow{\varphi} = \varphi_0, \ldots, \varphi_n$:

(iii) $Z = \bigcup_{\alpha\in\mathrm{On}} Z_\alpha$.

$$(*) \qquad \mathrm{ZF} \vdash \forall\alpha\exists\beta > \alpha(\overrightarrow{\varphi} \textit{ are absolute for } Z_\beta, Z).$$

Note: Formally here we are saying that if we have a term for Z and a term for the function F_Z, and we can prove in ZF that F_Z has properties (i) - (iii), then for any $\overrightarrow{\varphi}$, there is a proof in ZF of $(*)$. We are *not* saying that in ZF $\vdash$"$\forall\overrightarrow{\varphi}((*))$ *holds*". (Assertions such as the latter we shall see later are false.)

Proof: We apply Lemma 2.40 and try and find some $W = Z_\beta$ such that (ii) of the lemma applies. This will suffice. By lengthening the list if need be we shall assume that $\overrightarrow{\varphi}$ is subformula closed. For $i \le n$ we define functions $F_i : \mathrm{On} \longrightarrow \mathrm{On}$. If $\varphi_i \equiv \exists x\varphi_j(x, \vec{y})$ set:

$$\begin{aligned} G_i(\vec{y}) &= 0 \text{ if } \neg\exists x \in Z\varphi_j(x,\vec{y})^Z \\ &= \eta \text{ where } \eta \text{ is least so that } \exists x \in Z_\eta\varphi_j(x,\vec{y})^Z. \\ F_i(\xi) &= \sup\{G_i(\vec{y}) \mid \vec{y} \in Z_\xi\}. \end{aligned}$$

Note that G_i is a well defined function, and consequently so is F_i: G_i"$Z_\xi \in V$ by AxReplacement; hence $F_i(\xi) = \sup G_i$"Z_ξ is then a well defined term. Note also that each F_i is monotonic: $\zeta < \xi \longrightarrow F_i(\zeta) \le F_i(\xi)$. If φ_i is not of the above form, set $F_i(\xi) = 0$ everywhere.

Claim: $\forall\alpha\exists\beta > \alpha(\mathrm{Lim}(\beta) \wedge \forall\xi < \beta\forall i \le nF_i(\xi) < \beta)$.

Proof of Claim: Define by recursion on ω: $\lambda_0 = \alpha$;

$\lambda_{k+1} = \max\{\lambda_k + 1, F_0(\lambda_k), \ldots, F_n(\lambda_k)\}$; $\beta = \sup_k \lambda_k$.

Then $\lambda_k < \lambda_{k+1}$ implies that $\mathrm{Lim}(\beta)$. Hence if $\tau < \beta$ then $\tau < \lambda_k$ for some $k \in \omega$. Hence $F_i(\tau) \le F_i(\lambda_k) \le \lambda_{k+1} < \beta$. Q.E.D.(*Claim*)

Now that the Claim is proven, then we may verify the Lemma with such a β for Z_β and Z.

Q.E.D.

Exercise 2.43 Carry out this final verification.

We may immediately set Z to be V and Z_α to be V_α and obtain the corollary:

Theorem 2.43 **(Montague-Levy) The Reflection Theorem**. *Let $\overrightarrow{\varphi}$ be any finite list of formulae of $\mathcal{L}$. Then*

$$\mathrm{ZF} \vdash \forall\alpha\exists\beta > \alpha(\overrightarrow{\varphi} \textit{ are absolute for } V_\beta). \qquad \text{Q.E.D.}$$

As cautioned above, this is a *theorem scheme* again: it is one theorem of ZF for each choice of $\overrightarrow{\varphi}$. Notice that if, in particular, $\overrightarrow{\varphi}$ are sentences, we may write the conclusion as:

$$\mathrm{ZF} \vdash \forall\alpha\exists\beta > \alpha(\overrightarrow{\varphi} \longleftrightarrow (\overrightarrow{\varphi})^{V_\beta}).$$

Moreover if the $\overrightarrow{\varphi}$ are *axioms* of ZF we have that they are true in V. In this case we may write: $\mathrm{ZF} \vdash \forall\alpha\exists\beta > \alpha(\,(\bigwedge\overrightarrow{\varphi})^{V_\beta})$.

In other words: for any finite list from ZF we can find arbitrarily large β so that those axioms hold in V_β. We can state something stronger:

Corollary 2.44 *Let T be any set of axioms in $\mathcal{L}$ extending* ZF, *and $\vec{\varphi}$ a finite list of axioms from T. Then $T \vdash \forall\alpha\exists\beta > \alpha((\bigwedge\!\!\!\bigwedge \vec{\varphi})^{V_\beta})$.*

Proof: Since T extends ZF T proves the existence of the V_α hierarchy, and $T \vdash \varphi_i$ for each φ_i from $\vec{\varphi}$. Hence $T \vdash \bigwedge\!\!\!\bigwedge \vec{\varphi}$ trivially. And $T \vdash \forall\alpha\exists\beta > \alpha(\bigwedge\!\!\!\bigwedge \vec{\varphi} \longleftrightarrow (\bigwedge\!\!\!\bigwedge \vec{\varphi})^{V_\beta})$ Q.E.D.

At first blush it might look as if the restriction to finite lists of $\vec{\varphi}$ is unnecessary. Why could we not look at a recursive enumeration φ_i of all axioms of ZF say, and find some V_α in which they were all true? We know from the Gödel Second Incompleteness Theorem that there is no way to formalise that argument within ZF, since it would be tantamount to proving the existence of a *model* of the ZF axioms, and hence the *consistency* of ZF. So what goes wrong? Lemma 2.42 can only work for finite lists $\vec{\varphi}$: the statement "$\vec{\varphi}$ *are absolute for* Z_β, Z" involves a conjunction of the formulae from the list: we cannot write an infinitely long formula in $\mathcal{L}$, so we have no way of even expressing the absoluteness of such an infinite list. Another paraphrase on this is in the following Exercise.

Exercise 2.44 Show that for every formula φ of $\mathcal{L}$:

ZF $\vdash$"There is a c.u.b. class $C \subseteq$ On so that $\forall\alpha \in C\forall\vec{x} \in V_\alpha(\varphi(\vec{x}) \leftrightarrow (\varphi(\vec{x}))^{V_\alpha})$"

[Hint: The reasoning of Lemma 2.42 pretty much gives the relevant cub class as the closure points of the F_i.]
Remark: One might think that one could enumerate all the axioms of ZF $\varphi_0, \varphi_1, \ldots$, find the appropriate classes C_{φ_n} and take $D = \bigcap_n C_{\varphi_n}$. This appears then to be an intersection of only countable many c.u.b. classes and so must be c.u.b. in On? But for any element $\alpha \in D$ we'd have $(\mathrm{ZF})^{V_\alpha}$, and we appear to have proven the existence of models of ZF - contradicting Gödel. What is wrong with this reasoning?

Exercise 2.45 Find a sentence σ so that if σ is absolute for V_α then α is a limit ordinal. Repeat the exercise and find τ so that if τ is absolute for V_β then $\beta = \omega_\beta$ (the β'th infinite cardinal). [Hint: consider the statement: "*For every β ω_β exists*".]

As the last exercise shows, if we insist on finding a V_α which is absolute for any particular sentence, then we may need to find a very large α for this to happen. If we are content to merely find *a set* for which a formula is absolute, we can find a countable such set. More generally:

Lemma 2.45 *Let Z be a term, and $\vec{\varphi}$ be any finite list of formulae of $\mathcal{L}$. Then*

ZFC $\vdash \forall x \subseteq Z\exists y[x \subseteq y \subseteq Z \wedge \vec{\varphi}$ *are absolute for y, Z* $\wedge |y| \leq \max\{\omega, |x|\}]$.

Proof: We define from the term Z the term giving the function $F(\alpha) = Z \cap V_\alpha$ which we shall call Z_α. Again assume that $\vec{\varphi}$ is subformula closed. As x is a set, by the AxReplacement $G``x \in V$ where $G(u) =_{\mathrm{df}}$ the least $\alpha$ such that $u \in Z_\alpha$ (or $= 0$ if $u \notin Z$). Then $\sup G``x = \bigcup G``x \in V$. Call this ordinal β_0. By Lemma 2.42 find $\beta > \beta_0$ with $\vec{\varphi}$ absolute for Z_β, Z. By AC fix a wellorder $\lhd$ of Z_β. Without loss of generality we assume $\varnothing \in Z_\beta$. If φ_i is of the form $\exists x\varphi_j(x, y_1, \ldots, y_{k_j})$ (with $\mathrm{FVbl}(\varphi_i) \subseteq \{\vec{y}\}$) we define a function $G_i : {}^{k_j}Z_\beta \longrightarrow Z_\beta$ by the following clauses:

$h_i(\vec{y})$ = the $\lhd$-least $x \in Z_\beta$ so that $\varphi_j(x, y_1, \ldots, y_{k_j})^{Z_\beta}$ if such exists
$= \varnothing$ otherwise.

We also set h_i to be the constant $\varnothing$-function in the cases that φ_i is not of the above form, or that φ_i has no free variables. With h_i now defined in every case, we look for the least set y closed under the h_i. We can find such a y by repeatedly closing under the finitary functions h_i, and obtain a y with cardinality no greater than $\max\{\omega, |X|\}$ (see Exercise 2.46). We can then appeal to the criterion in Lemma 2.40, which asserts in this case that $\vec{\varphi}$ is absolute for y, Z_β. But $\vec{\varphi}$ is absolute for Z_β, Z, and thus the Lemma is proven. Q.E.D.

Exercise 2.46 Let x be any set, and $f_i : {}^{n_i}V \longrightarrow V$ for $i < \omega$ be any collection of finitary functions (meaning that $n_i < \omega$); show that there is a $y \supseteq x$ which is closed under each of the f_i (thus f_i“ ${}^{n_i}y \subseteq y$ for each i) and $|y| \leq \max\{\omega, |x|\}$. [Hint: no need for a formal argument here: build up a y in ω many stages $y_k \subseteq y_{k+1}$ at each step applying all the f_i.]

The last lemma then says that, *e.g.* , if φ were a finite list of axioms of ZFC, and $x = \varnothing$, then $\langle y, \in\rangle$ would be a countable structure in which those axioms were true.

Returning to our reflection results, we may apply the above to obtain corollaries to Lemma 2.45.

Corollary 2.46 *Let Z be a term, and $\vec{\varphi}$ be any finite list of formulae of $\mathcal{L}$. Then*

$$\mathrm{ZFC} \vdash \forall x \subseteq Z[\mathrm{Trans}(x) \longrightarrow \exists w[x \subseteq w \wedge \mathrm{Trans}(w) \wedge \vec{\varphi} \textit{ are absolute for } w, Z \wedge |w| \leq \max\{\omega, |x|\}]$$

Proof: We directly apply the Mostowski-Shepherdson Collapsing Lemma to the set y appearing in the statement of Lemma 2.45, thereby collapsing it to the transitive $w \supseteq x$ here. As $\langle w, \in\rangle \cong \langle y, \in\rangle$ we have $\varphi(\vec{v})^y \leftrightarrow \varphi(\pi(\vec{v}))^w$. Hence $\vec{\varphi}$ are absolute for w, Z. Obviously $|y| = |w|$. Q.E.D.

In the special case that $Z = V$ and $x = \omega$ in the above we may get:

Corollary 2.47 *Let T be any set of axioms in $\mathcal{L}$ extending* ZFC*, and $\vec{\varphi}$ a finite list from T, then*

$$T \vdash \exists y[\mathrm{Trans}(y) \wedge |y| = \omega \wedge \bigwedge\!\!\!\bigwedge(\vec{\varphi})^y].$$

Thus we can find for any finite set of ZFC axioms a countable transitive set model in which all those axioms come out true. Again the finiteness of $\vec{\varphi}$ is necessary.

2.6 Inaccessible Cardinals

We shall encounter in this section an example of a ‘large cardinal’: this is a cardinal whose existence does not follow from the axioms of ZFC. In general this is because such cardinals allow one to conclude that there are structures (typically V_κ where κ is the cardinal number under consideration) in which all the ZFCaxioms are true. If ZFC could prove the existence of such a κ then this would contradict the Gödel Second Incompleteness Theorem. From these further large cardinals can be defined, and although we give the briefest of illustrative examples, it is not the intention of the course to go down this route, rich as it is.

2.6.1 Inaccessible cardinals

Definition 2.48 *A cardinal $\kappa > \omega$ is a* strong limit cardinal, *if for any $\alpha < \kappa \longrightarrow 2^{|\alpha|} < \kappa$.*

DEFINITION 2.49 *A regular cardinal $\kappa > \omega$ is*

(i) weakly inaccessible *if it is a limit cardinal (Hausdorff 1908);*

(ii) (Sierpinski-Tarski (1930); Zermelo (1930)) strongly inaccessible *if in addition it is a strong limit cardinal.*

The idea behind the nomenclature is that an accessible cardinal κ is one that can be reached from below by either the successor cardinal operation, or else the power set operation, as per Note (1) that follows.

Notes (1) Another way of putting this is to say that a cardinal κ is weakly inaccessible if it is (a) regular and (b) $\alpha < \kappa \longrightarrow \alpha^+ < \kappa$. It is (strongly) inaccessible if it is both (a) regular and (c) $\alpha < \kappa \longrightarrow |\mathcal{P}(\alpha)| < \kappa$.

(2) The word 'strongly' is often omitted.

(3) If the GCH holds then the two notions coincide (for the simple reason that GCH $\longrightarrow 2^{|\alpha|} = |\mathcal{P}(\alpha)| = \alpha^+ < \kappa!$).

(4) The least strong limit cardinal is singular of cofinality ω (Check!) In particular if GCH holds then $\aleph_\omega$ is the least strong limit cardinal.

LEMMA 2.50 *(AC) Let $\omega < \kappa \in$ Reg. The following are equivalent:*

(i) κ is strongly inaccessible;

(ii) $V_\kappa = H_\kappa$;

(iii) $(\mathrm{ZFC})^{H_\kappa}$;

(iv) $\kappa = \beth_\kappa$.

PROOF: $(i) \Rightarrow (ii)$. Since $\kappa \in$ Card, we have $H_\kappa \subseteq V_\kappa$ (Lemma 2.33(ii)). But $x \in V_\kappa \Rightarrow \exists\alpha < \kappa(x \in V_\alpha)$. By induction on $\alpha < \kappa$ one shows that $|V_\alpha| < \kappa$: suppose true for $\beta < \alpha$: then $V_\alpha = \mathcal{P}(V_\beta)$ if $\alpha = \beta + 1$, and as $|V_\beta| < \kappa$, then $|\mathcal{P}(V_\beta)| = |2^{|V_\beta|}| < \kappa$ as κ is strongly inaccessible; if $\mathrm{Lim}(\alpha)$ then V_α is the union of less than κ many sets of size less than κ, and hence has cardinality less than κ. Hence, in either case V_α is a transitive set of size less than κ. Hence it is in H_κ.

$(ii) \Rightarrow (iii)$. We have already that $(\mathrm{ZFC}^-)^{H_\kappa}$ (by Lemma 2.34). Only Ax.Power is missing. But $(\mathrm{Ax\,.\,Power})^{V_\lambda}$ for any limit ordinal λ, and hence in particular for $\lambda = \kappa$.

$(iii) \Rightarrow (iv)$. We prove by induction that $\alpha < \kappa \longrightarrow \beth_\alpha < \kappa$. This suffices. Assume true for $\beta < \alpha$. If $\alpha = \beta + 1$ then $2^{\beth_\beta} = \beth_\alpha$. But $(\mathrm{AxPower} + AC)^{H_\kappa}$, hence $(\exists\tau \in \mathrm{On}(\tau \approx \mathcal{P}(\beth_\beta))^{H_\kappa}$. So $2^{\beth_\beta} = |\mathcal{P}(\beth_\beta)| \leq \tau < \kappa$. If $\mathrm{Lim}(\alpha)$ then $\beth_\alpha < \kappa$ by the inductive hypothesis and the regularity of κ.

$(iv) \Rightarrow (i)$. Recall that $|V_{\omega+\alpha}| = \beth_\alpha$ (Ex. 2.9). Our assumption yields that

$$\omega^2 \leq \alpha < \kappa \longrightarrow 2^{|\alpha|} = |\mathcal{P}(\alpha)| \leq |V_{\alpha+1}| = \beth_{\alpha+1} < \kappa$$

as required for strong inaccessibility. Q.E.D.

EXERCISE 2.47 Verify that κ is weakly inaccessible iff κ is regular and $\kappa = \aleph_\kappa$.

EXERCISE 2.48 Does $\kappa > \omega \wedge V_\kappa = H_\kappa$ imply that κ is strongly inaccessible?

DEFINITION 2.51 *(Mahlo 1911) A regular limit cardinal κ is called a* weakly Mahlo cardinal *in case* $\mathrm{Reg} \cap \kappa$ *is stationary below κ. κ is called* (strongly) Mahlo *if it is both weakly Mahlo and strongly inaccessible.*

LEMMA 2.52 *If κ is weakly Mahlo then in fact κ is the κ'th weakly inaccessible cardinal, and the class of weakly inaccessible cardinals below κ is stationary below κ. The same sentence is true with 'strongly' replacing 'weakly' throughout.*

PROOF: As $\mathrm{Reg} \cap \kappa$ is unbounded in κ, $(\mathrm{Reg} \cap \kappa)^*$ is c.u.b. below κ. But such are all limit cardinals. As $\mathrm{Reg} \cap \kappa$ is moreover stationary below κ, $D =_{\mathrm{df}} (\mathrm{Reg} \cap \kappa) \cap (\mathrm{Reg} \cap \kappa)^*$ is stationary below κ (see Ex.2.13). But all members of D are then weakly inaccessible cardinals. Q.E.D.

EXERCISE 2.49 Let λ be the least weakly inaccessible cardinal which is itself a limit of weakly inaccessible cardinals (meaning the weakly inaccessibles below λ are unbounded in λ). Show that λ is not weakly Mahlo. The same sentence is true with 'strongly' replacing 'weakly' throughout.

2.6.2 A MENAGERIE OF OTHER LARGE CARDINALS

We briefly consider some other notions of "large cardinal" stronger than Mahlo. (For a full account see Drake [2], Devlin [1], Jech [3].) We do this to give some flavour to the rich structure of even the so-called small large cardinals. They are called 'small' because, if they are consistent, then they are consistent with the statement that "$V = L$" - they can thus potentially be exemplified in L. Several depend upon the notion of a *homogeneous set* for a certain kind of function.

DEFINITION 2.53 *(i) $[\kappa]^n$ denotes the set of all n element subsets of κ.*
(ii) $[\kappa]^{<\omega}$ denotes the set of all finite subsets of κ

DEFINITION 2.54 *$H \subseteq \kappa$ is* homogeneous *for $f : [\kappa]^n \longrightarrow \lambda \Longleftrightarrow_{\mathrm{df}} |f``[H]^n| = 1$.*

A homogeneous set is one therefore that every n-tuple there from gets sent by f to the same ordinal $\xi < \lambda$. Often in applications $\lambda = 2 = \{0,1\}$ so we can think of f as partition of $[\kappa]^n$ into two colours. If H is homogeneous, then this means that all n-tuples from H are assigned the same colour. For λ colours the same applies. If a longer order type is specified on H then the harder it is to find such homogeneous sets. Large cardinals can then be specified by putting requirements on H and so forth as in the next two definitions.

DEFINITION 2.55 *A cardinal κ is* weakly compact *if for every $f : [\kappa]^2 \longrightarrow 2$ there is a homogeneous subset $H \subseteq \kappa$ with H unbounded in κ.*

DEFINITION 2.56 *(Jensen) A cardinal κ is* ineffable *if for every $f : [\kappa]^2 \longrightarrow 2$ there is a homogeneous subset $H \subseteq \kappa$ with H stationary in κ.*

By themselves the bare definitions may not mean too much. We give some equivalent formulations.

DEFINITION 2.57 *(i) A* tree *$\langle T, <_T \rangle$ is a wellfounded partial ordering so that for any $s \in T$, $\{s_0 \in T \mid s_0 <_T s\}$ is linearly ordered.*
(ii) A branch *through a tree T is a maximal linearly ordered set;*
(iii) $T_\alpha =_{\mathrm{df}} \{s \in T \mid \mathrm{rank}_T(s) = \alpha\}$ is the set of elements of the tree of tree-rank or 'level' α.

A tree thus looks how it sounds.

Definition 2.58 *Let us say that a cardinal* κ *has the* tree property *iff for every tree* $T = \langle \kappa, <_T \rangle$ *with* $\forall \alpha < \kappa (|T_\alpha| < \kappa)$ *has a branch of order type* κ.

There is no reason for a cardinal in general to satisfy the tree property. For example on ω_1 it may be the case that there is an uncountable tree $T = \langle \omega_1, <_T \rangle$, with field ω_1, with all levels T_α countable, yet without any branch of cardinality ω_1. (Such trees are called *Aronszajn trees*.) However the König Tree Lemma shows that ω_0 has the tree property.

Lemma 2.59 *For a cardinal* κ *the following are equivalent:*

(i) κ *is strongly inaccessible and satisfies the tree property;*

(ii) κ *is weakly compact;*

(iii) for every $A \subseteq \kappa$ *there is a transitive* M*, and a* B, j *with* $j : \langle V_\kappa, \in, A \rangle \longrightarrow \langle M, \in, B \rangle$ *an elementary embedding with* $j \upharpoonright \kappa = \mathrm{id} \upharpoonright \kappa$ *and* $j(\kappa) > \kappa$.

There are many further characterisations of weakly compact. See Jech, Drake. One property of weakly compact cardinals is that every stationary subset of κ reflects this property below κ, as in the following Exercise.

Exercise 2.50 $(*)$ Let κ be weakly compact. Show that for any stationary subset $S \subseteq \kappa$, there is $\lambda < \kappa$ so that $S \cap \lambda$ is stationary in λ. [Hint: Use (iii) of the last lemma: suppose the conclusion fails; then there is $C_\lambda \subseteq \lambda$ with $C_\lambda \cap S \cap \lambda = \varnothing$ for every cardinal $\lambda < \kappa$. Let $A = \{\langle \xi, \lambda \rangle \mid \xi \in C_\lambda\} \cup S \times \{0\}$. Let j, M, B be as in (iii) above. Let $C_\kappa = \{\xi \mid \langle \xi, \kappa \rangle \in B\}$. By elementarity of the embedding j the following holds in M : "C_κ is c.u.b.in κ, whilst $C_\kappa \cap S \cap \kappa = \varnothing$". But $(S \cap \kappa)_M = S$ - so this is a contradiction.]

Definition 2.60 *(Jensen) A cardinal* κ *is* subtle *iff*

For any sequence $\langle A_\alpha \mid \alpha < \kappa \rangle$ *with all* $A_\alpha \subseteq \alpha$ *and any c.u.b.* $C \subseteq \kappa$*, there is a pair of* $\alpha, \beta \in C$ *with* $\alpha < \beta \wedge A_\beta \cap \alpha = A_\alpha$*).*

Lemma 2.61 *(Jensen) For a cardinal* κ *the following are equivalent:*

(i) κ *is ineffable ;*

(ii) for any sequence $\langle A_\alpha \mid \alpha < \kappa \rangle$ *with all* $A_\alpha \subseteq \alpha$ *there is a set* $E \subseteq \kappa$*, so that*

$$\{\alpha < \kappa \mid A_\alpha = E \cap \alpha\} \text{ is stationary.}$$

Definition 2.62 *A cardinal* κ satisfies the partition relatio*n* $\kappa \longrightarrow (\gamma)_2^{<\omega} \Longleftrightarrow_{\mathrm{df}}$ *for any* $f : [\kappa]^{<\omega} \longrightarrow 2$ *there is an* $H \subseteq \kappa$*,* $\mathrm{ot}(H) \geq \gamma$*, which is* homogeneous *for* $f : [\kappa]^{<\omega} \longrightarrow \lambda$*, namely for all* $n < \omega$ — $f``[H]^n| = 1$.

The extra strength here is that f must assign the same colour to each n-tuple from H (although for a different $m \neq n$ a different colour may be chosen for all m-tuples from H). Such cardinals become rapidly stronger than those considered above, and quickly enter the realm of 'medium large cardinals'. This happens as soon as γ crosses the threshold from countable to uncountable. The cardinals here defined are in increasing strength, when measured in terms of where they are first exemplified in On: if κ is the least satisfying $\kappa \longrightarrow (\omega)_2^{<\omega}$ then κ is the κ'th ineffable cardinal. Similar if κ is the first ineffable, it is the κ'th subtle cardinal, and also the κ'th weakly compact cardinal. If κ is the first weakly compact cardinal, then it is the κ'th Mahlo cardinal. All the above are consistent with '$V = L$'; not however the existence of a cardinal κ satisfying $\kappa \longrightarrow (\omega_1)_2^{<\omega}$: if such a cardinal exists we may prove that $V \neq L$.

Chapter 3

Formalising semantics within ZF

The study of first order structures and the languages appropriate to them is the branch of mathematics called *model theory*. Like other parts of mathematics it can be formalised within set theory, and developed from the ZFaxioms. Whereas most mathematicians would not be seeing any great advantage in having their area of mathematics in doing this, as set theorists we shall see that formalising that part of model theory that handles structures of the form $\langle X, \in \rangle$, (or of $\langle X, \in, A_1, \ldots, A_n \rangle$ where $A_i \subseteq X$), will be of immense utility. Amongst other results it is at the heart of Gödel's construction of the *constructible hierarchy*, L.

We have defined the notion of absoluteness of formulae between structures or terms rather generally. However we have not been very specific about what kinds of concepts are actually absolute. We alluded to this problem at the end of Section 1.2, and in particular we noted the possible non-absoluteness of the power set operation. In general objects that have very simple definitions tend to be absolute for transitive sets and classes (thus $\varnothing, \{x, y\}, \omega$, "$f$ is a function", "x an ordinal") whilst more complex ones are not ($y = \mathcal{P}(x)$, "x is a cardinal").

3.1 Definite terms and formulae

The *definite* terms and formulae are amongst those that we are interested in being absolute between transitive ZF^- models. We address the question of which terms and formulae defining concepts can be so absolute. We shall define "definite term (and formula)" first and later show that such have this degree of being "absolutely definite".

Definition 3.1 (Definite terms and formulae)
(A) We define the definite terms *and* formulae *by a simultaneous induction on the complexity of formulae and of the terms' definition.*

(i) Any atomic formula $x = y$, $x \in y$ is definite ;
if φ, ψ are definite, then so are: $\neg\varphi$; $(\varphi \vee \psi)$; $\exists y \in x\varphi$

(ii) Any variable x is a definite term. If s, t are definite terms, so are:
$\bigcup s$, $\{s, t\}$, $s \backslash t$.

(iii) Suppose $t_0(x_1, \ldots, x_n)$ and $t_1, \ldots, t_n$ are definite terms. Then $t_0(t_1/x_1, \ldots, t_n/x_n)$ is a definite term. If $\varphi(x_1, \ldots, x_n)$ is a definite formula then so is $\varphi(t_1/x_1, \ldots, t_n/x_n)$.

(iv) If $\varphi(x_0, x_1, \ldots, x_n)$ and $t_1, \ldots, t_n$ are definite, then so are the terms:

$$y \cap \{x \mid \varphi(x, t_1/x_1, \ldots, t_n/x_n)\} \text{ and } \{t_1(y, x) \mid y \in z\}.$$

(v) ω.

(vi) If t is definite, and $\mathrm{Fun}(t)$, then the canonical function term f given by the recursion $f(y, \vec{x}) = t(y, \vec{x}, \{f(z, \vec{x}) \mid z \in y\})$ is definite.

Note (1): By (i) any Δ_0 formula of $\mathcal{L}$ is definite. (iv) gives a form of "definite separation" axiom, in the first part, and a kind of "definite replacement" in the second part. Note also that if s is a definite term then in particular "$x \in s$", "$\exists y \in s\varphi$" are definite formulae.

Lemma 3.2 (ZF^-) *If t is a definite term then: $\forall \vec{x}(t(\vec{x}) \in V)$.*

Proof: Formally this would be a proof by induction on the complexity of t; informally notice that the way we have defined definite terms uses methods, such as at (ii) where the ZF^- axioms yield these classes directly as sets, or in the case of (iii) and (iv) an appeal to Ax.Subsets would yield them as sets. In (vi) we appeal to the principle of recursion (which does not use Ax.Power) to ensure that f as defined there is a function of V^n to V (for some n). Q.E.D.

We shall be interested in terms and formulae that are absolute between any two transitive ZF^- models M, N. Such we shall call *absolutely definite*, a.d. for short. We shall be particularly interested in when they are so absolute between such an M and V. We shall readily be able to identify a whole host of terms and defining formulae as definite. We shall also be showing that any definite term or formula is a.d., and thus in one fell swoop be able to conclude they are absolute for such classes. As might be expected the proof proceeds by induction on the complexity of the term or formula.

Theorem 3.3 *Let $t(\vec{x})$ be a definite term, and $\varphi(\vec{x})$ a definite formula. Then (a) t and (b) φ are a.d., that is, they are absolute between any two transitive ZF^- models M, N.*

Proof: We shall first prove (a) and (b) by a simultaneous induction on the complexity of definite terms and formulae. We do this by referring to the construction clauses (i)-(vi) Def. 3.1 in turn. It suffices to prove this absoluteness between V and any transitive class term model of ZF^- W (note V is also a transitive ZF^- term). So let W be a transitive class term with $(\mathrm{ZF}^-)^W$. The atomic formulae of (i) are trivially so absolute, and the inductive steps in the more complex formulae are trivial except for the bounded existential quantifier; assume $y \in W$ and φ is absolute for W:

$$((\exists x \in y)\varphi)^W \leftrightarrow (\exists x(x \in y \wedge \varphi))^W \leftrightarrow \exists x \in W(x \in y \wedge \varphi^W) \leftrightarrow \exists x(x \in y \wedge \varphi^W) \leftrightarrow (\exists x \in y)\varphi$$

where we use the transitivity of W and hence that $y \subseteq W$, in the $\leftarrow$ direction of the third equivalence.

We remark that we have shown Lemma 2.41:

Corollary 3.4 *Let φ be a Δ_0 formula. Then φ is a.d.*

For (ii) suppose s, t are definite:

$(\bigcup s)^W = \{z \mid \exists y \in s(z \in y)\}^W = \{z \mid z \in W \wedge \exists y \in s^W(z \in y)\} = \{z \mid \exists y \in s^W(z \in y)\} = \bigcup s$ since $s^W \subseteq W$. $\{s, t\}$ and $s \backslash t$ are similar.

For (iii) suppose $t_0, \ldots, t_n$ are definite. Let $\vec{z} \supseteq \mathrm{Fvbl}\{t_1, \ldots, t_n\}$. Let $\{x_1, \ldots, x_n\} \supseteq \mathrm{Fvbl}(t_0)$. Then we make the inductive assumptions that for any $\vec{z} \in W : t_i(\vec{z})^W = t_i(\vec{z})$, and for any $\vec{x} \in W$ that $t_0(\vec{x})^W = t_0(\vec{x})$. By Lemma 3.2, if $t_i(\vec{z})$is defined for $\vec{z} \in W$ then we know that $t_i(\vec{z}) \in W$.

$$\begin{aligned}
(t_0(t_1(\vec{z})/x_1, \ldots, t_n(\vec{z})/x_n))^W &= t_0^W(t_1^W(\vec{z})/x_1, \ldots, t_n^W(\vec{z})/x_n) \\
&= t_0^W(t_1(\vec{z})/x_1, \ldots, t_n(\vec{z})/x_n) \\
&= t_0(t_1(\vec{z})/x_1, \ldots, t_n(\vec{z})/x_n)
\end{aligned}$$

The first equality is just the definition of relativisation to W and the next two are the inductive hypotheses outlined.

Entirely similarly,

$$(\varphi(t_1/x_1, \ldots, t_n/x_n))^W \leftrightarrow \varphi^W(t_1^W/x_1, \ldots, t_n^W/x_n) \leftrightarrow \varphi^W(t_1/x_1, \ldots, t_n/x_n) \leftrightarrow \varphi(t_1/x_1, \ldots, t_n/x_n)$$

where the new inductive hypothesis is now that $\varphi(\vec{x}) \leftrightarrow \varphi(\vec{x})^W$ for any $\vec{x} \in W$, and is used in the final equivalence. The first equivalence is Lemma 1.23.

For (iv): suppose $\varphi(x_0, x_1, \ldots, x_n)$ and $t_1, \ldots, t_n$ are definite, then:

$(y \cap \{x \mid \varphi(x/x_0, t_1/x_1, \ldots, t_n/x_n)\})^W$

$= y \cap W \cap (\{x \mid \varphi(x, t_1/x_1, \ldots, t_n/x_n)\})^W$

$= y \cap W \cap \{x \in W \mid (\varphi(x, t_1/x_1, \ldots, t_n/x_n))^W\}$

$= y \cap W \cap \{x \in W \mid \varphi(x, t_1/x_1, \ldots, t_n/x_n)\}$ (by (iii))

$= y \cap \{x \mid \varphi(x, t_1/x_1, \ldots, t_n/x_n)\}$ since $y \subseteq W$ as Trans(W).

Assume $z \in W$ and t_1 is definite. We make the inductive assumption that we have shown that $t_1(u, v)^W = t_1(u, v) \in W$ for any $u, v \in W$. Then

$\{t_1(y, x) | y \in z\}^W = \{t_1(y, x)^W | (y \in z)^W\} = \{t_1(y, x) | y \in z\}$

using that $z \subseteq W$ in the first equality.

For (v) we consider ω. We note that the following are expressible in a Δ_0 way and hence are absolute for W :

(a) $x = \varnothing \leftrightarrow \forall z \in x(z \neq z)$

(b) $\mathrm{Trans}(x) \leftrightarrow \forall y \in x \forall z \in y(z \in x)$;

(c) $x \in \mathrm{On} \leftrightarrow (\mathrm{Trans}(x) \wedge \forall y, z \in x(y \in z \vee z \in y \vee z = y))$;

(d) $\mathrm{Lim}(x) \leftrightarrow x \in \mathrm{On} \wedge x \neq \varnothing \wedge \forall y \in x \exists z \in x(y \in z)$;

(e) $x \in \omega \leftrightarrow x \in \mathrm{On} \wedge \neg \mathrm{Lim}(x) \wedge \forall y \in x \neg \mathrm{Lim}(y)$.

(f) $x = \omega \leftrightarrow x \in \mathrm{On} \wedge \mathrm{Lim}(x) \wedge \forall y \in x \neg \mathrm{Lim}(y)$

By (e) we have seen that $x \in \omega$ is given by a Δ_0 formula and hence is absolute for W. Now note that $\omega \subseteq W$: suppose $n \in \omega$ is least for which $n \notin W$. Then $0 = \varnothing \in W$ so $n = m + 1 =_{\mathrm{df}} m \cup \{m\}$. However if $m^W = m$ then by Ax.Pair and Union $(m \cup \{m\})^W \in W$ where

$(m \cup \{m\})^W = \{x \in W | (x \in m \vee x = m)^W\} = \{x \in W | (x \in m \vee x = m)\} = \{x | (x \in m \vee x = m)\} = m \cup \{m\}$.

Hence $\omega \subseteq W$. But then $\omega^W = \{x \in W | (x \in \omega)^W\} = \{x \in W | (x \in \omega)\}$ (by (e)
$= \{x | (x \in \omega)\}$ (since $\omega \subseteq W$).
$= \omega$.

Finally for (vi): we assume t is definite, and $\mathrm{Fun}(t)$, and f is the canonical function term given by:
$f(y, \vec{x}) = t(y, \vec{x}, \{f(z, \vec{x}) \mid z \in y\})$.

We thus have the inductive hypothesis that $t(y, \vec{x}, u)^W = t(y, \vec{x}, u)$ for any $y, \vec{x}, u \in W$. Let $y, \vec{x} \in W$. We prove the result by $\in$-induction, hence we also assume we have proven for any $z \in y$ that $f(z, \vec{x})^W = f(z, \vec{x}) \in W$. Then by (iv) we have: $\{f(z, \vec{x}) \mid z \in y\}^W = \{f(z, \vec{x}) \mid z \in y\} \in W$. Then:

$$\begin{aligned} f(y, \vec{x})^W &= (t(y, \vec{x}, \{f(z, \vec{x}) \mid z \in y\})^W \\ &= t(y, \vec{x}, \{f(z, \vec{x}) \mid z \in y\}^W) \\ &= t(y, \vec{x}, \{f(z, \vec{x}) \mid z \in y\}) \quad \text{(by the above comment)} \\ &= f(y, \vec{x}) \quad \text{as required.} \end{aligned}$$

Q.E.D.(Thm.3.3)

We now have a very powerful method for showing that all sorts of concepts and definitions are absolute for transitive structures in which ZF^- holds. For example all the ordinal arithmetic operations are defined by recursive clauses from definite terms. We can formally justify this as follows.

Lemma 3.5 *Suppose we define:*

$$\begin{array}{rcll} f(y, \vec{x}) &=& t_1(y, \vec{x}) & \text{if } \psi_1(y, \vec{x}) \\ &=& \vdots & \\ &=& t_n(y, \vec{x}) & \text{if } \psi_n(y, \vec{x}) \\ &=& \varnothing & \text{otherwise.} \end{array}$$

for some definite $t_1, \ldots, t_n$, and mutually exclusive (meaning at most one of $\psi_1(y, \vec{x}), \ldots, \psi_n(y, \vec{x})$ holds) but definite $\psi_1, \ldots, \psi_n$, then $f(y, \vec{x})$ is definite.

Proof: Note that $u_1 \cup \cdots \cup u_n = \bigcup\{u_1, \ldots, u_n\}$ so this is definite.
Then: $f(y, \vec{x}) = \{t_1(y, \vec{x}) | \psi_1(y, \vec{x})\} \cup \cdots \cup \{t_n(y, \vec{x}) | \psi_n(y, \vec{x})\}$. Q.E.D.

Corollary 3.6 *All the arithmetical functions $A_\alpha(\beta) = \alpha + \beta$; $M_\alpha(\beta) = \alpha.\beta$; $E_\alpha(\beta) = \alpha^\beta$ are definite and hence a.d.*

Proof: For example:

$$\begin{array}{rcll} A_\alpha(x) &=& \alpha & \text{if } x = \varnothing; \\ A_\alpha(x) &=& A_\alpha(y) + 1 & \text{if } x \in \mathrm{On} \wedge \mathrm{Succ}(x) \wedge x = S(y); \\ A_\alpha(x) &=& \sup\{A_\alpha(y) | y \in x\} & \text{if } x \in \mathrm{On} \wedge \mathrm{Lim}(x). \end{array}$$

The first and third conditions on the right we have already seen are definite at (a), (c), (d) above. But $\mathrm{Succ}(x) \leftrightarrow \exists y(x = y \cup \{y\}) \leftrightarrow \exists y \in x(x = y \cup \{y\})$. We note that $y \cup \{y\}$ is definite, and so by the Theorem 3.3 $\mathrm{Succ}(x)$ is definite. The three conditions are mutually exclusive we can appeal to the last lemma once we note that the three terms $\varnothing$, $y \cup \{y\}$, and $\bigcup z$ where z is a definite set by 3.1 (iv) in place of t_1, t_2,and t_3 are definite. The other functions are exactly the same. Q.E.D.

Note (1): $\mathcal{P}(x)$ is not definite: if it were we could conclude from Theorem 3.3 that for any transitive set satisfying $(\mathrm{ZF}^-)^W$ that $\mathcal{P}(x) \in W$ which is not true in general.

"x is countable" cannot be expressed by a definite formula $\varphi(x)$: again if it were, we should have that the concept is absolute for transitive W satisfying $(\mathrm{ZF}^-)^W$. We list some definite concepts.

LEMMA 3.7 *For any n: (i) $\bigcup^n x$, (ii) $\{x_1, \ldots x_n\}$, (iii) $\langle x, y\rangle$; $(u)_0$, $(u)_1$ where $u = \langle (u)_0, (u)_1\rangle$; (iv) $\langle x_1, \ldots, x_n\rangle$, (v) $x \times y$,(vi) ran(z), (vii) dom(z), (viii) $z``x$, (ix) $z \restriction x$, (x) z^{-1} are all definite terms.*
The following relations are definable by definite formulae:

(xi) $x \subseteq y$; (xii) $\mathrm{Trans}(y)$; (xiii) $\mathrm{Rel}(z)$; $\mathrm{Fun}(z)$; (xiv) $z(x) = y$; (xv) "z is a (1-1) function"; z is an onto function; (xvi)"x is unbounded in β"; "$z : \alpha \longrightarrow \beta$ is a cofinal function"; "$x \subseteq \beta$ is a closed and unbounded set";

(xvii) the terms $\mathrm{TC}(x)$, (xviii) $\rho(x)$ are definite terms.
Thus all the above are a.d.

PROOF: The first two are simply repeated applications of operations defined to be definite. Similarly (iii) $\langle x, y\rangle = \{\{x\}, \{x, y\}\}$; (iv) $\langle x_1, \ldots, x_n\rangle$ was defined by repeated application of $\langle -, -\rangle$ and hence is definite; (v) $x \times y = \bigcup\{x \times \{z\} | z \in y\} = \bigcup\{\{\langle w, z\rangle | w \in x\} | z \in y\}$.

(vi): $\mathrm{ran}(z) = \{u \in \bigcup^2 z | \exists w \in z \exists v \in \bigcup^2 z (w = \langle v, u\rangle\}$;

(vii): $\mathrm{dom}(z) = \{u \in \bigcup^2 z | \exists w \in z \exists v \in \bigcup^2 z (w = \langle u, v\rangle\}$;

(viii): $z``x = \{v \in \bigcup^2 z \mid \exists u \in x \exists w \in z(w = \langle u, v\rangle)\}$;

(ix), (x) Exercise.

(xi): $x \subseteq y \leftrightarrow \forall z \in x(z \in y)$, it is thus Δ_0 and so definite;

(xii) $\mathrm{Trans}(y) \leftrightarrow \forall z \in y(z \subseteq y)$;

(xiii) $\mathrm{Rel}(z) \leftrightarrow z \subseteq \mathrm{dom}(z) \times \mathrm{ran}(z)$;

$\mathrm{Fun}(z) \leftrightarrow \mathrm{Rel}(z) \wedge \forall x \in \mathrm{dom}(z) \forall u, v \in \mathrm{ran}(z)(v \neq u \longrightarrow (\langle x, u\rangle \in z \leftrightarrow \langle x, v\rangle \notin z))$;

(xv) $z(x) = y \leftrightarrow \mathrm{Fun}(z) \wedge \langle x, y\rangle \in z$;

(xvi) Exercise.

(xvii) $\mathrm{TC}(x) = t(x, \{\mathrm{TC}(y) | y \in x\})$ for a definite t using the definite recursion scheme.

(xviii): $s(z) = z \cup \{z\}$ is definite; then $\{s(v) \mid v \in u\}$ is definite as then is $t(u) = \bigcup\{s(v) \mid v \in u\}$ (by (iv) and (ii) resp. in Def.3.1). Using the definite recursion scheme (vi) we get $\rho(x) = t(\{\rho(y) \mid y \in x\})$. (Here we are just expressing that $\rho(x) = \sup\{\rho(y) + 1 \mid y \in x\}$.) Q.E.D.

EXERCISE 3.1 (i) Show that "*x is a total order of y*" can be expressed in a Δ_0 fashion.

(ii) Complete (ix),(x), (xvi), (xvii) of Lemma 3.7.

LEMMA 3.8 *The following are definite: ${}^n x$ (for any n); ${}^{<\omega} x =_{\mathrm{df}} \bigcup\{{}^n x | n \in \omega\}$; "x is finite". Hence $\mathcal{P}_{<\omega}(z) =_{\mathrm{df}} \{x \subseteq z | x \text{ is finite}\}$ is definite.*

PROOF: By induction on n : we define $F(n, x) = {}^n x$:

$F(0, x) = {}^0 x = \varnothing$.

$F(n + 1, x) = {}^{n+1} x = \{f \cup \{\langle n, y\rangle\} | f \in F(n, x) \wedge y \in x\}$;

$F(\omega, x) = {}^{<\omega} x = \bigcup\{F(n, x) | n \in \omega\}$.

This is given by definite recursion clauses, and so $F(n, x)$ is definite for $n \leq \omega$.

"x is finite"$\leftrightarrow \exists f \in {}^{<\omega}x(f$ is onto). And then:
$\{x \subseteq z|\ x$ is finite$\} = \{x|\exists f \in {}^{<\omega}z(x = \mathrm{ran}(f))\}$ Q.E.D.

Note: the absoluteness of finiteness implies that if $\mathrm{Trans}(W) \wedge (\mathrm{ZF}^-)^W$ then any finite subset of W is in W. This need not be true of course for infinite subsets of W.

EXERCISE 3.2 Suppose $\mathrm{Trans}(W)\wedge(\mathrm{ZF}^-)^W$. Show $(V_\alpha)^W = V_\alpha \cap W$. [Hint: use that the rank function is definite.]

Note: "$\mathrm{cf}(\alpha)$" along with "*x is a cardinal*" or "ω_1" are not definite, and so not absolute for such W in general (but see the next exercise). Neither then is "*x is a regular/singular cardinal*." However being a wellorder is so absolute as the next lemma shows.

EXERCISE 3.3 Let λ be a limit ordinal; show that the following are absolute for V_λ: (i) $\mathcal{P}(x)$ (ii) "α is a cardinal" (and hence $(Card)^{V_\lambda} = Card \cap \lambda$); (iii) $\mathrm{cf}(\alpha)$ (iv) "α is (strongly) inaccessible" (v) $y = V_\alpha$ (vi) $\aleph_\alpha$ (vii) $\beth_\alpha$.

LEMMA 3.9 *(i) "z is a wellorder of y"; (ii) "z is a wellfounded relation on y" are absolutely definite.*

PROOF: Suppose $\mathrm{Trans}(W) \wedge (\mathrm{ZF}^-)^W$, $z, y \in W$. For (i) "*z is a total order of y*" can be expressed in a Δ_0 way (Exercise). Suppose ("*z is a wellorder of y*")^W. Since we have Ax . Replacement holding in W we have that "*$\langle y, z\rangle$ is isomorphic to an ordinal*" holds in W. If $(\alpha \in \mathrm{On})^W$ and $(f : \langle y, z\rangle \cong \langle \alpha, <\rangle)^W$ then $\mathrm{dom}(f) = y$, $\mathrm{ran}(f) = \alpha$, "f is a bijection", *etc.*, are all absolute for W. Hence $f : \langle y, z\rangle \cong \langle \alpha, <\rangle$ holds in V. Consequently $\langle y, z\rangle$ is truly a wellorder.

Conversely if "*z is a wellorder of y*" with $z, y \in W$, then as for any $w \in V$ with $w \subseteq y$ we have w has a z-minimal element w_0 say, then $w_0 \in W$ (as $\mathrm{Trans}(W)$) and no $u \in W$ satisfies uzw_0. So if also $w \in W$ then ("*w_0 is an z-minimal element of w*")^W.

(ii) is only an amplification of (i), effected by defining an absolute rank function ρ_z of the wellfounded relation z. We leave this to the reader. Q.E.D.

The example of wellorder shows that being expressible by a Δ_0 formula is not a necessary condition for absoluteness: wellorder in general is a Π_1-concept when literally written out. However if $(\mathrm{ZF}^-)^W$ holds then we have Ax.Replacement available to turn this Π_1 concept into an existential Σ_1 statement and hence have that it is U-absolute for W. We may say that it is thus "$\Delta_1^{\mathrm{ZF}^-}$". If W is not a model of sufficient Replacement then this argument can fail.

EXERCISE 3.4 "$y = V_\alpha$" is Π_1, and "$\rho(x) < \alpha$" is Δ_1 expressible.

3.1.1 THE NON-FINITE AXIOMATISABILITY OF ZF

We use the Reflection Theorem together with our absoluteness results to prove the non-finite axiomatisability of ZF. (We say a set of axioms T *axiomatises* S if $T \vdash \sigma$ for every σ from S. A set S is *finitely axiomatisable* if there is a finite set T that axiomatises S.)

THEOREM 3.10 ***(The non-finite axiomatisability of*** **ZF*)*** *Let T be any set of axioms in $\mathcal{L}$, extending* ZF*, and T_0 be any finite subset of T; if from T_0 we can prove every axiom of T then T is inconsistent.*

In particular, with T as ZF*, no finite subset of* ZF *axioms will axiomatise all of* ZF*, unless* ZF *is inconsistent.*

Proof: Suppose $T \supseteq T_0$ were such sets of axioms, with all of T provable from T_0, for a contradiction. We have the assertion: $\mathrm{ZF} \vdash \forall\alpha\exists\beta > \alpha((\bigwedge\!\!\!\bigwedge T_0)^{V_\beta} \leftrightarrow \bigwedge\!\!\!\bigwedge T_0)$. Then as T_0 proves every axiom of ZF, it proves the following instance of the Reflection Theorem:

$$T_0 \vdash \forall\alpha\exists\beta > \alpha((\bigwedge\!\!\!\bigwedge T_0)^{V_\beta} \leftrightarrow \bigwedge\!\!\!\bigwedge T_0).$$

However trivially $T_0 \vdash \forall\alpha\exists\beta > \alpha((\bigwedge\!\!\!\bigwedge T_0)^{V_\beta})$
since $T_0 \vdash \bigwedge\!\!\!\bigwedge T_0$. Then, by the principle of ordinal induction:

$$T_0 \vdash \exists\beta_0[(\bigwedge\!\!\!\bigwedge T_0)^{V_{\beta_0}} \wedge \forall\delta < \beta_0(\neg(\bigwedge\!\!\!\bigwedge T_0)^{V_\delta}]. \quad (*)$$

We are assuming that T_0 proves all of ZF, so by the Soundness of first order predicate logic, Theorem 1.20, in the form that if $T_0 \vdash \psi$ and $(\bigwedge\!\!\!\bigwedge T_0)^{V_\gamma}$, then $(\psi)^{V_\gamma}$, we may deduce, $T_0 \vdash (\mathrm{ZF})^{V_{\beta_0}}$.

Then all our absoluteness results about transitive models hold for V_{β_0} for such a β_0 as in $(*)$. Also in particular :

$$T_0 \vdash \beta < \beta_0 \to (V_\beta)^{V_{\beta_0}} = V_\beta \cap V_{\beta_0} = V_\beta \text{ (Exercise 3.2)}$$

Again using Soundness, since $T_0 \vdash \exists\beta(\bigwedge\!\!\!\bigwedge T_0)^{V_\beta}$, and at $(*)$ we have $T_0 \vdash (\bigwedge\!\!\!\bigwedge T_0)^{V_{\beta_0}}$:

$$T_0 \vdash (\exists\beta(\bigwedge\!\!\!\bigwedge T_0)^{V_\beta})^{V_{\beta_0}}.$$

However, then we have
$T_0 \vdash \exists\beta < \beta_0(\bigwedge\!\!\!\bigwedge T_0)^{V_\beta}$ which contradicts $(*)$. So T_0 and hence T is inconsistent. Q.E.D.

3.2 Formalising syntax

We shall consider the language $\mathcal{L} = \mathcal{L}_{\dot\in,\dot=}$ that we have been using to date, that can be interpreted in $\in$-*structures*, that is any structure $\langle X, E\rangle$ with a domain a class of sets X and an interpretation E for the $\dot\in$ symbol. In what follows, we shall almost always be considering the *standard interpretation* of the $\dot\in$ symbol, where it is interpreted as the true set membership relation. The equality symbol $\dot=$ will without exception be interpreted as true equality $=$. Up to now the object language of our ZF theory has been floating free from our universe of sets, but we shall see how this language (indeed any reasonably given language) can be *represented* by using sets, just as we can represent the natural numbers $0, 1, 2, \ldots$ by the sets $\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \ldots$ We make therefore a choice of *coding* of the language $\mathcal{L}$ by sets in V. The method of coding itself is not terribly important, there are many ways of doing this, but the essential feature is that we want a mapping of the language into a class of sets, where the latter is ZF (in fact ZF^- or even much more simply) definable). As we are mainly interested in the first order language $\mathcal{L}$ we give the definitions in detail just for that. In principle we could do this for any language, for any structures.

Definition 3.11 *(Gödel code sets) We define by (a meta-theoretic) recursion on the structure of formulae φ of $\mathcal{L}$ the* code set $\ulcorner\varphi\urcorner \in V_\omega$.

(i) $\ulcorner v_i \dot= v_j\urcorner$ *is* $2^{i+1}\cdot 3^{j+1}$; $\ulcorner v_i \dot\in v_j\urcorner$ *is* $5^{i+1}\cdot 7^{j+1}$;

(ii) $\ulcorner\chi \vee \psi\urcorner$ *is* $\langle\ulcorner\chi\urcorner, \ulcorner\psi\urcorner\rangle$;

(iii) $\ulcorner\neg\psi\urcorner$ *is* $\langle 0, \ulcorner\psi\urcorner\rangle$;

(iv) $\ulcorner\exists v_i\psi\urcorner$ *is* $\langle 11^{i+1}, \ulcorner\psi\urcorner\rangle$.

Note (a) atomic formulae are the only ones coded by integers, (b) in each case, that if φ is non-atomic then the code set contains immediate subformula(e) codes as direct members. (c) If we had wanted to have further predicate symbols besides $\in$ in our language, *e.g.* monadic predicates $A_k(v_n)$ we could have added as codes $13^{k+1} \cdot 17^{n+1}$, and similarly for j-place predicates. The means of coding are completely flexible and in this setting any reasonable system can work.

Clearly given a code set u we may decode φ from it in a unique fashion, making use of the primes and the prime power coding. We give the formal counterpart of the above definition using finite functions from V_ω, a definite formula defining the characteristic function of the class of code sets of formulae of $\mathcal{L}$:

Definition 3.12 *(i)* $\mathrm{Fml}(u, f, n) = 1 \leftrightarrow f \in {}^{<\omega}V_\omega \wedge \mathrm{dom}(f) = n + 1 \wedge f(n) = u \wedge$
$\wedge \forall k \in \mathrm{dom}(f)[\exists i, j \in \omega(f(k) = 2^{i+1} \cdot 3^{j+1} \vee f(k) = 5^{i+1} \cdot 7^{j+1} \vee$
$\exists m, l < k[f(k) = \langle f(m), f(l)\rangle \vee f(k) = \langle 0, f(m)\rangle \vee \exists i \in \omega(f(k) = \langle 11^{i+1}, f(m)\rangle)]]\ )$;
$\mathrm{Fml}(u, f, n) = 0$ otherwise.
(ii) $\mathrm{Fmla}(u) = 1 \leftrightarrow \exists n \in \omega \exists f \in {}^{<\omega}V_\omega\, \mathrm{Fml}(u, f, n) = 1$; $\mathrm{Fmla}(u) = 0$ *otherwise.*

We thus may think of the formula φ as represented by, or coded by, $f(n)$, where f is the function that describes its construction according to the last definition, with $\mathrm{dom}(f) = n + 1$.

It should be noted that both the last definitions are built up using definite terms, and so are defined by definite formulae and thus are a.d.

3.3 Formalising the satisfaction relation

We now formalise the (first order) *satisfaction relation* due to Tarski, familiar from model theory.

Definition 3.13
(i) $Q_x =_{\mathrm{df}} \{h | \mathrm{Fun}(h) \wedge \mathrm{dom}(h) = \omega \wedge \mathrm{ran}(h) \subseteq x \wedge \exists n \in \omega \exists y \in x(\forall m \geq n h(m) = y)\}$.
(ii) If $h \in Q_x$, *and* $y \in x$, *then* $h(y/i)$ *is the function defined by:*

$$\forall j \in \omega(j \neq i \longrightarrow h(y/i)(j) = h(j)) \ \wedge\ h(y/i)(i) = y.$$

Again Q_x is definite: we may write

$$h \in Q_x \leftrightarrow \exists h_0 \in {}^{<\omega}x \exists y \in x(h = h_0 \cup \{\langle n, y\rangle | n \in \omega \wedge \mathrm{dom}(h_0) \leq n\}).$$

Thus although Q_x does not contain *finite* functions, any $h \in Q_x$ is essentially a finite function with a constant tail - and this makes it definite. (Again: ${}^\omega x$, like $\mathcal{P}(x)$, is *not* definite.) (ii) also specifies a definite relation between i, x, y, and h.

We next specify what it means for a finite function h to be an assignment of variables potentially occurring in a formula u to objects in x that makes u come out true in the structure $\langle x, \in\rangle$.

Definition 3.14 *(i) We define by recursion the term* $\mathrm{Sat}(u,x)$;

$$\begin{array}{rcl}
\mathrm{Sat}(\ulcorner v_i \dot{=} v_j \urcorner, x) & = & \{h \in Q_x | h(i) = h(j)\}; \\
\mathrm{Sat}(\ulcorner v_i \dot{\in} v_j \urcorner, x) & = & \{h \in Q_x | h(i) \in h(j)\}; \\
\mathrm{Sat}(\ulcorner \chi \vee \psi \urcorner, x) & = & \mathrm{Sat}(\ulcorner \chi \urcorner, x) \cup \mathrm{Sat}(\ulcorner \psi \urcorner, x)\}; \\
\mathrm{Sat}(\ulcorner \neg\psi \urcorner, x) & = & Q_x \setminus \mathrm{Sat}(\ulcorner \psi \urcorner, x)\}; \\
\mathrm{Sat}(\ulcorner \exists v_i \psi \urcorner, x) & = & \{h \in Q_x | \exists y \in x (h(y/i) \in \mathrm{Sat}(\ulcorner \psi \urcorner, x))]\}; \\
\mathrm{Sat}(u, x) & = & \varnothing \textit{ if } \mathrm{Fmla}(u) = 0.
\end{array}$$

(ii) We write $\langle x, \in \rangle \models u[h]$ *iff* $h \in \mathrm{Sat}(u,x)$.

Note: By design then we have $\langle x, \in \rangle \models \ulcorner \neg\psi \urcorner[h]$ iff it is not the case that $\langle x, \in \rangle \models \ulcorner \psi \urcorner[h]$ etc. (We write the latter as $\langle x, \in \rangle \not\models \ulcorner \psi \urcorner[h]$.) If, uninterestingly, $x = \varnothing$ then also $\mathrm{Sat}(u,x) = \varnothing$.

Lemma 3.15 $\mathrm{Sat}(u,x)$ *is defined by a definite recursion. Hence "*$\langle x, \in \rangle \models \ulcorner \varphi \urcorner[h]$*" is definite.*

Proof: This should be pretty clear, but we give an explicit recursive term t for Sat:

$\mathrm{Sat}(u,x) = \{h \in Q_x \mid \mathrm{Fmla}(u) = 1 \wedge$
$\exists i, j \in \omega[(u = 2^{i+1} \cdot 3^{j+1} \wedge h(i) = h(j)) \vee (u = 5^{i+1} \cdot 7^{j+1} \wedge h(i) \in h(j))]]\vee$
$\vee[\exists v \in \bigcup u(h \in \mathrm{Sat}(v,x))] \vee [0 \in \bigcup u \wedge \exists v \in \bigcup u(v \neq 0 \wedge h \notin \mathrm{Sat}(v,x)]\vee$
$\vee[\exists i \in \omega(11^{i+1} \in \bigcup u \wedge \exists v \in u \exists y \in x(h(y/i) \in \bigcup \mathrm{Sat}(v,x)\ ]\}$.

The specification here yields a definite term $\mathrm{Sat}(u,x) = t(x,u,\{\mathrm{Sat}(v,x)|v \in u\})$ noting that we have already established that all the concepts appearing here, such as "Q_x","$\mathrm{Fmla}(u)$", "ω", *etc.* are definite. Q.E.D.

By our work so far then then we may say that "the assignment h makes the formula φ true in the structure $\langle x, \in \rangle$" if $\langle x, \in \rangle \models \ulcorner \varphi \urcorner[h]$. Otherwise we say it is similarly "false".

If φ is a formula of $\mathcal{L}$ with free variables amongst $v_{j_0}, \ldots, v_{j_n}$ and $y_0, \ldots, y_n \in x$ then we abbreviate:
$\langle x, \in \rangle \models \ulcorner \varphi \urcorner[y_0, \ldots, y_n] \longleftrightarrow \langle x, \in \rangle \models \ulcorner \varphi \urcorner[h]$ for any $h \in Q_x$ with $h(j_i) = y_i$ all $i \leq n$.
This makes perfect sense, since the intepretation of the formula φ in the structure only depends on the assignment to the free variables of φ. If φ has no free variables at all, then it is deemed a sentence and either $\mathrm{Sat}(\ulcorner \varphi \urcorner, x) = Q_x$, in which we case we say the sentence φ is true in $\langle x, \in \rangle$ or else $\mathrm{Sat}(\ulcorner \varphi \urcorner, x) = \varnothing$ in which case it is false. In each case we simply write $\langle x, \in \rangle \models \ulcorner \varphi \urcorner$ or $\langle x, \in \rangle \not\models \ulcorner \varphi \urcorner$ accordingly, as then assignment functions h are superfluous.

3.4 Formalising definability: the function Def.

The following is the crucial function used to build up definable sets.

Definition 3.16 $\mathrm{Def}(x) =_{\mathrm{df}} \{\{w \in x \mid \langle x, \in \rangle \models u[h(w/0)]\}; \mathrm{Fmla}(u) = 1 \wedge h \in Q_x\}$.

Lemma 3.17 *"*$\mathrm{Def}(x)$*" is a definite term.*

Proof: First note that we have shown that "$\langle x, \in\rangle \models u[h(w/0)]$" is definite. Hence so is

$$\iota(x,u,h) =_{\text{df}} \{w \in x | \langle x, \in\rangle \models u[h(w/0)]\}.$$

Hence $\{\iota(x,u,h) | \text{Fmla}(u) = 1 \wedge h \in Q_x\}$ is definite. Q.E.D.

The class $\text{Def}(x)$ we think of as the "definable power set of x": it consists of those subsets $y \subseteq x$ so that membership in y is given by a formula $\varphi(v_0, v_1, \ldots, v_m)$ all of whose free variables are amongst those shown, together with a *fixed* assignment of some $y_1, \ldots, y_m$, and the members $y_0 \in y$ are determined by allowing v_0 to range over all of x. Those y_0 that when added to the fixed assigment $y_1, \ldots, y_m$, cause $\varphi[y_0, y_1, \ldots, y_m]$ to come out true in $\langle x, \in\rangle$ are then added to y. We may write slightly more informally:

$$\text{Def}(x) = \{z \,|\, z = \{w | \langle x, \in\rangle \models \varphi[w, y_1, \ldots, y_m]\}, \text{Fmla}(\varphi) = 1,\ \vec{y} \in {}^{<\omega}x\}$$

where it is implicitly understood that we should have written $\ulcorner\varphi\urcorner$ for φ and it is left unsaid that the free variables of φ have all been assigned some value in x by the assignment displayed.

Lemma 3.18 *(i)* $x \in \text{Def}(x)$; *(ii)* $\text{Trans}(x) \longrightarrow x \subseteq \text{Def}(x)$;
(iii) $\forall z \subseteq x(|z| < \omega \to z \in \text{Def}(x))$;
(iv) (AC) $|x| \geq \omega \longrightarrow |\text{Def}(x)| = |x|$.

Proof: (i) $x = \{w \,|\, \langle x, \in\rangle \models \ulcorner v_0 = v_0\urcorner[w]\}$ and so $x \in \text{Def}(x)$.

(iv) Assume x is infinite. Then Q_x has the same cardinality as ${}^{<\omega}x$, namely $|x|$. Also, $F =_{\text{df}} \{u \mid \text{Fmla}(u) = 1\}$ is a countable set. Since $\text{Def}(x)$ is the class of subsets of x given by a definition involving a formula $u \in \text{Fmla}$ together with a finite parameter string $y_1, \ldots, y_n$ we see that: $|\text{Def}(x)| \leq |F|.|Q_x| = \omega.|x| = |x|$. That $|x| \leq |\text{Def}(x)|$ follows from (iii). (ii) and (iii) are left as an exercise. Q.E.D.

Exercise 3.5 Finish (ii) and (iii) of Lemma 3.18.

Exercise 3.6 Let $\langle x, \in\rangle$ be a transitive $\in$-model. Show that $\text{Trans}(\text{Def}(x))$. If $y, z \in x$ then is $\langle y, z\rangle \in \text{Def}(x)$? Is $\{x\}$? [Hint (for the last question): If $\rho(x) = \alpha$, compute $\rho(\text{Def}(x))$ and compare this with the given sets.]

Exercise 3.7 Let us say that w is *outright definable in the set* $\langle x, \in\rangle$ if for some formula φ with only free variable v_0 then w is the unique element in x so that $\langle x, \in\rangle \models \varphi[w]$. We may thus define a variant on the Def function by:

$$\text{Def}_0(x) = \{z \,|\, \{z\} = \{w \in x | \langle x, \in\rangle \models \varphi[w]\}, \text{Fmla}(\varphi) = 1, \text{FVbl}(\varphi) = \{v_0\}\ , w \in x\}$$

of the sets outright definable in $\langle x, \in\rangle$, definable without use of parameters. Show that $|\text{Def}_0(x)| \leq \omega$ for any x.

Definition 3.19 *We say that a set z is* ordinal definable* *("$z \in OD^*$") if for some β, $z \in \text{Def}_0(V_\beta)$.*

(This definition is just a placeholder for the official - but equivalent - definition of ordinal definability to come.)

Exercise 3.8 (i) Show that: (a) $On \subseteq OD^*$; (b) $\forall\beta V_\beta \in OD^*$; (c) $\forall x(x \in OD^* \to \{x\} \in OD^*)$. (ii)($*$) Show that there is a (countable) set X so that for unboundedly many ordinals β in On, $X \in \text{Def}_0(V_\beta)$. [Hint: consider the theory of each V_β: the set of all codes of sentences σ so that $\langle V_\beta, \in\rangle \models \ulcorner\sigma\urcorner$. This is a subset of V_ω.]

3.5 More on correctness and consistency

The next theorem illustrates that our definitions are 'correct': we have formulated two ways of talking about a statement φ being 'true in a structure' W, firstly we considered relativised formulae and spoke from an exterior perspective of 'φ holds or is true in W' by asserting 'φ^W'. The formula φ from $\mathcal{L}$ we consider to be in our language in which we wish to state our axioms about the structure consisting of our intuitive universe of sets. We have now a second interior method through the formalised version of the language which consists of sets coding formulae as for $\ulcorner\varphi\urcorner$ above together with the satisfaction relation. This relation was between (codes of) formulae and structures or 'models'. The next theorem asserts that these two methods are in harmony.

Theorem 3.20 **(Correctness Theorem)** *Suppose φ is a formula of $\mathcal{L}$ with free variables $\vec{v} = v_{j_1}, \ldots, v_{j_m}$ then:*

$$\mathrm{ZF}^- \vdash \forall x \forall \vec{y} \in {}^m x[(\langle x, \in\rangle \models \ulcorner\varphi\urcorner[\vec{y}/\vec{v}]) \longleftrightarrow (\varphi(\vec{y}/\vec{v}))^{\langle x,\in\rangle}].$$

• This would be a proof by induction on the complexity of φ (we shall omit the details). It is again a *theorem scheme*, being one theorem for each φ.

The ZF and ZFCaxiom collections themselves have formal counterparts as sets: just as each formula φ is mapped to its code set $\ulcorner\varphi\urcorner$ as above, we can also find *sets* that collect together the code sets of those sentences φ that are axioms of ZF (or ZFC). Namely, there is an algorithm for listing the axioms of ZFas $\varphi_0, \varphi_1, \ldots, \varphi_n, \ldots$

Definition 3.21 $\ulcorner\mathrm{ZF}\urcorner =_{\mathrm{df}} \{u | \mathrm{Fmla}(u) = 1 \wedge (\mathrm{Ax}\,0(u) \vee \mathrm{Ax}\,1(u) \vee \cdots \vee \mathrm{Ax}\,8(u))\}$.
$\ulcorner\mathrm{ZFC}\urcorner$ *is defined similarly by adding* "$\vee\, \mathrm{Ax}\,9(u)$".

In the above by '$\mathrm{Axj}(u)$' we mean that u is a code set for an axiom of the type Axj written out in our official language. Thus Axo is the Ax.Extensionality. If this latter axiom is written out using only $\neg$, $\vee, \exists$ *etc.* as σ then we have $\mathrm{Ax}\,0(u) \longleftrightarrow u = \ulcorner\sigma\urcorner$. The other axioms similarly must be written out in the formal language, and then coded according to our prescription. Some axioms are in fact axiom schemata: infinite sets of axioms. So for $\mathrm{Ax}\,6(u)$ (for Ax.Replacement) we should demand that u conforms to the right shape of formula that is an instance of the axiom of replacement when written out in this correct manner. $\mathrm{Ax}\,6(u)$ will then be an infinite set, as will be $\ulcorner\mathrm{ZF}\urcorner$.

Lemma 3.22 *(i) "$u \in \ulcorner\mathrm{ZF}\urcorner$", "$u \in \ulcorner\mathrm{ZFC}\urcorner$" are definite. (ii) If φ is an axiom of* ZF *then*

$$\mathrm{ZF}^- \vdash \ulcorner\varphi\urcorner \in \ulcorner\mathrm{ZF}\urcorner.$$

Similarly if φ is an axiom of ZFC *then* $\mathrm{ZF}^- \vdash \ulcorner\varphi\urcorner \in \ulcorner\mathrm{ZFC}\urcorner$.

This would again be a proof by induction on the structure of φ. The intuitive meaning that it captures is that "$\mathrm{ZF} \subseteq \ulcorner\mathrm{ZF}\urcorner$". The point again is that the definitions of $\ulcorner\mathrm{ZF}\urcorner$ and $\ulcorner\mathrm{ZFC}\urcorner$ are again *definite*. These details are uninteresting and somewhat tedious, but the *idea* that this can be done is very interesting. (ii) is again a theorem scheme, one for each axiom φ.

Definition 3.23 "$\langle x, \in\rangle \models \ulcorner\mathrm{ZF}\urcorner$"$\Longleftrightarrow_{\mathrm{df}} \forall u \in \ulcorner\mathrm{ZF}\urcorner\, \langle x, \in\rangle \models u$. *("$\langle x, \in\rangle \models \ulcorner\mathrm{ZFC}\urcorner$" similarly.)*

We have that, *e.g.* "$\langle x, \in\rangle \models \ulcorner \mathrm{ZF} \urcorner$" and "$\langle x, \in\rangle \models \ulcorner \mathrm{ZFC} \urcorner$" are definite, and so a.d. Then in this case we say that $\langle x, \in\rangle$ "is a model of ZF(C)".

Corollary 3.24 *(to the Correctness Theorem) For φ any axiom of* ZF^- *then*

$$\mathrm{ZF}^- \vdash \forall x(\langle x, \in\rangle \models \ulcorner \mathrm{ZF}^- \urcorner \longrightarrow \varphi^{\langle x, \in\rangle});$$

similarly for φ any axiom of ZFC

$$\mathrm{ZF}^- \vdash \forall x(\langle x, \in\rangle \models \ulcorner \mathrm{ZFC} \urcorner \longrightarrow \varphi^{\langle x, \in\rangle}).$$

Exercise 3.9 Suppose κ is strongly inaccessible. Verify that $\langle V_\kappa, \in\rangle \models \ulcorner \mathrm{ZFC} \urcorner$.

Exercise 3.10 $(*)$ (E) Let $\mathcal{A}, \mathcal{B}$ be structures. We write $\mathcal{A} \prec \mathcal{B}$ if for every formula u, every $h \in Q_{\mathcal{A}}$ if $\mathcal{B} \models u[h]$ then $\mathcal{A} \models u[h]$. Suppose that κ, λ are such that $\langle V_\kappa, \in\rangle \prec \langle V_\lambda, \in\rangle$. Show that κ is a strong limit cardinal and that both $\langle V_\kappa, \in\rangle, \langle V_\lambda, \in\rangle$ are models of ZFC.

Exercise 3.11 $(*)$ (E) Suppose there is λ which is strongly inaccessible. Show that there is κ with $\langle V_\kappa, \in\rangle$ a model of ZFC, and with $\mathrm{cf}(\kappa) = \omega$. [Hint: Use the Reflection Theorem proof on V_λ, which we now have assumed to be a ZFC model, to show that every formula φ of ZFC now "reflects" down to a cub $C_\varphi \subseteq \lambda$ set of ordinals. Now intersect over all φ. This method shows that in fact there is a cub set of points $\kappa < \lambda$ with $\langle V_\kappa, \in\rangle$ not only a model of ZFC, but also $\langle V_\kappa, \in\rangle \prec \langle V_\lambda, \in\rangle$]

Exercise 3.12 Suppose $\langle X, \in\rangle \models T$ for some set of sentences T including Ax.Ext. Show that there is a countable transitive x with $\langle x, \in\rangle \models T$. [Hint: The Downward-Löwenheim Skolem Theorem says for any cardinal λ with $\omega \le \lambda \le |X|$ there is a Y with $\langle Y, \in\rangle \prec \langle X, \in\rangle$ and $|Y| = \lambda$. Then use the Mostowski-Shepherdson Collapsing Lemma.] In particular if there is an $\in$-structure which is a model of ZFC then there is a countable transitive one.

3.5.1 Incompleteness and Consistency Arguments

In general when we say that a theory T is *consistent* we mean that for no sentence σ do we have $T \vdash \sigma$ and $T \vdash \neg\sigma$. We abbreviate this as "$\mathrm{Con}(T)$". Of course if T is inconsistent then we may prove anything at all from T and we can then say (assuming that T is in a language in which we formulate arithmetic axioms) that "$T \vdash 0 = 1$" encapsulates the notion that T is inconsistent. The heart of Gödel's argument is that it is possible to formulate the concept of a formal proof from an algorithmically or recursively given axiom set T extending PA, Peano Arithmetic, in such a way that "v_0 *codes a proof from* $\ulcorner T \urcorner$ *of* v_1", abbreviated $\mathrm{Pf}^T(v_0, v_1)$, can be *represented* in the theory T. Then we may use "$\neg\exists v_0\, \mathrm{Pf}^T(v_0, \ulcorner 0 = 1 \urcorner)$", abbreviated as "$\mathrm{Con}^T$", to capture the formal assertion that T is consistent. He then showed that $T \nvdash \mathrm{Con}^T$. In short we thus formalise the notions of "proof", "contradiction", "axiom" *etc. within* the theory T, starting with the formalisation of syntax that we have already effected. We are not going here to go down the route of investigating Gödel's proof in its entirety, however we can rather easily obtain a weak version of Gödel's Second Incompleteness Theorem which suffices for our purposes. (Compare the proof of Theorem 3.10)

THEOREM 3.25 **(Gödel)** $\mathrm{Con(ZF)} \Rightarrow \mathrm{ZF} \nvdash \exists x(\mathrm{Trans}(x) \wedge \langle x, \in\rangle \models \ulcorner \mathrm{ZF} \urcorner)$.

PROOF: Suppose σ abbreviates the sentence $\exists x(\mathrm{Trans}(x) \wedge \langle x, \in\rangle \models \ulcorner \mathrm{ZF} \urcorner)$. Suppose that $\mathrm{ZF} \vdash \sigma$. Then:
$\mathrm{ZF} \vdash \exists z(\mathrm{Trans}(z) \wedge \langle z, \in\rangle \models \ulcorner \mathrm{ZF} \urcorner \wedge$" $\forall w(\mathrm{Trans}(w) \wedge \rho(w) < \rho(z) \to \langle w, \in\rangle \not\models \ulcorner \mathrm{ZF} \urcorner)$" $(*)$

Let z satisfy the last formula. By the last Corollary for any axiom φ of ZF we have $\varphi^{\langle z,\in\rangle}$. That is $(\mathrm{ZF})^{\langle z,\in\rangle}$. As $\mathrm{ZF} \vdash \sigma$ we shall have that $(\sigma)^{\langle z,\in\rangle}$. In other words $(\exists x(\mathrm{Trans}(x) \wedge \langle x, \in\rangle \models \ulcorner \mathrm{ZF} \urcorner))^{\langle z,\in\rangle}$. So let $y \in z$ satisfy this, namely

$(\mathrm{Trans}(y) \wedge \langle y, \in\rangle \models \ulcorner \mathrm{ZF} \urcorner)^{\langle z,\in\rangle}$.

But this is a definite formula and so is absolute for the transitive structure $\langle z, \in\rangle$ as $(\mathrm{ZF}^-)^{\langle z,\in\rangle}$. Hence we really do have:

$y \in z \wedge \mathrm{Trans}(y) \wedge \langle y, \in\rangle \models \ulcorner \mathrm{ZF} \urcorner$.

But $\rho(y) < \rho(z)$. This contradicts $(*)$. Hence ZF is inconsistent. Q.E.D.

However, providing we have done our formalisation of $\ulcorner \mathrm{ZF} \urcorner$ and $\mathrm{Pf}^T(v_0, v_1)$ *etc.* sensibly, we shall have that in ZF we can prove the Gödel *Completeness Theorem*: that any consistent set of sentences in any first order theory whatsoever has a model, and thus shall have:

$\mathrm{ZF} \vdash$"$\mathrm{Con}^{\mathrm{ZF}} \longrightarrow \exists X, E[|X| = \omega \wedge \langle X, E\rangle \models \ulcorner \mathrm{ZF} \urcorner]$" $(**)$

But there is no indication that E should be the natural set membership relation on the countable set X, or that $\mathrm{Trans}(X)$. X, E arise simply from the proof of the Completeness Theorem. In general E will not be wellfounded, and will be completely artificial.

Taking this line further: if there is a set which is a transitive model of ZF, let us assume $\langle x, \in\rangle \in V$ is such. We additionally assume such an x is chosen of least rank. The assumed existence of $\langle x, \in\rangle$ implies $\mathrm{Con}^{\mathrm{ZF}}$, and as this latter assertion is expressed as a definite sentence, and $\langle x, \in\rangle$ is a transitive ZF^- model, we have $(\mathrm{Con}^{\mathrm{ZF}})^{\langle x,\in\rangle}$. By $(**)$ $(\exists X, E(|X| = \omega \wedge \langle X, E\rangle \models \ulcorner \mathrm{ZF} \urcorner))^{\langle x,\in\rangle}$. Then the model $\langle X, E\rangle \in x$ *cannot* be a model with E wellfounded (it is an exercise to check this using $\rho(X) < \rho(x)$ - cfEx. 3.15 below).

What we shall attempt with Gödel's construction of L is to show:

$(+)\quad \mathrm{Con(ZF)} \Rightarrow \mathrm{Con}(\mathrm{ZF} + \Phi)$

where Φ will be various statements, such as AC or the GCH.

A statement such as the above (+) should be considered as a statement about the two axioms sets displayed: if the former derives no contradiction neither will the latter. The import is that if we regard ZF as "safe", as a theory, then so will be $\mathrm{ZF} + \Phi$. (One usually claims that these arguments about the relative consistency of recursively given axiom sets are theorems of a particular kind in Number Theory and themselves can be formalised in PA - but we ignore that aspect.)

EXERCISE 3.13 $(*)$ (E) We say that a set x is *outright definable* in a model $\langle M, E\rangle$ of ZFC if there is a formula $\varphi(v_0)$ with the only free variable shown, so that x is the unique set so that $(\varphi[x])^M$ holds. Suppose Con(ZFC). Show that there is a model $\langle M, E\rangle$ of ZFC in which every set is outright definable.

EXERCISE 3.14 $(*)$ (E) Show that there is no formula $\varphi(v_0)$ with just the free variable v_0 so that $\{y \mid \varphi(y)\}$ is the class of outright definable (in $(V, \in)$) sets. [Hint: use a form of Richard's Paradox. Suppose there is such a φ. The least ordinal γ not outright definable is a countable ordinal, but now let $\psi(v_0)$ be "v_0 *is an ordinal*" $\wedge \forall v_1 < v_0 \varphi(v_1)$. Then $\gamma = \{\tau \mid \psi(\tau)\}$.]

EXERCISE 3.15 $(**)$ (E) Suppose that there are transitive models of ZF. Let $\langle x, \in\rangle$ be such, chosen with $\rho(x)$ least. Then (by Ex.3.16) if $\langle X, E\rangle \in x$ is such that $\langle X, E\rangle \models \ulcorner \mathrm{ZF} \urcorner$, show that $\langle X, E\rangle$ cannot be an 'ω-model', that is $\omega^{\langle X,E\rangle} \neq \omega$. (Thus $\langle X, E\rangle$ contains non-standard integers, and in particular codes for non-standard formulae. More particularly still, $(\ulcorner \mathrm{ZF} \urcorner)^{\langle X,E\rangle}$ will contain non-standard axioms besides the standard ones.)

EXERCISE 3.16 $(**)$ (E) Suppose $\langle M, E\rangle$ is a model of ZF. Show that there is an element $\langle N, E'\rangle$ of M with $\langle N, E'\rangle$ a model of ZF.

3.5.2 SATISFACTION OVER V

In the above we defined satisfaction, and so truth, over any set structure $\langle x, \in\rangle$. In particular for x as any V_α. Can we define satisfaction for the whole universe $\langle V, \in\rangle$? The answer is no, not in ZFC alone. That is, there is no single formula $Sat_\omega(v_0, \cdots, v_n)$ so that

$$\langle V, \in\rangle \models Sat_\omega[u, x_1, \cdots, x_n] \Leftrightarrow \mathrm{Fml}(u) \wedge \langle V, \in\rangle \models u[x_1, \cdots, x_n]. \qquad (*)_\omega$$

EXERCISE 3.17 Show that if there were such a formula Sat_ω satisfying the above then ZF would be inconsistent. [Hint: Use the Reflection Theorem on Sat_ω.]

However for any fixed n there is a formula Sat_n that works for the above as long as u is restricted to those formulae that are at level Σ_n in the Levy hierarchy. We let the reader devise a function $\mathrm{Fml_n}$ so that $\mathrm{Fml_n}(u) = 1$ precisely when u codes a formula that is at some level Σ_k for $k \leq n$. Now run Definition 3.14 but restricting it to Σ_k formulae for $k \leq n$ to define Sat_n. This is a legitimate definite recursion defined over $\langle V, \in\rangle$. This gives us a true equivalence $(*)_n$ where Sat_n is in place of Sat_ω in the above.

EXERCISE 3.18 Show that for any natural number n ZF proves that there is an α (indeed a cub class of α) so that $\langle V_\alpha, \in\rangle \prec_{\Sigma_n} \langle V, \in\rangle$. [Hint: for an informal argument, just use reflection on Sat_n. More formally for the second part let $\ulcorner S_n \urcorner$ denote the codes of Σ_n formulae of $\mathcal{L}$ in the Levy hierarchy. Show that there is a term $c_n \subseteq On$ for a closed unbounded class of ordinals, so that $\mathrm{ZF} \vdash \forall \delta \in c_n \forall \ulcorner \varphi \urcorner \in \ulcorner S_n \urcorner \forall \vec{x} \in V_\delta (\varphi(\vec{x}) \leftrightarrow \langle V_\delta, \in\rangle \models \ulcorner \varphi \urcorner [\vec{x}])$. This we should naturally, but informally, also abbreviate as '$\langle V_\delta, \in\rangle \prec_{\Sigma_n} \langle V, \in\rangle$'.]

Just as in the Reflection Theorem, that 'for any natural number n' is on the outside of what ZF proves. That n is metatheoretic and not one of the objects in V.

EXERCISE 3.19 $(**)$ (E) Suppose the language $\mathcal{L}_{\dot{\delta}}$ is the standard language of set theory augmented by a single constant symbol $\dot{\delta}$. Suppose we consider the following scheme of axioms Γ stated in $\mathcal{L}_{\dot{\delta}}$: for each axiom φ of ZFC we adopt the axiom $\varphi_{\dot{\delta}}$: $\forall \vec{x}(\mathrm{Fr}(\varphi) \subseteq \vec{x} \longrightarrow (\varphi[\vec{x}])^{V_\delta} \longleftrightarrow \varphi[\vec{x}]))$. (Thus φ is declared absolute for $V_{\dot{\delta}}$.) Γ consists of all the axioms $\varphi_{\dot{\delta}}$. Informally, taken together then, Γ says that $\langle V_\delta, \in\rangle \prec \langle V, \in\rangle$ where δ interprets $\dot{\delta}$. However the existence of a δ satisfying the latter relation is not provable in ZFC (by the Gödel Incompleteness Theorem). Nevertheless show that $\mathrm{Con(ZFC)} \Rightarrow \mathrm{CON(ZFC} + \Gamma)$. Why does this not contradict Gödel?

Alfred Tarski (1902 Warsaw -1983 Berkeley, USA)

Chapter 4

The Constructible Hierarchy

In this chapter we define the constructible hierarchy due to Gödel, and prove its basic properties. Besides its original purpose used by Gödel to prove the relative consistency of AC and GCH to the other axioms of ZF, we can exploit properties of L to prove other theorems in algebra, analysis, and combinatorics. In set theory itself, properties of L can tell us a lot about V even if $V \neq L$.

Kurt Gödel (1906 Brno - 1983 Princeton, USA)

4.1 The L_α-hierarchy

We use the Def function to define a cumulative hierarchy based on the notion of *definable power set* operation: the Def function.

Definition 4.1 (Gödel) *(i)* $L_0 = \varnothing$; $L_{\alpha+1} = \mathrm{Def}(\langle L_\alpha, \in \rangle$;
$\mathrm{Lim}(\lambda) \longrightarrow L_\lambda = \bigcup\{L_\alpha \mid \alpha < \lambda\}$.
(ii) $L = \bigcup\{L_\alpha \mid \alpha < \mathrm{On}\}$.

Lemma 4.2 *The term L_α is definite, and hence absolute for transitive W satisfying $(\mathrm{ZF}^-)^W$.*

Proof: The Def function is definite and the $\alpha \rightarrowtail L_\alpha$ function is defined by definite recursion from it. Q.E.D.

We thus have defined a class term function $F(\alpha) = L_\alpha$ by a transfinite recursion on On, and so also the term L itself. It is natural to define the notion of "constructible rank" or L-rank, by analogy with ordinary V-rank.

Definition 4.3 *For $x \in L$ we define the L-rank of x, $\rho_L(x) =_{\mathrm{df}}$ the least α so that $x \in L_{\alpha+1}$.*

We give some of the basic properties of the L_α-hierarchy. Many are familiar properties common with the V_α-hierarchy: all of the following are true with L_α replaced by V_α.

Lemma 4.4 *(i) $\beta < \alpha \longrightarrow L_\beta \subseteq L_\alpha$;*
(ii) $\beta < \alpha \longrightarrow L_\beta \in L_\alpha$;
(iii) $\mathrm{Trans}(L_\alpha)$;
(iv) $\alpha = \rho(L_\alpha)$;
(v) $\alpha = \mathrm{On} \cap L_\alpha$.
Hence $\mathrm{Trans}(L)$ and $\mathrm{On} \subseteq L$.

Proof: We prove this by a simultaneous induction for (i)-(v). These are trivial for $\alpha = 0$. Suppose proven for α and we show they hold for $\alpha + 1$.

(i): It suffices to prove that $L_\alpha \subseteq L_{\alpha+1}$ since by the inductive hypothesis, for $\delta < \alpha$ we already know $L_\delta \subseteq L_\alpha$. (Actually this is just an instance of Lemma 3.18(ii), noting that $\mathrm{Trans}(L_\alpha)$ by (iii), but we prove it again.) Let $x \in L_\alpha$. By (iii) for α, $\mathrm{Trans}(L_\alpha)$ and hence $x \subseteq L_\alpha$.

$$x = \{y \in L_\alpha | \langle L_\alpha, \in\rangle \models \ulcorner v_0 \dot{\in} v_1 \urcorner [y, x]\} \in \mathrm{Def}(\langle L_\alpha, \in\rangle) = L_{\alpha+1.}$$

(ii) Again it suffices to show that $L_\alpha \in L_{\alpha+1}$. However $L_\alpha \in \mathrm{Def}(\langle L_\alpha, \in\rangle)$ by Lemma 3.18 (i).
(iii) $L_{\alpha+1} \subseteq \mathcal{P}(L_\alpha)$ hence $x \in L_{\alpha+1} \longrightarrow x \subseteq L_\alpha \subseteq L_{\alpha+1}$ by (i).
(iv) By the inductive hypothesis $\rho(L_\alpha) = \alpha$. By (ii) $L_\alpha \in L_{\alpha+1}$, hence $\alpha = \rho(L_\alpha) < \rho(L_{\alpha+1})$. Hence $\alpha + 1 \leq \rho(L_{\alpha+1})$. For the reverse inequality note that: $x \in L_{\alpha+1} \longrightarrow x \subseteq L_\alpha$, and so $\rho(x) \leq \rho(L_\alpha) = \alpha$. This means that

$\rho(L_{\alpha+1}) =_{\mathrm{df}} \sup\{\rho(x) + 1 \mid x \in L_{\alpha+1}\} \leq \alpha + 1.$

(v) By the inductive hypothesis and (i) $\alpha \subseteq L_\alpha \subseteq L_{\alpha+1}$, so it suffices to show that $\alpha \in L_{\alpha+1}$ in order to show that $\alpha + 1 \subseteq L_{\alpha+1}$. Thus:

$$\alpha = \{\delta \in L_\alpha \mid \delta \in \mathrm{On}\} = \{\delta \in L_\alpha \mid \langle L_\alpha, \in\rangle \models \ulcorner v_0 \dot{\in} \mathrm{On} \urcorner [\delta]\} \in \mathrm{Def}(\langle L_\alpha, \in\rangle) = L_{\alpha+1}.$$

That $\mathrm{On} \cap L_{\alpha+1} \subseteq \alpha + 1$: $\mathrm{On} \cap L_{\alpha+1} \subseteq \{\delta \in \mathrm{On} \mid \rho(\delta) < \alpha + 1\}$ by (iv). But the latter is just $\alpha + 1$.

We now assume $\mathrm{Lim}(\lambda)$ and (i)-(v) hold for $\alpha < \lambda$. Then (i)-(iii) and (v) are immediate. For (iv) : $\rho(L_\lambda) = \sup\{\rho(x) + 1 \mid x \in L_\lambda\} \leq \sup\{\alpha | \alpha \in \lambda\} = \lambda$. Conversely $\lambda \subseteq L_\lambda \longrightarrow \rho(L_\lambda) \geq \lambda$. Q.E.D.

LEMMA 4.5 *(i) For all $\alpha \in \mathrm{On}$, $\rho_L(\alpha) = \rho(\alpha) = \alpha$.*
(ii) For $n \le \omega$ $L_n = V_n$.
(iii) For all $\alpha \ge \omega$ $|L_\alpha| = |\alpha|$.

PROOF: (i) and (ii): Exercise. For (iii) we prove this by induction on α. For $\alpha = \omega$ this follows from (ii) and $|V_\omega| = \omega$. Suppose proven for α. $|L_{\alpha+1}| = |\mathrm{Def}(\langle L_\alpha, \in\rangle)| = |L_\alpha| = |\alpha| = |\alpha + 1|$ by Lemma 3.18 (iv). For $\mathrm{Lim}(\lambda)$: $|L_\lambda| = |\bigcup_{\alpha<\lambda} L_\alpha| \le |\lambda| \cdot |\lambda| = |\lambda|$ as by the inductive hypothesis $|L_\alpha| = |\alpha| \le |\lambda|$ for $\alpha < \lambda$.
Q.E.D.

EXERCISE 4.1 (i) Verify that for all $\alpha \in \mathrm{On}$, $\rho_L(\alpha) = \rho(\alpha) = \alpha$ (ii) Prove that for $n \le \omega$ $L_n = V_n$.

Remark: (i) shows that as far as ordinals go, they appear at the same stage in the L-hierarchy as in the V-hierarchy. However it is important to note that this is not the case for all constructible sets: there are constructible subsets of ω that are not in $L_{\omega+1}$.

DEFINITION 4.6 *(i) Let T be a set of axioms in $\mathcal{L}$. Let W be a class term. Then W is an* inner model *of T, if (a)* $\mathrm{Trans}(W)$*; (b)* $\mathrm{On} \subseteq W$ *; (c)* $(T)^W$ *, that is, for each σ in T, $(\sigma)^W$.*
(ii) If (i) holds we write $\mathrm{IM}(W, T)$ *and if T is* ZF *then simply* $\mathrm{IM}(W)$.

THEOREM 4.7 **(Gödel)** *L is an inner model of* ZF, $\mathrm{IM}(L)$. *In particular* $(\mathrm{ZF})^L$.

Remark: again this is to be read as saying: for each axiom φ of ZF, $\mathrm{ZF} \vdash (\varphi)^L$.
PROOF: We already have (a) and (b) by Lemma 4.4, so it remains to show $(\mathrm{ZF})^L$. We justify this by considering each axiom (or axiom schema) in turn. We use all the time, without comment the fact that each L_α is transitive.

Ax 0 ***Empty*** is trivial as $\varnothing = \varnothing^L \in L$.

Ax1: ***Extensionality***: This is Lemma 1.21, since we have Trans(L).

Ax2: ***Pairing Axiom*** Let $x, y \in L_\alpha$. Then
$\{x, y\} = \{z \in L_\alpha \mid \langle L_\alpha, \in\rangle \models \ulcorner v_0 \dot{=} v_1 \vee v_0 \dot{=} v_2 \urcorner [z/0, x/1, y/2]\} \in \mathrm{Def}(L_\alpha) = L_{\alpha+1} \subseteq L$.
By Lemma 1.24 then Ax2 holds in L.

Ax3 ***Union Axion*** Let $x \in L_\alpha$. This follows from Lemma 1.25 once we show:
$\bigcup x = \{z \in L_\alpha \mid \langle L_\alpha, \in\rangle \models \ulcorner \exists v_1 (v_1 \dot{\in} v_2 \wedge v_0 \dot{\in} v_1 \urcorner [z/0, x/2]\} \in \mathrm{Def}(L_\alpha)$.

Ax4 ***Foundation Scheme*** Let a be a term. Then:
$(a \ne \varnothing \longrightarrow (\exists x \in a(x \cap a = \varnothing)))^L \leftrightarrow (a^L \ne \varnothing \longrightarrow \exists x \in a^L(x \cap a^L = \varnothing))$. But the right hand side of the equivalence here is simply an instance of the Foundation scheme in V and thus is true.

Ax5 ***Separation Scheme*** Again let a be a class term. Suppose

$$a = \{z | \varphi(z/0, y_1/1, \ldots, y_n/n)\}.$$

Suppose $x, \vec{y} \in L_\gamma$. We apply Lemma 2.42 to the hierarchy $Z_\alpha = L_\alpha$, $Z = L$ to obtain a $\beta > \gamma$ so that
$\mathrm{ZF} \vdash \forall z \in L_\beta(\ (\varphi(z, y_1, \ldots, y_n))^L \leftrightarrow (\varphi(z, y_1, \ldots, y_n))^{L_\beta} \ )$.

By the Correctness Theorem 3.20

$$(\varphi(z, y_1, \ldots, y_n))^{L_\beta} \leftrightarrow \langle L_\beta, \in\rangle \models \ulcorner \varphi \urcorner [z, y_1, \ldots, y_n].$$

Hence, putting it all together:

$$\{z \in x \mid \varphi(z, y_1, \dots, y_n)\}^L = \{z \in x \mid \varphi(z, y_1, \dots, y_n)^{L_\beta}\} =$$
$$= \{z \in L_\beta \mid \langle L_\beta, \in\rangle \models \ulcorner \varphi \wedge v_0 \dot{\in} v_{n+1} \urcorner [z, y_1, \dots, y_n, x]\} \in \mathrm{Def}(L_\beta).$$

Ax6 ***Replacement Scheme*** Suppose f is a term, $x \in L$, and $\mathrm{Fun}(f)^L$. Let ρ_L be the constructible rank function. Then by the Replacement Scheme (in V) $(\rho_L \circ f^L)$"$x \in V$. Let α be its supremum. Then f^L"$x \subseteq L_\alpha$. Let $\beta \geq \alpha$ be sufficiently large so that by the Reflection Theorem

$$\mathrm{ZF} \vdash \forall y, z \in L_\beta(\ (f(z) = y)^L \leftrightarrow (f(z) = y)^{L_\beta}\ ).$$

Then again using the Correctness theorem we have that

$$f^L\text{“}x = \{y \in L_\beta | \langle L_\beta, \in\rangle \models \ulcorner \exists v_1 \dot{\in} v_2 (f(v_1) = v_0)\urcorner [y/0, x/2]\} \in \mathrm{Def}(L_\beta).$$

Ax7 ***Infinity Axiom*** Just note that $\omega \in L_{\omega+1}$.

Since we have shown the requisite sets are all in L we apply the appropriate cases of Lemma 1.25 and conclude Ax3,5,6,7 hold in L. We are thus left with:

Ax8 ***PowerSet Axiom*** $(\forall x \exists y (y = \mathcal{P}(x)))^L \leftrightarrow (\forall x \exists y \forall z (z \subseteq x \leftrightarrow z \in y))^L \leftrightarrow$

$$\leftrightarrow \forall x \in L \exists y \in L \forall z \in L (z \subseteq x \leftrightarrow z \in y)$$
$$\leftrightarrow \forall x \in L \exists y \in L (y = \mathcal{P}(x) \cap L).$$

So we verify the latter: let $x \in L$ be arbitrary. $\mathcal{P}(x) \cap L \in V$ by Axiom of Power and Separation in V. By Ax.Replacement ρ_L"$\mathcal{P}(x) \cap L \in V$. Let α be its supremum. Then, as required:

$$\mathcal{P}(x) \cap L = \{z \in L_\alpha | \langle L_\alpha, \in\rangle \models \ulcorner v_0 \dot{\subseteq} v_1 \urcorner [z, x]\} \in \mathrm{Def}(L_\alpha).$$ Q.E.D.

Suppose we define $\mathrm{im}_0(W)$ to be the variant on $\mathrm{IM}(W)$ that, keeping (a) and (b), replaces (c) by the statement that "$\forall x \subseteq W \exists y \in V (x \subseteq y \wedge \mathrm{Trans}(y) \wedge \mathrm{Def}(\langle y, \in\rangle \subseteq W)$" then a close reading of the last proof reveals that we in fact may show:

Theorem 4.8 *Suppose W is a class term and* $\mathrm{im}_0(W)$. *then* $\mathrm{IM}(W)$.

Exercise 4.2 $(*)$ (E) Prove this last theorem.

Exercise 4.3 Show that "*x is a cardinal*" and "*x is regular*" are downward absolute from V to L. Deduce that if κ is a (regular) limit cardinal then (κ *is a (regular) limit cardinal*)^L.

4.2 The Axiom of Choice in L

The very regular construction of the L_α-hierarchy ensures that the Axiom of Choice will hold in the constructible universe L. Indeed, it holds in a very strong form: whereas the Axiom of Choice is equivalent to the statement that any set can be wellordered, for L there is a class term that wellorders the whole universe of L in one stroke. Essentially what is at the heart of the matter is that we may wellorder the countably many formulae of the language $\mathcal{L}$, and then inductively define a wellorder $<_{\alpha+1}$ for $L_{\alpha+1}$ using a wellorder $<_\alpha$ for L_α. This latter wellorder $<_\alpha$ gives us a way of ordering all finite k-tuples of elements of L_α, and thus, putting these together, we get a wellorder of all possible definitions that go into making up new objects in $L_{\alpha+1}$. We shall additionally have that the ordering $<_{\alpha+1}$ *end-extends* that of $<_\alpha$. This means that if $y \in L_{\alpha+1} \setminus L_\alpha$ then for no $x \in L_\alpha$ do we have that $y <_{\alpha+1} x$. Taking $<_L = \bigcup_{\alpha \in \mathrm{On}} <_\alpha$ gives us the term for a global wellordering of all of L. We now proceed to fill out this sketch.

Let $x \in V$ and suppose we are given a wellorder $<_x$ of x. We define from this a wellorder $<_{Q_x}$. For $f \in Q_x$ we let $\mathrm{lh}(f) =_{\mathrm{df}}$ the least n so that $\forall m \geq n(f(m) = f(n))$. We then define for $f, g \in Q_x$:

$$f <_{Q_x} g \longleftrightarrow_{\mathrm{df}} \mathrm{lh}(f) < \mathrm{lh}(g) \vee (\mathrm{lh}(f) = \mathrm{lh}(g) \wedge \exists k \leq \mathrm{lh}(f)(\forall n < k f(n) = g(n) \wedge f(k) <_x g(k))).$$

Exercise 4.4 Check that $<_{Q_x}$ is definite. Moreover if $<_x \in \mathrm{WO}$ then $<_{Q_x} \in \mathrm{WO}$.

We now suppose we also have fixed an ordering $\varphi_0, \varphi_1, \ldots, \varphi_n, \ldots$ of the countably many elements of Fml which have at least v_0 amongst their free variables (we may define such a listing from any map $g : \omega \longleftrightarrow \mathrm{Fml}$). We assume that the function f given by $f(n) = \ulcorner \varphi_n \urcorner$ is a definite term.

Definition 4.9 *We define by recursion the ordering* $<_\alpha$ *of* L_α. $<_0 = \varnothing$; *let* $x, y \in L_{\alpha+1}$:

$$x <_{\alpha+1} y \leftrightarrow_{\mathrm{df}}$$
$$(x \in L_\alpha \wedge y \notin L_\alpha) \vee (x, y \in L_\alpha \wedge x <_\alpha y) \vee (x, y \notin L_\alpha \wedge \exists n \in \omega \exists f \in Q_{L_\alpha}(x = \iota(L_\alpha, \varphi_n, f) \wedge$$
$$\forall m \in \omega \forall g \in Q_{L_\alpha}(y = \iota(L_\alpha, \varphi_m, g) \longrightarrow n < m \vee (m = n \wedge f <_{Q_{L_\alpha}} g)))).$$

$\mathrm{Lim}(\lambda) \longrightarrow <_\lambda = \bigcup_{\alpha<\lambda} <_\alpha$; $<_L =_{\mathrm{df}} \bigcup_{\alpha \in \mathrm{On}} <_\alpha$.

Lemma 4.10 *(i)* $<_\alpha$ *is definite; (ii) the ordering* $<_\beta$ *is a wellordering and end-extends* $<_\alpha$ *if* $\alpha \leq \beta$ *; (iii) if* κ *is an infinite cardinal then* $<_\kappa$ *has order type* κ*;* $<_L$ *has order type On. Thus* $(\mathrm{AC})^L$.

Proof: (i) $f(\alpha) =_{\mathrm{df}} <_\alpha$ is defined by a definite recursion. (ii) By an obvious induction on α. (iii) Exercise. Q.E.D.

Exercise 4.5 Show that $\mathrm{ot}(L_\kappa, <_\kappa) = \kappa$ for κ an infinite cardinal; deduce that $\mathrm{ot}(L, <_L) = \mathrm{On}$.

4.3 The Axiom of Constructibility

Definition 4.11 *The Axiom of Constructibility is the assertion "*$V = L$*" which abbreviates "*$\forall x \exists \alpha\, x \in L_\alpha$*."*

The Axiom of Constructibility thus says that every set appears somewhere in this hierarchy. Since the model L is defined by a restricted use of the power set operation, many set theorists feel that the Def function is too restricted a method of building *all* sets. Nevertheless, the *inner model* L of the constructible sets, possesses a very rich structure.

Lemma 4.12 *(i) Let* W *be a transitive class term, and suppose* $(\mathrm{ZF}^-)^W$*. Then*

$$\begin{aligned} (L)^W &= L \quad \text{if } \mathrm{On} \cap W = \mathrm{On} \\ &= L_\theta \quad \text{if } \mathrm{On} \cap W = \theta. \end{aligned}$$

(ii) There is a finite conjunction σ_1 *of* ZF^- *axioms, so that in (i) the requirement that* $(\mathrm{ZF}^-)^W$ *can be replaced by* $(\sigma_1)^W$ *and the conclusion is unaltered.*

Proof: (i) The function term L_α is definite. Hence is absolute for such a W. Note that in the case that $\mathrm{On} \cap W = \theta \in \mathrm{On}$ then indeed $\mathrm{Lim}(\theta)$. But in either case for any $\alpha \in W$ $(L_\alpha)^W = L_\alpha$. Hence

$$(L)^W = (\bigcup\{L_\alpha | \alpha \in \mathrm{On}\})^W = \bigcup\{L_\alpha | \alpha \in \mathrm{On} \cap W\}$$

which yields the above result.

(ii) σ_1 is simply the conjunction of sufficiently many axioms needed for the proof that the function term L_α is definite, plus "*there is no largest ordinal*". Q.E.D.

Corollary 4.13 (ZF) $(V = L)^L$.

Proof: $\mathrm{Trans}(L)$ and $(\mathrm{ZF}^-)^L$. But $(V = L)^L \leftrightarrow V^L = L^L$. As $V^L = L$ and by Lemma 4.12 $(L)^L = L$ we are done. Q.E.D.

Theorem 4.14 $\mathrm{Con}(\mathrm{ZF}) \Rightarrow \mathrm{Con}(\mathrm{ZF} + V = L)$.

Proof: Suppose $\mathrm{ZF} + V = L$ is inconsistent. Suppose $\mathrm{ZF} + V = L \vdash (\varphi \wedge \neg\varphi)$.

$\mathrm{ZF} \vdash (\mathrm{ZF} + V = L)^L$ by the last Corollary and Theorem 4.7, then:

$\mathrm{ZF} \vdash (\varphi \wedge \neg\varphi)^L$, and hence:

$\mathrm{ZF} \vdash \varphi^L \wedge (\neg\varphi)^L$. Hence:

$\mathrm{ZF} \vdash \varphi^L \wedge \neg(\varphi^L)$. Hence ZF is inconsistent. Q.E.D.

Remark 4.15 P. Cohen (1962) showed $\mathrm{Con}(\mathrm{ZF}) \Rightarrow \mathrm{Con}(\mathrm{ZF} + V \neq L)$ by an entirely different method, that of "*forcing*". This method can be construed as either constructing models in a Boolean valued (rather than a 2-valued) logic; or else akin to some kind of syntactic method of construction. (An entirely different method was needed - see Exercise 4.7.) He further showed that $\mathrm{Con}(\mathrm{ZF}) \Rightarrow \mathrm{Con}(\mathrm{ZF} + \neg\,\mathrm{AC})$ and $\mathrm{Con}(\mathrm{ZF}) \Rightarrow \mathrm{Con}(\mathrm{ZF} + \neg\,\mathrm{CH})$. His methods are now much elaborated to prove a wealth of "*relative consistency*" statements such as these.

Theorem 4.16 **(Gödel 1939)** $\mathrm{Con}(\mathrm{ZF}) \Rightarrow \mathrm{Con}(\mathrm{ZF} + \mathrm{AC})$

Proof: We have shown $\mathrm{ZF} \vdash (\mathrm{AC})^L$, but also $\mathrm{ZF} \vdash (\mathrm{ZF})^L$, and thus $\mathrm{ZF} \vdash (\mathrm{ZF} + \mathrm{AC})^L$, Hence if $\mathrm{ZF} + \mathrm{AC} \vdash \varphi \wedge \neg\varphi$ for some φ then we should have $\mathrm{ZF} \vdash (\varphi \wedge \neg\varphi)^L$ as in the last proof, and hence not $\mathrm{Con}(\mathrm{ZF})$. Q.E.D.

Exercise 4.6 Suppose there is a transitive set model of ZFC. Show that there is a *minimal (transitive) model* of ZFC, that is for some countable ordinal β_0, $L_{\beta_0} \models \ulcorner\mathrm{ZFC}\urcorner$ and that L_{β_0} is a subclass of any other such transitive set model of ZF .

Remark We note the following observation on argumentation. For any formula χ we have the equivalence: $\mathrm{ZF} \vdash \chi^L$ if and only if $\mathrm{ZF} + V = L \vdash \chi$. For the $(\Rightarrow)$ direction, we have that, clearly, $\mathrm{ZF} + V = L \vdash \chi^L$. But it also proves that $\chi^L \leftrightarrow \chi^V \leftrightarrow \chi$. For the converse, we have that $\mathrm{ZF} \vdash (\mathrm{ZF})^L$, but also we saw at Cor.4.13 that $\mathrm{ZF} \vdash (V = L)^L$. By the Soundness Theorem we thus have $\mathrm{ZF} \vdash \chi^L$. This observation can make proofs of properties of L, which can then be proven under the additional hypothesis of $V = L$, easier.

4.4 The Generalised Continuum Hypothesis in L.

We first prove a simple lemma, but one of great utility.

Lemma 4.17 **(The Condensation Lemma)** *Let σ_1 be the finite conjunction of axioms of* ZF$^-$ *from Lemma 4.12, and suppose x and α are such that $\langle x, \in\rangle \prec \langle L_\alpha, \in\rangle$ and $(\sigma_1)^{L_\alpha}$. Then there is $\gamma \leq \alpha$ with $\langle x, \in\rangle \cong \langle L_\gamma, \in\rangle$.*

Proof: As $(\sigma_1)^{L_\alpha}$, and and so by that Lemma 4.12, we have $(V = L)^{L_\alpha}$. So by the Correctness Theorem, we have that $\langle L_\alpha, \in\rangle \models \ulcorner V = L\urcorner$. Hence $\langle x, \in\rangle \models \ulcorner V = L\urcorner$. Let $\pi : \langle x, \in\rangle \longrightarrow \langle y, \in\rangle$ be the Mostowski Shepherdson Collapse with $\mathrm{Trans}(y)$. Then $\langle y, \in\rangle \models \ulcorner\sigma_1\urcorner \wedge \ulcorner V = L\urcorner$ (as π is an isomorphism). By the first conjunct, and Correctness again, we have $(\sigma_1 \wedge V = L)^y$. By Lemma 4.12, $L^y = L \cap y = L_{\mathrm{On}\cap y}$. But by the second conjunct then, this equals y itself. So we may take $\gamma = \mathrm{On} \cap y$. Q.E.D.

Note: It can be shown that the assumption that $(\sigma_1)^{L_\alpha}$ can be very much reduced: all that is needed for the conclusion of the lemma is that $\mathrm{Lim}(\alpha)$, and with a lot more fiddling around even this condition can be dropped, and we have Condensation holding for every L_α.

Theorem 4.18 $\mathrm{ZF} \vdash (\omega \leq \kappa \in \mathrm{Card} \longrightarrow H_\kappa = L_\kappa)^L$. *Hence* $\mathrm{ZF} \vdash (\mathrm{GCH})^L$ *and thus* $\mathrm{ZF} + V = L \vdash \mathrm{GCH}$.

Proof: By the Remark at the end of the last subsection it suffices to prove $\mathrm{ZF} + V = L \vdash \mathrm{GCH}$. So we shall assume $V = L$. We have that $L_\omega = V_\omega = H_\omega$ already and hence the conclusion for $\kappa = \omega$. Assume $\omega < \kappa \in \mathrm{Card}$. If $\alpha < \kappa$ then by Lemma 4.5(iii) $|L_\alpha| = |\alpha| < \kappa$. Hence $L_\alpha \in H_\kappa$. Thus $L_\kappa \subseteq H_\kappa$. Now for the reverse inclusion suppose $z \in H_\kappa$. Find an α sufficiently large with $\{z\}, \mathrm{TC}(z) \in L_\alpha$ and by the Reflection Theorem $(\sigma_1)^{L_\alpha}$. As $z \in H_\kappa \longrightarrow \mathrm{TC}(z) \in H_\kappa$, we may apply the Downward Löwenheim-Skolem theorem in L and find $\langle x, \in\rangle \prec \langle L_\alpha, \in\rangle$ with $\mathrm{TC}(\{z\}) = \mathrm{TC}(z) \cup \{z\} \subseteq x$, and $(|x| = \max\{|\mathrm{TC}(\{z\})|, \aleph_0\} < \kappa)$.

As the transitive part of x contains all of $\mathrm{TC}(\{z\})$, we have that $\pi(z) = z$ where π is the transitive collapse map mentioned in the Condensation Lemma, taking $\pi : \langle x, \in\rangle \longrightarrow \langle y, \in\rangle = \langle L_\gamma, \in\rangle$ for some $\gamma \leq \alpha$. However we know that $|x| = |L_\gamma| = |\gamma| < \kappa$ by design. Hence $z \in L_\gamma \in L_\kappa$.

As $z \in H_\kappa$ was arbitrary we conclude that $L_\kappa \supseteq H_\kappa$. We thus have shown $H_\kappa = L_\kappa$. To show GCH it suffices to show that for all infinite cardinals κ that $2^\kappa = \kappa^+$. However $2^\kappa \approx \mathcal{P}(\kappa)$ and $\mathcal{P}(\kappa) \subseteq H_{\kappa^+} = L_{\kappa^+}$. Hence $|\mathcal{P}(\kappa)| \leq |L_{\kappa^+}| = \kappa^+$. By Cantor's Theorem we conclude $|\mathcal{P}(\kappa)| = \kappa^+$.

This argument establishes that $\mathrm{ZF} \vdash \mathrm{GCH}$. If we additionally assume $V = L$ we have the conclusion of the Theorem. Q.E.D.

The proof of the next is identical to that of Cor. 4.16:

Corollary 4.19 **(Gödel 1939)** $\mathrm{Con}(\mathrm{ZF}) \Rightarrow \mathrm{Con}(\mathrm{ZF} + \mathrm{GCH})$.

Exercise 4.7 (E) (Shepherdson) Show that there is no class term W so that $\mathrm{ZFC} \vdash \mathrm{IM}(W)$ and $\mathrm{ZFC} \vdash (\neg\,\mathrm{CH})^W$. [This Exercise shows that Gödel's argument was essentially a "one-off": there is no way one can define in ZFC alone an inner model and hope that it is a model of all of ZF plus, *e.g.*, $\neg$ CH.]

Exercise 4.8 Show that if there is a weakly inaccessible cardinal κ then $(\mathrm{ZFC})^{L_\kappa}$. Hence ZFC $\nvdash \exists\kappa(\kappa$ a weakly inaccessible cardinal.) [Hint: Use the fact that $(\mathrm{GCH})^L$.]

Exercise 4.9 Show that if κ is weakly inaccessible then $\forall\alpha < \kappa\exists\beta < \kappa(\beta > \alpha \wedge L_\beta \models \ulcorner\mathrm{ZFC}\urcorner)$. [Hint: use the Condensation Lemma and Downward Löwenheim-Skolem Theorem.]

Exercise 4.10 Assume $V = L$. When does $L_\alpha = V_\alpha$?

Exercise 4.11 (E)($*$) Show that if $\alpha < \omega_1$ is any limit ordinal, which is countable in L, then there is β, countable in L, so that $\mathcal{P}(\omega) \cap L_\beta = \mathcal{P}(\omega) \cap L_{\beta+\alpha}$. [This shows that although by the GCH proof all constructible reals will have appeared by stage $(\omega_1)^L$, there are arbitrarily long 'gaps' of countable length in the constructible hierarchy below $(\omega_1)^L$, where no new real numbers appear,. Hint: Suppose $V = L$. Let $\mathfrak{A} = \langle L_{\omega_2}, \in\rangle$ and, by the Downward Löwenheim-Skolem Theorem, let $Y \supseteq \alpha+1$ be a countable elementary substructure of $\mathfrak{A}$: $Y \prec \mathfrak{A}$. Let $\pi : Y \longrightarrow M$ be the transitive collapse of Y and as in the GCH proof, $M = L_\gamma$ for some γ. Consider $\beta = \pi(\omega_1)$. Then $\gamma > \beta > \alpha$.]

Exercise 4.12 Show that (i) if κ is a weakly inaccessible cardinal, then $(\kappa$ *is strongly inaccessible*$)^L$; (ii) if κ is a weakly Mahlo cardinal, then $(\kappa$ *is strongly Mahlo*$)^L$.[Hint: See Exercises 4.3 & 4.8. For (ii) show that the property of being cub in κ is preserved upwards from L to V.]

Exercise 4.13 (i) Let $\langle x, \in\rangle \prec L_{\omega_1}$ where $\omega_1 = (\omega_1)^L$. Show that already $\mathrm{Trans}(x)$ and so $x = L_\gamma$ for some $\gamma \leq \omega_1$. [Hint: For $\delta < \omega_1$ note that $(|\delta| = |L_\delta| = \omega)^{L_{\omega_1}}$. Hence for $\delta \in x$, in L_{ω_1}, and thus in x, there is an onto map $f : \omega \longrightarrow L_\delta$. Thus, as $\omega \subseteq x \wedge f \in x$ we deduce that $\mathrm{ran}(f) = L_\delta \subseteq x$. Deduce that $\mathrm{Trans}(x)$.]

(ii) ($*$) Now let $\langle x, \in\rangle \prec L_{\omega_2}$ where $\omega_2 = (\omega_2)^L$. Show that $\mathrm{Trans}(x \cap L_{\omega_1})$ and so $x \cap L_{\omega_1} = L_\gamma$ for some γ.

Exercise 4.14 ($*$) Assume $V = L$. N. Schweber defined a countable ordinal τ to be *memorable* if for all sufficiently large $\beta < \omega_1$, $\tau \in \mathrm{Def}_0(\langle L_\beta, \in\rangle)$. Show:
(i) The memorable ordinals form a countable, so proper, initial segment of $(\omega_1, \in)$.
(ii) Let δ be the least non-memorable ordinal. Show that δ is also the least ordinal η so that for arbitrarily large $\gamma < \omega_1$, $L_\eta \prec L_\gamma$.

4.5 Ordinal Definable sets and *HOD*

Gödel's method of defining the inner model L of constructible sets was not the only way to obtain the consistency of the Axiom of Choice with the other axioms of ZF. Another model can be defined, the inner model of the *hereditarily ordinal definable sets* or "*HOD*" in which the *AC* can be shown to hold. (The GCH is not provably true there, and the absoluteness of the construction of L - which allowed us to show that $L^L = L$ is not available: it is consistent that $HOD^{HOD} \neq HOD$.) We investigate the basics here. There is some evidence that Gödel was aware of this approach, as he suggested looking at the ordinal definable sets for a model of *AC*. However the construction requires essential use of the Reflection Theorem that was not proven until the end of the 1950's by Levy and Montague. Some see these remarks of Gödel as indicating that he was aware of the Reflection Principle, even if he did not publish a proof.

Definition 4.20 *We say that a set z is* ordinal definable *($z \in OD$) if and only if for some formula $\varphi(v_0, v_1, \ldots, v_m)$ with free variables shown, for some ordinals $\alpha_1, \ldots, \alpha_m$ then z is the unique set so that $\varphi[z, \alpha_1, \ldots, \alpha_m]$.*

We next need to show that the expression '$z \in OD$' is definable within ZF. (At the moment the last clause of the last definition has loosely talked about "definability (in $\langle V, \in \rangle$)" - which is not definable in $\langle V, \in \rangle$.) We do this by showing it is equivalent to the alternative definition given in Def.3.19, which involved only the definable sets V_β and the definable function $\mathrm{Def}_0(x)$.

EXERCISE 4.15 (Richard's Paradox) Let T be the set of those z so that for some closed term $\{x \mid t(x)\}$ (that is one without free variables) $z = \{x \mid t(x)\}$. Show that there is no formula $\psi(v_0)$ (with just the one free variable shown), so that $T = \{z \mid \psi[z]\}$, and thus T is not definable by such a formula, thus T is not such a closed term. [Hint: as there are only countably many closed terms, there will only be countably many ordinals in T. Suppose for a contradiction that $\psi(v_0)$ does define the set of elements of T (meaning that it is true of just the elements of T). Consider the term $\{\alpha \mid \forall \beta \leq \alpha(\psi[\beta])\}$.]

EXERCISE 4.16 Let $\vec{\gamma} = \gamma_0, \ldots, \gamma_{n-1} \in {}^n On$ for some n. Then there is β so that $\vec{\gamma} \in \mathrm{Def}_0(V_\beta)$. [Hint: Let $<^n$ be the wellorder of ${}^n On$ as above at Ex. 1.10. Let $\varphi_0(\vec{\alpha})$ express "$\vec{\alpha}$ is the $<^n$-least sequence so that $\forall \beta\, (\vec{\alpha} \notin \mathrm{Def}_0(V_\beta))$". But if $\varphi_0(\vec{\alpha})$ were true, it would reflect to some δ. But then $\vec{\alpha} \in \mathrm{Def}_0(V_\delta)$.]

EXERCISE 4.17 (Scott) For any formula $\psi(v_0, \ldots, v_{m-1})$ with free variables $v_0, \ldots, v_{m-1}$,

$$\mathrm{ZF} \vdash \forall \alpha_0, \ldots, \alpha_{n-1} \exists \beta(\alpha_0, \ldots, \alpha_{n-1} \in \mathrm{Def}_0(V_\beta) \wedge \forall x_0, \ldots, x_{m-1} \psi(x_0, \ldots, x_{m-1}) \leftrightarrow (\psi(x_0, \ldots, x_{m-1}))^{V_\beta}).$$

[Hint: Another use of the Richard Paradox argument. Expand Ex.4.16. Suppose the displayed formula is $\forall \alpha_0, \ldots, \alpha_{n-1} \varphi_1(\vec{\alpha})$, and suppose $\varphi_1(\vec{\alpha})$ false for some $<^n$-least $\alpha_0, \ldots, \alpha_{n-1}$. Let β be any sufficiently large ordinal (so greater than $\max\{\vec{\alpha}\}$) that reflects $\varphi_1 \wedge \psi$. But now, as in Ex.4.16, $\alpha_0, \ldots, \alpha_{n-1} \in \mathrm{Def}_0(V_\beta)$ and V_β reflects ψ too which is a contradiction.]

THEOREM 4.21 *'$z \in OD$' is expressible by the single formula in* ZF: $\varphi_{OD}(z)$: "$\exists \beta(z \in \mathrm{Def}_0(V_\beta))$".

PROOF: Let OD^* denote the class of sets z satisfying the Definition 3.19, that is the formula $\varphi_{OD}(z)$ above. It suffices to show then $OD^* = OD$. ($\subseteq$) is clear. Suppose x is the unique set satisfying $\varphi[x, \alpha_0, \ldots, \alpha_{n-1}]$. By Ex. 4.17 there is β with $\alpha_0, \ldots, \alpha_{n-1} \in \mathrm{Def}_0(V_\beta)$ and V_β reflects φ with $x \in V_\beta$. Then $\varphi[x, \alpha_0, \ldots, \alpha_{n-1}]$ defines x in V_β. But amalgamating the definitions of the sequence $\vec{\alpha}$ with that given by φ we have a definition $\varphi'[x]$ in V_β without the use of ordinal parameters. Thus $x \in \mathrm{Def}_0(V_\beta)$. Q.E.D.

THEOREM 4.22 *OD has a definable wellordering.*

PROOF: We use a definable wellorder $<^{HF}$ of HF to impose a wellordering on the Gödel code sets of formulae with one free variable. As $OD = \{z \mid \exists \beta(z \in \mathrm{Def}_0(V_\beta))\}$ for any $z \in OD$ we can set $\beta(z) =_{df}$ the least β so that $z \in \mathrm{Def}_0(V_\beta)$. Let ϕ_z be the least, in the ordering $<^{HF}$, formula with the single free variable v_0, that defines z in $V_{\beta(z)}$.

Now define

$$x <_{OD} z \Leftrightarrow x, z \in OD \wedge (\beta(x) < \beta(z) \vee (\beta(x) = \beta(z) \wedge \phi_x <^{HF} \phi_z)).$$

One can check this is a wellorder of OD. Q.E.D.

LEMMA 4.23 *Let A be any class that has a definable set-like wellorder given by some $\varphi(v_0, v_1)$ ("set-like" meaning for any $z_0 \in A$, $\{z \in A \mid \varphi(z, z_0)\}$ is a set). Then $A \subseteq OD$.*

PROOF: By assumption we can define by recursion a rank function $r(z) = \sup\{r(y)+1 \mid y \in A \wedge \varphi(y,z)\}$. Then $\mathrm{ran}(r) \subseteq On$. But now for each $z \in A$ for some α we have $r(z) = \alpha$ and we may define z as "that unique z with $r(z) = \alpha$". Q.E.D.

COROLLARY 4.24 $L \subseteq OD$

PROOF: By Ex. 4.5 the ordering $<_L$ of L is both definable and a wellorder in order type On. It is thus "set-like" as described above. Hence $L \subseteq OD$. Q.E.D.

COROLLARY 4.25 $Con(\mathrm{ZF}) \Rightarrow Con(\mathrm{ZF} + V = OD)$.

PROOF: $V = L$ implies $V = OD$ by the last corollary. So this follows from Theorem 4.14. Q.E.D.

On the other hand it is not provable in ZF that $V = OD$ (or that OD is transitive; even $(AxExt)^{OD}$ may fail, see Ex.4.20 below). Thus we cannot prove that OD is an inner model of ZFC. For this we need to consider the closely related subclass of *hereditarily ordinal definable sets*.

DEFINITION 4.26 *(The hereditarily ordinal definable sets - HOD)*

$$z \in HOD \Leftrightarrow z \in OD \wedge TC(z) \subseteq OD.$$

We thus require not only that z be in OD but this fact propagates down through the $\in$-relation below z. By definition HOD is a transitive class of sets, and it contains all ordinals.

EXERCISE 4.18 Show $z \in HOD \leftrightarrow z \in OD \wedge \forall y \in z(y \in HOD)$. Show that $\mathcal{P}(\omega) \cap OD = \mathcal{P}(\omega) \cap HOD$.

THEOREM 4.27 $(\mathrm{ZFC})^{HOD}$, *that is for each axiom τ of* ZFC, *we have* τ^{HOD}.

PROOF: By transitivity of HOD we have $\varnothing \in HOD$ and AxExtensionality holds in HOD. It is easy to check that $x, y \in HOD \to \{x,y\}, \bigcup x \in HOD$. Likewise as any ordinal is in HOD (*e.g.*, by induction using the last exercise), so is $\omega \in HOD$. For AxPower: suppose $x \in HOD$. It will suffice to show that $\mathcal{P}(x)^{HOD} = \mathcal{P}(x) \cap HOD \in OD$. The first equality is obvious as "$y \subseteq x$" is Δ_0. But notice that $\mathcal{P}(x) \cap HOD = \mathcal{P}(x) \cap OD$. (If $y \subseteq x \wedge y \in OD$ then $y \in HOD$ by the exercise.) So it suffices to show $\mathcal{P}(x) \cap OD \in OD$. Let $\gamma_0 = \rho(x)$; then $\mathcal{P}(x) \cap OD \subseteq V_{\gamma_0+1}$. Let φ_{OD} be as above. As $x \in OD$ there are $\psi, \beta_1, \ldots, \beta_n$ with $\{x\} = \{x \mid \psi(x,\vec{\beta})\}$. By the Reflection Theorem on φ_{OD} and ψ we can find $\gamma_1 > \gamma_0, \vec{\beta}$ with $z = \mathcal{P}(x) \cap OD \Leftrightarrow V_{\gamma_1} \models \exists x(\psi(x,\vec{\beta}) \wedge z = \{y \mid \varphi_{OD}(y) \wedge y \subseteq x\})$.

For AxSeparation: let a be a class term, and let $x \in HOD$. We require that $a^{HOD} \cap x \in HOD$. Suppose $a = \{z \mid \varphi(z,\vec{y})\}$ for some φ, some $\vec{y} \in HOD$. By the Reflection Theorem we can find a sufficiently large γ which is reflecting for φ and the defining formula for HOD, and with $a^{HOD} \cap x \in V_\gamma$. Then we have:

$$u = a^{HOD} \cap x \Leftrightarrow V_\gamma \models u = \{z \in x \mid \varphi(z,\vec{y})^{HOD}\}.$$

From the right hand side here, we see that u is definable over V_γ but using the parameters x and $\vec{y}$. However these are all in OD and so we may replace them by their (finitely many) definitions using just ordinal parameters, thereby rendering the right hand side a term purely with ordinal parameters. Hence u is in OD and thence in HOD (as $u \subseteq x \subseteq HOD$).

AxReplacement is similar: let F be a function given by a term, and let $x \in HOD$. We require $F^{HOD}``x \in HOD$. In $V$, by AxReplacement, let $F^{HOD}``x \subseteq V_\gamma$, but then it also is a subset of $V_\gamma \cap HOD$. It thus suffices to show $V_\gamma \cap HOD \in HOD$, as then the AxSeparation will separate out from $HOD \cap V_\gamma$ exactly the set $F^{HOD}``x$. This is the next Exercise.

Finally for AxChoice, note that we have the stronger principle of a Global Wellorder of HOD from which AC is obviously derivable. To define such a wellorder $<_{HOD}$, just restrict the Global Wellorder $<_{OD}$ (Theorem 4.22) to elements of $HOD \subseteq OD$. Q.E.D.

Exercise 4.19 Show that for any β, $V_\beta \cap HOD \in HOD$.

Exercise 4.20 Show that the following are equivalent: (i) $V = OD$, (ii) $V = HOD$, (iii) $Trans(OD)$, (iv) $(AxExt)^{OD}$. [Hint: Use that for any α $V_\alpha \in OD \wedge V_\alpha \cap OD \in OD$.]

Exercise 4.21 Show that $HOD \cap \mathcal{P}(\omega)$ is the largest subset of $\mathcal{P}(\omega)$ with a definable wellorder. [Hint: Use Lemma 4.23 and Ex. 4.18.]

Exercise 4.22 Suppose that W is a term defining an inner model of ZF and there is a definable global wellorder of W (that, as in L, there is a formula defining a wellorder $<_W$ of the whole of W in order type On). Show that $W \subseteq HOD$. (Consequently HOD is the largest inner model W with a definable bijection $F : On \leftrightarrow W$.)

Exercise 4.23 Define "Π_2-OD" (and Π_2-HOD) just as we did for OD and HOD but now restrict the formulae allowed in definitions to be Π_2 only. Show that Π_2-$OD = OD$ and Π_2-$HOD = HOD$. Now do the same for Σ_2-OD and Σ_2-HOD.

Exercise 4.24 * Show that there is a *single* formula $\varphi_0(v_0)$ with just the free v_0, so that OD is the class of all those x so that $x \in \mathrm{Def}_0(V_\beta)$ for some β, but only using φ_0; that is, the class of those x so that for some β, $\{x\} = \{z \mid V_\beta \models \varphi_0[z]\}$.

Again it is consistent that $V = L = HOD$, $V \neq L = HOD$ and $V \neq L \neq HOD$ as well as further combinations such as HOD^{HOD} may or may not equal HOD. CH may fail in HOD (see the next Exercise).

Exercise 4.25 $(*)(E))$ This shows that we may have $(\neg CH)^{HOD}$. Let $C_\alpha = \{n \in \omega \mid 2^{\aleph_{\alpha+n}} = \aleph_{\alpha+n+1}\}$. Suppose $|\{C_\alpha \mid \alpha \in On\}| \geq \aleph_2$ (this can be shown consistent with ZFC), then $(\neg CH)^{HOD}$.

We can define OD_x and HOD_x as before but now we allow sets $z \in x$ as parameters in our definitions as well as ordinals. HOD_x will be an inner model of ZF as before, but it will only be a model of Choice if there is an HOD_x-definable wellorder of x itself to start with.

Exercise 4.26 The Leibniz-Mycielski Principle (LM []) is the following:

$$\forall x \neq y(\exists^\ulcorner \varphi(v_0) \wedge \mathrm{FVbl}(\varphi) = \{v_0\} \wedge \exists \beta (x, y \in V_\beta \wedge \langle V_\beta, \in \rangle \models \varphi[x] \not\leftrightarrow \varphi[y]))).$$

Show that $V = OD$ implies LM.

4.6 Criteria for Inner Models

It is possible to give a definition for when a class term W is an inner model, $IM(W)$ of the ZF axioms which is formalisable in ZF. We first give an equivalent axiomatisation of ZF.

Definition 4.28 *We set* ZF* *to be the theory that consists of the Axioms* **Ax0-4**, **Ax7-8** *and:*
Ax5* (Δ_0-**Separation Scheme**) *For every Δ_0-term a:* $x \cap a \in V$
where by a Δ_0-term a we mean a term $a = \{x \mid \varphi(x, \vec{y})\}$ *where φ is a Δ_0-formula.*
Ax6* **(Collection Scheme)** *For every formula φ:* $\forall \vec{y} \exists v \varphi(v, \vec{y}) \longrightarrow \forall w \exists t (\forall \vec{y} \in w \exists v \in t \varphi(v, \vec{y}))$.

The weakening of the **Ax5** is made up for by the strengthening of **Ax6** which is less about the range of functions than 'collecting' together the ranges of relations on sets z. Note we could have expressed **Ax6***, arguably more awkwardly, as: "For any term r if $\forall y\, r``\{y\} \neq \varnothing$ then $\forall w \exists t \forall y \in w (r``\{y\} \cap t \neq \varnothing)$".

Theorem 4.29 ZF $\vdash$ ZF* *and* ZF* $\vdash$ ZF, *and thus the two theories are equivalent.*

Proof: (ZF $\vdash$ ZF*) As ZF already proves **AxSep** for all terms we only have to show that ZF proves the stronger Collection scheme. Suppose φ satisfies the antecedent, and let w be any set. By the argument of the Reflection Principle there is η so that $w \in V_\eta$ and $\forall \vec{y} \in w \exists v \varphi(v, \vec{y})$ is absolute between V and V_η. So we may take $t = V_\eta$.
(ZF* $\vdash$ ZF) We first show that ZF* $\vdash$ **Ax5**. Let φ be $Q_1 v_1 \cdots Q_n v_n \psi(v_0, \vec{v})$ where ψ is quantifier free, and we have taken (by Logic) φ in prenex normal form, and v_0 is the only free variable of φ. Let z_0 be any set. We wish to show that $\{v_0 \in z_0 \mid \varphi\} \in V$. By induction we define further sets z_i for $0 < i \leq n$. Assume z_{i-1} is defined.
Case 1 Q_i is $\forall$.

$$\forall v_0 \ldots v_{i-1} \in z_{i-1} (\exists v_i \neg Q_{i+1} v_{i+1} \cdots Q_n v_n \psi(v_0, \vec{v}) \to \exists v_i \in z'_{i-1} \neg Q_{i+1} v_{i+1} \cdots Q_n v_n \psi(v_0, \vec{v})).$$

Case 1 Q_i is $\exists$.

$$\forall v_0 \ldots v_{i-1} \in z_{i-1} (\exists v_i Q_{i+1} v_{i+1} \cdots Q_n v_n \psi(v_0, \vec{v}) \to \exists v_i \in z'_{i-1} Q_{i+1} v_{i+1} \cdots Q_n v_n \psi(v_0, \vec{v})).$$

Then in either case the existence of z'_{i-1} is implied by the Collection scheme, and we may let $z_i = z_{i-1} \cup z'_{i-1}$. We thus have, again for both cases:

$$\forall v_0 \ldots v_{i-1} \in z_{i-1} (Q_i v_i Q_{i+1} v_{i+1} \cdots Q_n v_n \psi(v_0, \vec{v}) \leftrightarrow Q_i v_i \in z_i Q_{i+1} v_{i+1} \cdots Q_n v_n \psi(v_0, \vec{v})).$$

Hence:

$$\begin{aligned} \forall v_0 \in z_0 (Q_1 v_1 \cdots Q_n v_n \psi(v_0, \vec{v}) &\leftrightarrow \\ &\leftrightarrow Q_1 v_1 \in z_1 Q_2 v_2 \cdots Q_n v_n \psi(v_0, \vec{v}) \\ &\vdots \\ &\leftrightarrow Q_1 v_1 \in z_1 \cdots Q_n v_n \in z_n \psi(v_0, \vec{v})). \end{aligned}$$

Consequently the term $\{v_0 \in z_0 \mid \varphi\} = \{v_0 \in z_0 \mid Q_1 v_1 \in z_1 \cdots Q_n v_n \in z_n \psi(v_0, \vec{v})\}$. By Δ_0-Separation on the formula defining the right hand term, we have that $\{v_0 \in z_0 \mid \varphi\}$ is a set as required.

We need to show that **AxRep** is derivable in ZF*. Let $Fun(f) \wedge w \in V$. Suppose $f = \{\langle x, y\rangle \mid \varphi(x, y)\}$. As $Fun(f)$ then $\forall x \exists y \varphi(x, y)$. By Collection then there is $t \in V$ so that $\forall x \in w \exists y \in t \varphi(x, y)$. Then $f``w = \{y \in t \mid \exists x \in w \varphi(x, y)\}$. The latter is a set in V since we have just proven full Separation.

Q.E.D.

THEOREM 4.30 *If W is any term, let $im_0(W)$ be the sentence: "$\forall x \subseteq W \exists y (x \subseteq y \wedge Trans(y) \wedge \mathrm{Def}(\langle y, \in\rangle) \subseteq W$". Then $IM(W)$ is equivalent to $im_0(W)$. That is $IM(W) \vdash_{\mathrm{ZF}} im_0(W)$ and conversely $im_0(W) \vdash_{\mathrm{ZF}} IM(W)$.*

PROOF: ($IM(W) \vdash_{\mathrm{ZF}} im_0(W)$): Since ZF^W (by $IM(W)$) we have that $(\forall x \exists \alpha (x \in V_\alpha))^W$. Thus the rank function ρ^W is definable in W, but it is also absolute, and thus $\rho^W = \rho$. Let $x \subseteq W$ and let $z = \rho``x$. Then $z \in V$ by the **AxRep** applied in V to ρ. Let $\eta = \sup z$. Then $\eta \in On$. Note that if $y = V^W_\eta$ then $Trans(y)$ (as V^W_η is transitive in W which itself is transitive). Lastly $\mathrm{Def}(\langle y, \in\rangle) \subseteq W$, as by absoluteness of the Def function $\mathrm{Def}(\langle y, \in\rangle) = (\mathrm{Def}(\langle y, \in\rangle))^W$.
($im_0(W) \vdash_{\mathrm{ZF}} IM(W)$):
(1) $Trans(W)$.
PROOF: Note first $im_0(W)$ implies that $W \neq \varnothing$. Let $x \in W$. Then $\{x\} \subseteq V$ and by $im_0(W)$, there is y with $x \in y \wedge Trans(y) \wedge \mathrm{Def}(\langle y, \in\rangle) \subseteq W$. As $Trans(y)$ we then have $x \subseteq y$. Let $z \in x$ and so $z \in y$. Then $z = \{t \in y \mid t \in z\}$ (as $z \in y \wedge Trans(y)$). This is equal to $\{t \in y \mid (t \in z)^y\}$ by the absoluteness of $\in$ to transitive sets, and so equals

$$\{t \in y \mid \langle y, \in\rangle \models (v_0 \dot{\in} v_1)[t, z]\} \in \mathrm{Def}(\langle y, \in\rangle) \subseteq W.$$

Therefore $z \in W$ as required.
(2) $On \subseteq W$.
PROOF: As $Trans(W)$ either $On^W = On$ (in which case we are done) or $On^W = \eta$ for some $\eta \in On$. Then $im_0(W)$ implies there is a transitive $y \subseteq W$ with $\eta \subseteq y$, and $\mathrm{Def}(\langle y, \in\rangle) \subseteq W$. Then $\eta = \{z \in y \mid z \in On\}$ (as $y \subseteq W$). But then $\eta = \{z \in y \mid \langle y, \in\rangle \models (v_0 \dot{\in} \dot{O}n)[z]\}$. But $\mathrm{Def}(\langle y, \in\rangle) \subseteq W$ and so $\eta \in W$, contradicting that $\eta = On^W$.

We need to show that $(\mathrm{ZF})^W$ holds. By the previous theorem it suffices to show $(\mathrm{ZF}^*)^W$. That $Trans(W)$ and $On \subseteq W$ already yields that AxEmpty and AxInf hold in W; Pairing, Union, Power, Foundation and Δ_0-Separation in W can be left as exercises.
(3)(**Ax6***)^W.
PROOF: Let $\varphi(x, y, \vec{p})$ be a formula of ZF with free variables shown. Then:

$$(\forall \vec{p} (\forall x \exists y \varphi \to \forall u \exists v \forall x \in u \exists y \in v \varphi))^W \leftrightarrow$$
$$\forall \vec{p} \in W (\forall x \in W \exists y \in W \varphi^W \to \forall u \in W \exists v \in W \forall x \in u \exists y \in v \varphi^W).$$

So, taking $\vec{p}, u \in W$, we assume $\forall x \in W \exists y \in W \varphi^W$. By Collection in V there is a set $t \in V$ so that $\forall x \in u \exists y \in t (y \in W \wedge \varphi^W)$. By $im_0(W)$ there is a transitive $z \supseteq t \cap W$, and with $\mathrm{Def}(\langle z, \in\rangle) \subseteq W$. But then $z \in W$ and so $\forall x \in u \exists y \in z \varphi^W$.

Q.E.D.

The utility of the last theorem is that often it is a simple matter to verify for any given W the three assertions that it is transitive, contains all ordinals and is a model of $im_0(W)$. For example, HOD is easily seen to have these three properties. Whilst the statement "ZF^W" (or indeed "$IM(W)$") is metatheoretic in nature: it requires assertions of the infinitely many formulae contained in "ZF^W". The theorem shows that within ZF, that a term W is an inner model is truly a first order expression about W.

Exercise 4.27 Show that there is a finite set of axioms of ZF so that if $On \subseteq W$ and W is a transitive class model of just these axioms then it is a model of all the axioms of ZF. Why does this not contradict the non-finite axiomatisability of ZF, Theorem 3.10?

Exercise 4.28 Show that if M is a class term, and ZF proves $IM(M)$ and $(\neg CH)^M$, then ZF is inconsistent.

Exercise 4.29 Call a set x *non-typical* if $\exists y(x \in y \in OD \wedge |y| = \omega)$, and write $x \in NT$. Say, as usual, that a set x is *hereditarily non-typical*, and write $X \in HNT$, if $x \in NT$ and $TC(x) \subseteq NT$. Show that (i) the class $HNT \supseteq HOD$; (ii) ZF^{HNT}. (It need not be the case that AC^{HNT}; (Tzouvaras [**?**]).)

4.6.1 Further examples of inner models

Relative constructibility

There are several ways to generalise Gödel's construction of L.

(I) The $L(A)$-hierarchy.

Here we start out, not with the empty set as L_0 but with the set A:

Definition 4.31

$$\begin{aligned} L_0(A) &= A \cup \{A\}; \\ L_{\alpha+1}(A) &= \mathrm{Def}(\langle L_\alpha(A), \in\rangle); \\ \mathrm{Lim}(\lambda) \to L_\lambda(A) &= \bigcup\{L_\alpha(A) \mid \alpha < \lambda\}. \\ L(A) &= \bigcup\{L_\alpha(A) \mid \alpha < \mathrm{On}\}. \end{aligned}$$

In this model the arguments for L can be straightforwardly used to show that all axioms of ZF are valid in $L(A)$. However the Axiom of Choice need not hold, unless in $L(A)$ there is a $L(A)$-definable wellorder of A. Of course if $V = L$ then $A \in L$ and the construction of L inside the ZF-model $L(A)$ reveals that "$V = L$" holds, in which case $AC^{L(A)}$ trivially holds. Matters become more interesting when $V \neq L$, and an important model here is when $A = \mathbb{R}$. The model $L(\mathbb{R})$ contains all the reals (and so the structure of mathematical analysis). Consequently anything definable in the structure of analysis resides in the model. Moreover anything obtained by 'iterated definability over analysis' is also here: it would be definable using ordinals and the set of reals. Thus it is thought, the broadest methods of definability over analysis would produce sets in this model. Consequently it is in some sense a laboratory for generalised definability in analysis. However it is not thought in general that there must be wellorder of $\mathbb{R}$ that is *definable* over $\mathbb{R}$, or indeed in $L(\mathbb{R})$. (This was one approach that Cantor took to look at CH: to try to find a definable wellorder of $\mathbb{R}$; but it is consistent with the axioms of ZF that there is no such wellorder.) Consequently when. set theorists investigate $L(\mathbb{R})$ they do not assume that AC holds there, although it is taken to hold hold in the wider universe V.

(II) The $L[A]$-hierarchy.

The next hierarchy instead enlarges the language of set theory to incorporate a one place predicate symbol $\dot{A}$. Thus $A(x)$ either will or will not be true of sets x. The Def operator is enlarged to an operator $\mathrm{Def}_{\dot{A}}$ that now defines new sets over some structure in this new language

Definition 4.32

$$\begin{array}{rcl} L_0[A] & = & \varnothing; \\ L_{\alpha+1}[A] & = & \mathrm{Def}_{\dot{A}}(\langle L_\alpha[A], \in, A\rangle); \\ \mathrm{Lim}(\lambda) \to L_\lambda[A] & = & \bigcup\{L_\alpha[A] \mid \alpha < \lambda\} \\ L[A] & = & \bigcup\{L_\alpha[A] \mid \alpha < \mathrm{On}\}. \end{array}$$

The predicate A is usually taken to be a set in V, but the definition is perfectly good, and can be formulated in ZF if A is a definable proper class of sets. In either case A may impose quite a 'wild' behaviour on the model $L[A]$. That is not the case for the following very important inner model: unlike L, this model can accommodate the large cardinal called a '*measurable cardinal*'.

Definition 4.33 *($L[\mu]$) $L[\mu]$ is the above hierarchy where μ is a κ-*complete ultrafilter *on $\mathcal{P}(\kappa)$ in the sense of the discussion at the end of Section 2.1.2.*

The inner model $L[\mu]$ is much studied (μ is a κ-*complete ultrafilter on* $\mathcal{P}(\kappa))^{L[\mu]}$ and moreover, is the least inner model with this property. It has an absolute construction property similar to L within in any other inner model with such a ultrafilter or 'measure' on κ. It can be shown that $(\mathrm{GCH})^{L[\mu]}$, although the Condensation Lemma strictly speaking, fails in $L[\mu]$.

Exercise 4.30 Show for any A that $(\mathrm{ZF})^{L(A)}$ and that $(\mathrm{ZFC})^{L[A]}$. [Hint: Just modify the same arguments for L.]

Exercise 4.31 (i) Show that in $L[A]$, for $A \subseteq \kappa$, that for any $\gamma \geq \kappa$, $2^\gamma = \gamma^+$. (Thus, in $L[A]$ the GCH holds 'above κ'.) [Hint: Again modify the argument for L; this can only work above κ since A could be completely general, and we have no knowledge how $L_\kappa[A]$ may look.]
(ii) However improve the last exercise, by showing that in $L[A]$, for $A \subseteq \kappa = \delta^+$, that for any $\gamma \geq \delta$, $2^\gamma = \gamma^+$.

Higher Order Constructibility

We do not give the details, but for the reader familiar with notions of higher order logics, in particular n'th-order logics for $n < \omega$, we may construct L^n using n'th order logical definablility $\mathrm{Def}^{\mathrm{n}}$ (where our previous Def is now Def^1. Remarkably these notions do not form a hierarchy for $n \geq 3$, but instead all collapse:

Theorem 4.34 (Myhill-Scott) *For $n \geq 2$ $L^n = HOD$.*

4.7 The Suslin Problem

It is well known (in fact it is a theorem of Cantor) that if $\langle X,<\rangle$ is a totally ordered continuum that satisfies

(i) $\langle X,<\rangle$ *has no first or last end points* ;

(ii) $\langle X,<\rangle$ *has a* countable dense subset Y *(that is* $\forall x,z\in X\exists y\in Y(x<y<z)$*)*;

then $\langle X,<\rangle$ *is isomorphic to* $\langle\mathbb{R},<\rangle$.

(By *continuum* one requires that for any bounded subset of an interval in $\langle X,<\rangle$ has a supremum in X (and likewise an infimum in X.)

Suslin asked (1925) whether (ii) could be replaced with the seemingly weaker

(iii) $\langle X,<\rangle$ has the *countable chain condition (c.c.c.)* (that, if $I_\alpha=(x_\alpha,y_\alpha)$ for $\alpha<\omega_1$ is a family of open intervals in $\langle X,<\rangle$ then $\exists\alpha\exists\beta(I_\alpha\cap I_\beta\neq\varnothing)$).

Notice that (ii) implies (iii) : every open interval I_α must contain an element of Y; however Y only has countably many elements.

The question is thus: do (i) and (iii) also characterise the real line $\langle\mathbb{R},<\rangle$? Suslin hypothesised that they did. This became known as Suslin's hypothesis (SH). The problem can be reduced to the following question concerning *trees* on ordinals.

Definition 4.35 *A* tree $\langle T,<\rangle$ *is a partial ordering such that* $\forall x\in T(\{y\mid y<x\})$ *is wellordered.*

(i) The height *of* x *in* T *,* $\mathrm{ht}(x)$*, is* $\mathrm{ot}(\{y\mid y<x\},<)$ *(also called the* rank *of* x *in* T*).*

(ii) The height *of* T *is* $\sup\{\mathrm{ht}(x)\mid x\in T\}$*;*

(iii) $T_\alpha=_{\mathrm{df}}\{x\in T\mid \mathrm{ht}(x)=\alpha\}$.

Thus T_0 consists of the bottommost elements of the tree, and so are called *root(s)* (we shall assume there is only one root). A *chain* in any partial order $\langle T,<_T\rangle$ is any subset of T linearly ordered by $<_T$ and an *antichain* is any subset of T no two elements of which are $<_T$ −comparable. For a tree T a subset $b\subseteq T$ is a *branch* if it is a maximal linearly ordered (and so wellordered) set under $<_T$. A branch need not necessarily have a top-most element of course.

Definition 4.36 *Let* κ *be a regular cardinal. A* κ - Suslin tree *is a tree* $\langle T,<\rangle$ *such that*

(i) $|T|=\kappa$ *;*

(ii) Every chain and antichain in T *has cardinality* $<\kappa$.

We shall be concerned with ω_1- Suslin trees (and we shall drop the prefix "ω_1"). König's Lemma states that every countable tree with nodes that "split" finitely, has an infinite branch. This paraphrased says, *a fortiori*, that there are no ω-Suslin trees.

It turns out (see Devlin [1]) that the Suslin Hypothesis is equivalent to:

(SH): "*There are no* ω_1*-Suslin trees*"

(Although this requires proof which we omit.) So do such trees exist?

Theorem 4.37 *(Jensen) Assume* $V=L$*; then there is an* ω_1*-Suslin tree.*

Hence:

Corollary 4.38 $\mathrm{Con(ZF)}\Rightarrow\mathrm{Con(ZFC+CH+\neg\,SH)}$

It turns out that there is a construction principle for Suslin trees that is in itself of immense interest: it can be considered a strong form of the Continuum Hypothesis. It has been widely used in set theory and topology and has been much studied.

Definition 4.39 *(The Diamond Principle).* $\diamondsuit$ *is the assertion that there exists a sequence* $\langle S_\alpha | \alpha < \omega_1 \rangle$ *so that (i)* $\forall \alpha (S_\alpha \subseteq \alpha)$
(ii) $\forall X \subseteq \omega_1 \{\alpha | X \cap \alpha = S_\alpha\}$ *is stationary.*

$\diamondsuit$ thus asserts that there is a single sequence of S_α's that approximate any subset of ω_1 "very often". In particular note that $\diamondsuit \longrightarrow$ CH: if $x \subseteq \omega$ is any real then $x = S_\alpha$ for "stationarily" many $\alpha < \omega_1$. Thus the $\diamondsuit$ sequence incorporates an enumeration of the real continuum with each real occurring ω_1 many times in that enumeration. However it does much more beside.

Theorem 4.40 *(Jensen) In L,* $\diamondsuit$ *holds. That is* $\mathrm{ZF} \vdash (\diamondsuit)^L$.

Proof: Assume $V = L$. We have to define a $\diamondsuit$-sequence $\langle S_\alpha | \alpha < \omega_1 \rangle$. We define by recursion $\langle S_\alpha, C_\alpha \rangle$ for $\alpha < \omega_1$: $\langle S_\alpha, C_\alpha \rangle$ is the $<_L$-least pair of sets $\langle S, C \rangle$ so that
(a) $S_\alpha \subseteq \alpha$
(b) C is c.u.b. in α ;
(c) $\forall \beta \in C (S \cap \beta \neq S_\beta)$
if there is such a pair, and $\langle S_\alpha, C_\alpha \rangle = \langle \varnothing, \varnothing \rangle$ otherwise.

Thus, somewhat paradoxically, $\langle S_\alpha, C_\alpha \rangle$ is chosen to be the $<_L$- least "counterexample" to a $\diamondsuit$-sequence of length α.

Let $\mathcal{S} = \langle S_\alpha | \alpha < \omega_1 \rangle$. As we are assuming $V = L$, we have just constructed $\mathcal{S} \in H_{\omega_2} = L_{\omega_2}$. Looking a little more closely, since $\mathcal{P}(\omega_1) \subseteq L_{\omega_2}$, we have actually defined $\mathcal{S}$ by a recursion which only involved inspecting objects in L_{ω_2} which had certain definite properties. L_{ω_2} is a model of ZF$^-$ so these properties are absolute between L_{ω_2} and V which is L by assumption. In short the recursion as defined in L_{ω_2} defines the same $\mathcal{S}$ as in V: $\forall \alpha < \omega_2 (\langle S_\alpha, C_\alpha \rangle)_{L_{\omega_2}} = \langle S_\alpha, C_\alpha \rangle$ and indeed $(\mathcal{S})_{L_{\omega_2}} = \mathcal{S}$.

If $\mathcal{S}$ is not a $\diamondsuit$-sequence then:
(1) *There is an* $<_L$*-least pair* $\langle S, C \rangle$ *with*
(a) $S \subseteq \omega_1$; (b) $C \subseteq \omega_1$ *and* C *cub in* ω_1; (c) $\forall \beta \in C (S \cap \beta \neq S_\beta)$.

Given that we have $\mathcal{S} \in L_{\omega_2}$ the quantifiers in (1) are referring only to sets in L_{ω_2}. (1) thus holds relativised to L_{ω_2}. Expressing that in semantical terms we have:
(2) $\langle L_{\omega_2}, \in \rangle \models$" $\langle S, C \rangle$ *is the* $<_L$*-least pair with*
(a) $S \subseteq \omega_1$; (b) $C \subseteq \omega_1$ *and* C *cub in* ω_1; (c) $\forall \beta \in C (S \cap \beta \neq S_\beta)$."

By appealing to the Löwenheim-Skolem Theorem we can find $X \subseteq L_{\omega_2}$ with:
(3) $\langle X, \in \rangle \prec \langle L_{\omega_2}, \in \rangle$ *with* $\mathcal{S}, \langle S, C \rangle, \omega_1 \in X, \omega \subseteq X$, *and* $|X| = \omega$.

By Exercise 4.13 (ii) we have that $X \cap L_{\omega_1}$ is transitive and so in fact is some L_γ for some $\gamma < \omega_1$. If we now apply the Mostoski-Shepherdson Collapsing Lemma we have there is a π and a τ with:
(4) $\pi : \langle L_\tau, \in \rangle \cong \langle X, \in \rangle$ *with* $\pi \restriction L_\gamma = \mathrm{id}$.
(Recall that as $L_\gamma \subseteq X$ and is transitive π will be the identity on L_γ.)
(5) $\pi(\gamma) = \omega_1$, *and if* $\overline{S}, \overline{C}$ *are such that* $\pi(\overline{S}) = S, \pi(\overline{C}) = C$, *then* $\overline{S} = S \cap \gamma, \overline{C} = C \cap \gamma$.

Proof: $\pi^{-1}(\omega_1) = \{\pi^{-1}(\xi) \mid \xi \in \omega_1 \cap X\}$

$= \{\xi \mid \xi \in \omega_1 \cap X\}$
$= \omega_1 \cap X = \gamma.$

Similarly $\overline{S} = \pi^{-1}(S) = \{\pi^{-1}(\xi) \mid \xi \in S \cap X\}$
$= \{\pi^{-1}(\xi) \mid \xi \in S \cap \gamma\}$
$= \{\xi \mid \xi \in S \cap \gamma\}$ (using (4))
$= S \cap \gamma.$

That $\overline{C} = C \cap \gamma$ is entirely the same. Q.E.D.(5)

(6) *If* $\overline{\mathcal{S}} = \pi^{-1}(\mathcal{S})$ *then* $\overline{\mathcal{S}} = \mathcal{S} \upharpoonright \gamma$.

Proof: Note that $\pi^{-1}(\langle S_\xi | \xi < \omega_1\rangle) = \pi^{-1}(\{\langle \xi, S_\xi\rangle \mid \xi < \omega_1\})$
$= \{\pi^{-1}(\langle \xi, S_\xi\rangle) \mid \xi < \pi^{-1}(\omega_1)\}$
$= \{\langle \pi^{-1}(\xi), \pi^{-1}(S_\xi)\rangle \mid \xi < \gamma\}$
$= \{\langle \xi, S_\xi\rangle \mid \xi < \gamma\}$ since both $\xi, S_\xi \in L_\gamma$.

Similar equalities hold for $\pi^{-1}(\langle C_\xi | \xi < \omega_1\rangle)$. Hence $\pi^{-1}(\langle\langle S_\xi, C_\xi\rangle | \xi < \omega_1\rangle) = \mathcal{S} \upharpoonright \gamma$. Q.E.D.(6)

Appealing to (2) and (4) we have:

(7) $\langle L_\tau, \in\rangle \models$" $\langle \overline{S}, \overline{C}\rangle$ *is the* $<_L$*-least pair with*
(a) $\overline{S} \subseteq \gamma$; (b) $\overline{C} \subseteq \gamma$ *and* $\overline{C}$ *cub in* γ; (c) $\forall \beta \in \overline{C}(\overline{S} \cap \beta \neq S_\beta)$."

As $<_{L_\tau} = (<_L)_{L_\tau}$ and $<_L$ is an end-extension of $<_{L_\tau}$ and since (a)-(c) are absolute for transitive ZF$^-$ models, we have that (a)-(c) are really true in V of $\overline{S}, \overline{C}$, *i.e.* :

(8) $\langle \overline{S}, \overline{C}\rangle$ *is the* $<_L$*-least pair with*
(a) $\overline{S} \subseteq \gamma$; (b) $\overline{C} \subseteq \gamma$ *and* $\overline{C}$ *cub in* γ; (c) $\forall \beta \in \overline{C}(\overline{S} \cap \beta \neq S_\beta)$.

That is, $\overline{S}, \overline{C}$ really are the candidates to be chosen at the next, γ'th, stage of the recursion:

(9) $\langle \overline{S}, \overline{C}\rangle = \langle S_\gamma, C_\gamma\rangle$.

Now note that $\gamma \in C$ as $\overline{C} = C \cap \gamma$ is unbounded in the closed set C. Also, using (5), $S \cap \gamma = \overline{S} = S_\gamma$. This contradicts (1)! Q.E.D.

Exercise 4.32 (*) Formulate a principle $\diamondsuit_\kappa$ which asserts similar properties for a sequence $\langle S_\alpha \mid \alpha < \kappa\rangle$ where κ is any regular cardinal, and prove that it holds in L

Exercise 4.33 (**) Show that $\diamondsuit$ implies the existence of a family $\langle A_\xi \mid \xi < \omega_2\rangle$ of stationary subsets of ω_1, such that the intersection of any two of them is countable.

Theorem 4.41 *(Jensen)* $\diamondsuit$ *implies the existence of a Suslin tree.*

Proof: We shall construct by recursion a tree T of cardinality ω_1, using countable ordinals. In fact we shall have that $T = \omega_1$ itself, the construction thus delivers $<_T$. T will be the union of its levels T_α all of which will be countable, and $<_T = \bigcup_{\alpha<\omega_1} <_{T_{\leq\alpha}}$ where (a) $T_{<\alpha} = \bigcup_{\beta<\alpha} T_\beta$ and (b) $<_{T_{<\alpha}}$ is the tree ordering constructed so far on $T_{<\alpha}$. We shall ensure that every $<_T$-branch is countable, and likewise every maximal antichain. Then $\langle T, <_T\rangle$ will be Suslin. The recursion will ensure a *normality* condition: for every $\xi \in T$, and if $\xi \in T_\alpha$, then for every $\alpha < \beta < \omega_1$ there is $\zeta \in T_\beta$ with $\xi <_T \zeta$; every node then in the tree has tree-successors of arbitrary height below ω_1.

We let $T \upharpoonright 1 = T_0 = \{0\}$ and $T_{<1} = \varnothing$. Assume $\mathrm{Lim}(\alpha)$ and $T_\beta, <_{T_{<\beta}}$ defined for all $\beta < \alpha$. Then $T \upharpoonright \alpha = \bigcup_{\beta<\alpha} T_\beta$ and $<_{T_{<\alpha}} = \bigcup_{\beta<\alpha} <_{T_{<\beta}}$. Normality as described above, is then trivially conserved.

Assume now $\alpha = \beta + 1$ and that $\mathrm{Succ}(\beta)$. We assume that $T \upharpoonright \beta$, and $<_{T_{<\beta}}$ have been defined. We thus have defined T_γ where $\gamma + 1 = \beta$. For each $\xi \in T_\gamma$ we allot in turn the next ω sequence of ordinals available $\{\xi_i \mid i < \omega\}$. (We thus go through T_γ say by induction on the ordinals $\xi \in T_\gamma$ and we define $<_{T_{<\alpha}}$ by adding to the ordering $<_{T_{<\beta}}$ (which equals in the obvious sense $<_{T_{\le\gamma}}$) the pairs $\langle \xi, \xi_i \rangle$ (and also the pairs $\langle \zeta, \xi_i \rangle$ for those $\zeta <_{T_{\le\gamma}} \xi$ to complete the ordering.) Thus at successor stages of the tree it is infinitely branching. Again normality is obvious. This defines $T \upharpoonright \alpha$ and $<_{T_{\le\beta}} = <_{T_{<\alpha}}$.

Finally if $\alpha = \beta + 1$ but $\mathrm{Lim}(\beta)$ we need to define $T \upharpoonright \alpha$ and make a careful choice of which maximal branches through $T_{<\beta}$ (thus those of order type β) that we may extend with impunity to have nodes at level β, *i.e.* in T_β, thus fixing $<_{T_{<\alpha}}$. This is where we use $\diamond$.

Case 1 $S_\beta (\subseteq \beta)$ is a maximal antichain in the tree so far defined: $\langle T_{<\beta}, <_{T_{<\beta}} \rangle$.

In this case for any $\xi \in T_{<\beta}$ there must be some $\sigma \in S_\beta$ with either $\sigma <_{T_{<\beta}} \xi$ or $\xi \le_{T_{<\beta}} \sigma$. Either way by the normality of the tree $\langle T_{<\beta}, <_{T_{<\beta}} \rangle$ so far, we pick a branch b_ξ through $T_{<\beta}$ with both $\sigma, \xi \in b_\xi$. Let $B = \{b_\xi \mid \xi \in T_{<\beta}\}$. This is a countable set of branches. We enumerate B as $\{b_n \mid n < \omega\}$ and choose the next ω many ordinals ξ_n for $n < \omega$, with $\xi_n \notin T_{<\beta}$. We extend the branch b_n to have ξ_n as a final node, and enlarge $<_{T_{<\beta}}$ appropriately to $<_{T_{<\alpha}}$. (Thus if ζ is on the branch b_n extended with the new point ξ_n,we add the ordered pair $\langle \zeta, \xi_n \rangle$ to $<_{T_{<\beta}}$; we thus obtain $<_{T_{<\alpha}}$.) Then we have $T_\beta = T_{<\beta} \cup \{\xi_n \mid n < \omega\}$ and so we have $T \upharpoonright \alpha$. By construction again we preserve normality: every $\zeta \in T_{<\alpha}$ has a successor in T_β.

Case 2 Otherwise.

Then we let T_β be any set consisting of the next ω many ordinals not used so far, and extend the ordering of $<_{T_{<\beta}}$ to T_β in any fashion as long as normality is preserved. (In other words we can just enumerate $T_{<\beta}$ as $\langle \zeta_n | n < \omega \rangle$ and go through adding on some new ordinals ξ_n to some branch through ζ_n that has order type β -if need be- as long as we ensure ζ_n has *some* successor at height β.

This ends the construction. We claim that if we set $T = \bigcup_{\alpha<\omega_1} T_\beta$ and $<_T = \bigcup_{\alpha<\omega_1} <_{T_\beta}$ then $\langle T, <_T \rangle$ is a Suslin tree. First we see that it has no uncountable antichain. Suppose there were such, and let $A \subseteq \omega_1$ be a maximal uncountable antichain (which exists by Zorn's Lemma).

Claim $C = \{\alpha \mid A \cap T_{<\alpha}$ is a maximal antichain in $T_{<\alpha}\}$ is cub in ω_1.

Proof: Let $\beta_0 < \omega_1$ be arbitrary. As $T_{<\beta_0}$ is countable, there exists $\beta_1 < \omega_1$ with every element of $T_{<\beta_0}$ compatible with some element of $A \cap T_{<\beta_1}$. Repeating this, we find $\beta_2 > \beta_1$ so that every element of $T_{<\beta_1}$ compatible with some element of $A \cap T_{<\beta_2}$; and similarly $\beta_{n+1} > \beta_n$ so that every element of $T_{<\beta_n}$ compatible with some element of $A \cap T_{<\beta_{n+1}}$. If $\gamma = \sup_n \beta_n$ then $A \cap T_{<\gamma}$ is a maximal antichain in $T_{<\gamma}$. C is thus unbounded in ω_1. That C is closed is immediate. Q.E.D. *Claim*

By our requisite property that $\langle S_\alpha \mid \alpha < \omega_1 \rangle$ is a $\diamond$-sequence, now that C is cub and $A \subseteq \omega_1$, there must be $\beta \in C$ with $S_\beta = A \cap \beta$. Thus S_β is a maximal antichain in $<_{T_{<\beta}}$. However at precisely this point in the construction we would have chosen T_β so that every element of $T_{<\beta}$, and so every element of $A \cap T_{<\beta}$, has a tree successor at height β in T_β. Note that all elements of the tree at greater heights $\delta > \beta$ are extensions of the tree above these elements on T_β. Thus $A \cap \beta$ is a maximal antichain in $<_T$! But $A \cap \beta$ must be A and be countable! Contradiction! Q.E.D.

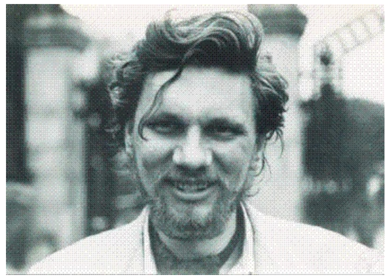
Exercise 4.34 (**) Show that $\diamondsuit$ implies the existence of two non-isomorphic Suslin trees.

One could further ask whether SH depends on CH. It is completely independent of CH as the following states.

- Con(ZF) implies the consistency of any of the following theories:

ZF + CH + SH; ZF+CH+¬ SH; ZF+¬ CH +¬ SH: ZF+¬ CH + SH

The second of these is Cor. 4.38 above. The other consistencies can be shown by using variations on Cohen's forcing methods, for which see [4]. Some of the arguments are very subtle.



Ronald Jensen

Appendix A

Logical Matters

A.1 The formal languages - syntax

We outline formal first order languages of predicate logic with axioms for equality. We do this for our language $\mathcal{L} = \mathcal{L}_{\in}$ which we shall use for set theory, but it is completely general:

(i) set variables; $v_0, v_1, \ldots, v_n, \ldots$ (for $n \in \mathbb{N}$)

(ii) two binary predicates: $\dot{=}, \dot{\in}$; an optional n-ary relation symbol $\dot{R}v_1 \cdots v_n$ (other languages would contain further function symbols $\dot{F}_i$ and relations symbols $\dot{R}_j$ of different -arities).

(iii) logical connectives: $\vee, \neg$

(iv) brackets: (,)

(v) an existential quantifier: $\exists$.

A formula is finite string of our symbol set; the formulae *of $\mathcal{L}$ (*'Fml'*) are defined inductively in a way similar for any first order language.*

1) $x = y$ and $x \in y$ are the atomic formulae *where x, y stand for any of the variables v_i, v_j. (If we opt for variants where we have the relation or function symbols, then $Rv_1 \cdots v_n$ and $Fv_1 \cdots v_n = v_{n+1}$ are also atomic.)*

2) Any atomic formula is a formula;

3) If φ and ψ are formulae then so is $\neg\varphi$ and $(\varphi \vee \psi)$, $\exists x \varphi$ where x is any variable;

4) φ is only a formula if it is so by repeated applications of 1)-3).

Inherent in the induction is the idea that a formula has *subformulae* and that a formula is built up from atomic formulae according to some *finite* tree structure. Further, given the formula we may identify the unique tree structure. Indeed we think of this as an *algorithm* that given a symbol string tests whether it is a formula by winding the recursion backwards to try to discover the underlying tree structure. Using this fact we can then perform recursions over the class of formulae using the clauses 1)-3) as part of our recursive definition. Clause 4) then ensures that our recursion will cover all formulae.

Definition A.1 *For φ a formula we define*

(A) the set of variables of φ, $\mathrm{Vbl}(\varphi)$ *by:*

$\mathrm{Vbl}(v_i = v_j) = \mathrm{Vbl}(v_i \in v_j) = \{v_i, v_j\}; \mathrm{Vbl}(Rv_1 \cdots v_n) = \{v_1, v_2, \dots, v_n\};$
$\mathrm{Vbl}(\neg\varphi)) = \mathrm{Vbl}(\varphi);\ \ \mathrm{Vbl}((\varphi \vee \psi)) = \mathrm{Vbl}(\varphi) \cup \mathrm{Vbl}(\psi);\ \mathrm{Vbl}(\exists x\varphi) = \mathrm{Vbl}(\varphi) \cup \{x\}.$

(B) the set of free variables *of* φ, $\mathrm{FVbl}(\varphi)$ *would be obtained exactly as above but changing the clause for* $\exists x\varphi$ *to:* $\mathrm{FVbl}(\exists x\varphi) = \mathrm{FVbl}(\varphi) - \{x\}$

(C) φ *is a* sentence *if* $\mathrm{FVbl}(\varphi) = \varnothing$.

By the above remarks in (B) we have defined the free variable set for all formulae. Note the crucial very final clause in (B) concerning the $\exists$ quantifier. The set of official *logical connectives* is minimal, it is just $\neg$ and $\vee$. But it is well known that the other connectives, $\wedge, \to, \leftrightarrow$ can be defined in terms of them, as can $\forall$, from $\exists$ and $\neg$. We shall use formulae freely involving these connectives, without comment. Here is another example.

Definition A.2 *For* φ *a formula, we define the* set of subformulae of φ, $\mathrm{Subfml}(\varphi)$, *by:*
$\mathrm{Subfml}(v_i = v_j) = \mathrm{Subfml}(v_i \in v_j) = \mathrm{Subfml}(Rv_1 \cdots v_n) = \varnothing;$
$\mathrm{Subfml}(\neg\varphi)) = \mathrm{Subfml}(\exists x\varphi) = \mathrm{Subfml}(\varphi) \cup \{\varphi\};$
$\mathrm{Subfml}((\varphi \vee \psi)) = \mathrm{Subfml}(\varphi) \cup \mathrm{Subfml}(\psi) \cup \{\varphi, \psi\}.$

Deductive systems

A deductive system of *predicate calculus* is (I) a set of axioms from which we can make pure logical deductions together with (II) those rules of deduction. There are many examples. The following is the simplest to explain (but rather difficult to use naturally) but this allows us to prove things about the system as simply as possible.

(I) Axioms of predicate calculus (for a language with relational symbols, and equality):
For any variables x, y and any φ, ψ, χ in Fml:
$\varphi \to (\psi \to \varphi)$
$(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$
$(\neg\psi \to \neg\varphi) \to ((\neg\psi \to \varphi) \to \psi)$
$\forall x\phi(x) \to \phi(y/x)$ where y is *free for* x (this has a slightly technical meaning).
$\forall x(\varphi \to \psi) \to (\varphi \to \forall x\psi)$ (where $x \notin \mathrm{Fr}(\varphi)$)
$\forall x(x = x)$
$x = y \to (\varphi(x, x) \to \varphi(x, y))$
(II) Rules of Deduction
(1) *Modus Ponens*. From $(\varphi \to \psi)$ and φ deduce: ψ.
(2) *Universal Generalisation*: From φ deduce $\forall x\varphi$.

In general a *theory* is a set of sentences, T, in a language (such as $\mathcal{L}_\in$). A *proof* of a sentence σ is then a *finite* sequence of formulae: $\varphi_0, \varphi_1, \cdots, \varphi_n = \sigma$ such that for any formula φ_i on the list either: (i) φ_i is an instance of a pure axiom of predicate calculus; or (ii) φ_i is in T; or (iii) φ_i follows from one or more earlier members of the list by an application of a deduction rule.

In which case we shall say that the list is a proof from the set of *axioms* T, and write $T \vdash \sigma$. If $T = \varnothing$ then we shall call this a proof in *first order logic* alone. We shall want to be able to say that it is a mechanical, or algorithmic, process to *check a proof*. Given a finite list which purports to be a proof, it is indeed a mechanical process to check (i) or (iii) for any φ_i on the list. In order to ensure the whole process is algorithmic it is usual (and overwhelmingly the case) that the set of axioms T is either finite itself, or for any formula ψ there is an algorithm or recursive process which can decide whether ψ is in T or not. In which case we say that T is a *recursive set of axioms*.

A.2 Semantics

We have defined so far only *syntactical concepts*. We have not associated any *meaning*, or *interpretation* to our language. We want to know what it means for a sentence to be *true (or false) in an interpretation.* If I wish to express the commutative law in group theory say, then I may write something down such as $\forall x \forall y(\circ(x,y) = \circ(y,x))$ (with a binary function symbol $\circ(v_i, v_j)$). For this to be *true in the group* G say, we need that for any interpretation of the variables x, y as group elements g, h in G that $g.h = h.g$ holds for the group multiplication.

We can give a recursive definition of what it means for a sentence to be 'true in a structure', but as can be seen, the recursion involves at the same time defining *satisfaction* of a formula φ by an assigment of elements to the free variables of φ, again by recursion on the structure of formulae. We'll keep with the example of a group $\langle G, e, \cdot, ^{-1} \rangle$ for a language containing a binary function symbl $F_\circ$, a unary function symbol I and a constant symbol E which are interpreted as $\cdot, ^{-1}$, and e respectively. Then '$I(v_j) = v_k$' and '$F_\circ(v_i, v_j) = v_k$' now count as atomic formulae.

We let $Q_G = {}^{<\omega\,\mathrm{Vbl}}G$ be the set of maps from finite sequences of variables of the language to G to G. For φ a formula let $\mathrm{Vbl}(\varphi)$, be the set of all variables occurring in φ. For $h \in Q_G$, $v_i \in \mathrm{dom}(h)$ and $g \in G$ we let $h(g/i)$ be the function that is defined everywhere like h except that $h(g/i)(v_i) = g$.

Definition A.3 *(i) We define by recursion the term* $\mathrm{Sat}(\varphi, G)$;

$$\mathrm{Sat}(v_i = v_j, G) = \{h \in Q_G | h(i) = h(j)\} \quad ;$$
$$\mathrm{Sat}(I(v_j) = v_k, G) = \{h \in Q_G | h(j)^{-1} = h(k)\}$$
$$\mathrm{Sat}(F_\circ(v_i, v_j) = v_k, G) = \{h \in Q_G | h(i) \cdot h(j) = h(k)\} \quad ;$$
$$\mathrm{Sat}(\chi \vee \psi, G) = (\mathrm{Sat}(\chi, G) \cup \mathrm{Sat}(\psi, G)) \cap \{h \in Q_G \mid \mathrm{dom}(h) \supseteq \{\mathrm{Vbl}(\chi) \cup \mathrm{Vbl}(\psi)\}\};$$
$$\mathrm{Sat}(\neg\psi, G) = Q_G \setminus \mathrm{Sat}(\psi, G)\} \quad ;$$
$$\mathrm{Sat}(\exists v_i \psi, G) = \{h \in Q_G | \mathrm{dom}(h) \supseteq \mathrm{Vbl}(\psi) \cup \{v_i\} \& \exists g \in G(h(g/i) \in \mathrm{Sat}(\psi, G))]\};$$
$$\mathrm{Sat}(u, G) = \varnothing \quad \textit{if } u \textit{ is not a formula.}$$

(ii) We write $\langle G, e, \cdot, ^{-1} \rangle \models \varphi[h]$ *iff* $h \in \mathrm{Sat}(\varphi, G)$.

Note: By design then we have $\langle G, e, \cdot, ^{-1} \rangle \models \neg\psi[h]$ iff it is not the case that $\langle G, e, \cdot, ^{-1} \rangle \models \psi[h]$ etc. (We write the latter as $\langle G, e, \cdot, ^{-1} \rangle \not\models \psi[h]$.)

If φ is a sentence then we write

$$\langle G, e, \cdot, ^{-1} \rangle \models \psi \text{ iff for some } h \in Q_G \text{ with } \mathrm{dom}(h) \supseteq \mathrm{Vbl}(\varphi)\ \langle G, e, \cdot, ^{-1} \rangle \models \psi[h]$$
$$(\text{equivalently for all } h \in Q_G \text{ with } \mathrm{dom}(h) \supseteq \mathrm{Vbl}(\varphi)\ \langle G, e, \cdot, ^{-1} \rangle \models \psi[h]\ ).$$

If T is a set of sentences in a language, and $\mathfrak{A}$ is a structure appropriate for that language, we write $\mathfrak{A} \models T$ *iff for all* ψ *in* T $\mathfrak{A} \models \psi$.

Definition A.4 *(Logical Validity) Let* $T \cup \{\sigma\}$ *be a theory in a language; then* $T \models \sigma$ *if for every structure*

$\mathfrak{A}$ *appropriate for the language,*

$$\mathfrak{A} \models T \Rightarrow \mathfrak{A} \models \sigma.$$

Theorem A.5 *(Gödel Completeness Theorem) Predicate Calculus is* sound*, that is,*

$$T \vdash \sigma \Rightarrow T \models \sigma.$$

and is moreover complete*, that is* $T \models \sigma \Rightarrow T \vdash \sigma$.

The substantial part here is the Completeness direction: it is an *adequacy* result in that it shows that Predicate Calculus is sufficient to deduce from a theory T all those sentences that are true *in all structures* that are models of a particular theory. That it deduces *only those* sentences true in all structures satisfying the theory, is the soundness direction.

This theorem should not be confused with Gödel's *Incompleteness Theorems*. These concerned whether sets of axioms T were *consistent*, that is whether from the axioms of T we cannot prove a contradiction such as '$0 = 1$'. Here if we take T to be PA - *Peano Arithmetic* the accepted set of axioms for the natural number structure $\mathbb{N} = \langle \mathbb{N}, 0, \mathrm{Succ}\rangle$, then Gödel showed that there was a suitable mapping $\varphi \rightarrowtail \ulcorner\varphi\urcorner$ taking formulae in the language appropriate for $\mathbb{N}$, into *code numbers* of these formulae (called *gödel codes*). If formulae could be coded as elements of $\mathbb{N}$, so can a finite list of formulae - in other words potential proofs. Using the fact that the axioms of PA are capable of being recursively listed, he showed that there was a formula defining a function on pairs of numbers, $F(n, k)$, to 1/0 with:

$\mathrm{PA} \vdash \forall n \forall k F(n,k) \in \{0,1\} \wedge$
$F(n,k) = 1 \Leftrightarrow$ *n is a code number of a proof from* PA *of the formula φ with* $\ulcorner\varphi\urcorner = k$.

He then showed that if PA is consistent, then in fact $\mathrm{PA} \nvdash \forall n F(n, \ulcorner 0 = 1\urcorner) = 0$. The right hand side here is a statement about F and numbers, but has the interpretation that "PA is a consistent system (in other words that '$0 = 1$ is not deducible'). This is commonly abbreviated '$\mathrm{Con}(\mathrm{PA})$'; so he showed that even if PA is consistent, $\mathrm{PA} \nvdash \mathrm{Con}(\mathrm{PA})$ (the *Second Incompleteness Theorem*). In fact the theorem has wider applicability as he noted after considering Turing's work on computability: for *any* consistent, computably given set of axioms T say, if from T we can deduce the Peano axioms, then $T \nvdash \mathrm{Con}(T)$. The axioms of set theory ZF, if consistent, are of course such a T.

Exercise A.1 Let x be any set, and $f_i : {}^{n_i}V \longrightarrow V$ for $i < \omega$ be any collection of finitary functions (meaning that $n_i < \omega$); show that there is a $y \supseteq x$ which is closed under each of the f_i (thus f_i“ ${}^{n_i}y \subseteq y$ for each i) and $|y| \leq \max\{\omega, |x|\}$. [Hint: no need for a formal argument here: build up a y in ω many stages $y_k \subseteq y_{k+1}$ at each step applying all the f_i.]

Definition A.6 *Let* $\mathfrak{A} = \langle A, =, \overrightarrow{R_i}, \overrightarrow{F_j}\rangle$ *be any structure for any (first order) language* $\mathcal{L}_{\mathfrak{A}}$. *We write* $\mathfrak{B} \prec \mathfrak{A}$ *("$\mathfrak{B}$ is an* elementary substructure of $\mathfrak{A}$*"), where* $\mathfrak{B} = \langle B, =, \overrightarrow{R_i} \upharpoonright B, \overrightarrow{F_j} \upharpoonright B\rangle$*, to mean that every formula* $\varphi(v_0, \dots v_{n-1})$ *of the language of* $\mathcal{L}_{\mathfrak{A}}$*, and every n-tuple of elements* $y_0, \dots, y_{n-1}$ *from* $\mathfrak{B}$*, then*

$$\mathfrak{A} \models \varphi[y_0/v_0, \dots, y_{n-1}/v_{n-1}] \Leftrightarrow \mathfrak{B} \models \varphi[y_0/v_0, \dots, y_{n-1}/v_{n-1}].$$

The *Tarski-Vaught criterion* yields when one substructure $\mathfrak{B}$ is an elementary substructure of $\mathfrak{A}$.

Lemma A.7 (Tarski-Vaught criterion) $\mathfrak{B} \prec \mathfrak{A}$ *iff for all formulae* $\varphi(v_0, \ldots, v_n)$,

$$\forall b_1, \ldots, b_n \in B(\exists a \in A\, \mathfrak{A} \models \varphi[a, \vec{b}] \to \exists b \in B\, \mathfrak{B} \models \varphi[a, \vec{b}]).$$

Definition A.8 (Skolem Function) *Let* $\exists x \varphi(x, y_0, \ldots, y_n)$ *be any formula in the language* $\mathcal{L}_{\mathfrak{A}}$ *appropriate for the structure* $\mathfrak{A}$. *Suppose there is a wellorder* $\lhd$ *of the domain A. The* skolem function h_φ *for* φ *is the (partial) function:*

$$h_\varphi(y_0, \ldots, y_n) \approx \textit{ the } \lhd\textit{-least } x \textit{ such that } \mathfrak{A} \models \varphi[x, y_0, \ldots, y_n].$$

Notice that there are as many skolem functions as formulae in the language - which will be countable in the cases of interest to us. There are situations where the skolem functions h_φ are already present amongst the functions $\vec{F_j}$ of the structure $\mathfrak{A}$. In particular we may have that that a wellorder $\lhd$ of $\mathfrak{A}$ is itself one of the relations $\vec{R_j}$ of $\mathfrak{A}$. In that case we do not need the skolem functions h_φ to be amongst the $\vec{F_j}$, since we can outright define them, within $\mathfrak{A}$ and not referring to anything external to $\mathfrak{A}$; namely, just use the displayed definition within $\mathfrak{A}$ to pick out the least x.

The following theorem will be used in applications.

Theorem A.9 **Löwenheim-Skolem Theorem** *Let* $\mathfrak{A}$ *be any infinite structure for any language as above of cardinality* ρ. *Suppose* $X \subseteq A$. *Then there is a elementary substructure* $\mathfrak{B}$ *of* $\mathfrak{A}$, $\mathfrak{B} \prec \mathfrak{A}$, *with* $X \subseteq B \subseteq A \wedge |B| = \max\{|X|, \rho\}$.

Proof: The idea is to find the closure of X under the finitary skolem functions h_φ. Let H be the set of such functions. Then $|H|$ we are told is ρ. Let $X_0 = X$, and let

$$X_{n+1} = \bigcup\{h_\varphi``X_n \mid h_\varphi \in H\}; \quad Y = \bigcup_{n<\omega} X_n.$$

The idea is that by closing up in this way we have ensured that the Tarski-Vaught criterion can be applied. However $|X_{n+1}| = \rho \otimes |X_n| = \rho \otimes |X_0|$. Hence $B = \bigcup_n X_n$ satisfies $|B| = \rho \otimes |X| = \max\{\rho, |X|\}$. Now if we take any $y_0, \ldots, y_n \in B$ we shall have that $y_0, \ldots, y_n \in X_m$ for some $m < \omega$. But then if $\mathfrak{A} \models \varphi(z, \vec{y})$ then $\exists x \in X_{m+1} (\mathfrak{B} \models \varphi(x, \vec{y}))$. Q.E.D.

Corollary A.10 *Any infinite structure* $\mathfrak{A}$ *has a* countable *substructure* $\mathfrak{B} \prec \mathfrak{A}$.

A.3 A Generalised Recursion Theorem

Definition A.11 *If* $\langle A, R\rangle$ *is a partial order, we let* $A_x =_{df} \{y \mid y \in A \wedge yRx\}$. *We sometimes write* $A_x = \mathrm{pred}_{\langle A,R\rangle}(x)$ *if we wish to be clear about which order on A is concerned.*

A_x is thus the set of R-predecessors of x that are in A.

Definition A.12 *If $\langle A, R\rangle$ is a wellfounded relation, R is said to be is* set-like on A, *if for every $x \in A$, $A_x =_{df} \mathrm{pred}_{\langle A,R\rangle}(x) =_{df} \{y \mid y \in A \wedge yRx\}$ is a set.*

One can prove a recursion theorem for wellfounded relations, but observe that such relations are not necessarily transitive orderings. We remedy this by defining R^* - the *transitive closure or transitivisation* of R in A, where for $x, y \in A$ we want to put

$$xR^*y \leftrightarrow_{\mathrm{df}} xRy \vee \exists n > 0 \exists z_1 \in A, \ldots, z_n \in A(xRz_1Rz_2 \cdots Rz_nRy).$$

This is a somewhat informal definition, but the intention is clear: xR^*y if there is a finite R-path using elements from A from x to y.

Definition A.13 *$\langle A, R\rangle$ be a relation. For $x \in A$ we define $\bigcup_R x =_{df} \bigcup_{z \in x} A_z$. We let*
$\bigcup_R^0 x = A_x; \bigcup_R^{n+1} x = \bigcup \left\{A_z \mid z \in \bigcup_R^n x\right\}$.
For $x, y \in A$ we set yR^*x iff $y \in \bigcup\{\bigcup_R^n x | n \in \mathbb{N}\}$. R^* is called the ancestral *or* transitive closure *of R.*

The reader should check that a) $\bigcup_R^{n+1} x = \{y \in A \mid \exists z_0 \in A \cdots \exists z_n \in A(yRz_0Rz_1 \cdots Rz_nRx)\}$, and b) with R as the $\in$-relation itself $y \in^* x \leftrightarrow y \in \mathrm{TC}(x)$.

Lemma A.14 *Let $\langle A, R\rangle$ be a relation. Then:*
(i) R^* is transitive on A. If $x \in A$ then R^* is transitive on $A_x \cup \{x\}$.
(ii) If R is set-like, then so is R^*.

Proof: (i) is obvious. (ii) We first note that $\bigcup_R z$ is a set; this is because z is a set and R is set-like on A which implies that A_x is a set for each $x \in z$, and the Axiom of Unions allows us to conclude that $\bigcup_{x \in z} A_x \in V$. Hence by induction, so is each $\bigcup_R^{n+1} z$, and then another application of Replacement and Union ensures that $\bigcup\{\bigcup_R^n\{x\}|n \in \mathbb{N}\} \in V$; but this latter set is then the set of R^* predecessors of x. Q.E.D.

Theorem A.15 *(Transfinite Induction on Wellfounded Relations). Suppose $\langle A, R\rangle$ is a wellfounded relation, with R set-like on A. Let $t \subseteq A$ be non-empty class term. Then there is $u \in t$ which is R-minimal amongst all elements of t.*

Proof: Let $A_x^* =_{df} \mathrm{pred}_{\langle A,R^*\rangle}(x)$ be the set of R^*-predecessors of x. Note this is a set and is a subset of A. Let x be any element of t, and let u be an R-minimal member of the set $(t \cap A_x^*) \cap \{x\}$. Q.E.D.

One should note that we do need to prove the above theorem, since the definition of $\langle A, R\rangle$ being wellfounded (Def. 1.14) entails only that every non-empty *set* $z \subseteq A$ has an R-minimal element. The theorem then says that this holds for classes t too.

Theorem A.16 ***(Generalized Transfinite Recursion Theorem)***
Suppose $\langle A, R\rangle$ is a wellfounded relation, with R set-like on A. If $G : V \times V \to V$ then there is a unique function $F : A \to V$ satisfying:

$$\forall x F(x) = G(x, F \upharpoonright A_x).$$

Proof: We shall define G as a union of *approximations* where $u \in V$ is an approximation if (a) Fun(u); (b) $\mathrm{dom}(u) \subseteq A$ is R-transitive - meaning $y \in \mathrm{dom}(u) \to A^*_y \subseteq \mathrm{dom}(u)$; and (c) $\forall y \in \mathrm{dom}(u) u(y) = G(y, u \restriction A_y)$. We call an approximation u an *x-approximation* if $x \in \mathrm{dom}(u)$. So u satisfies the defining clauses for F throughout its domain. Notice that if u is an x-approximation, then v is also an x approximation, where $v = u \restriction \{x\} \cup A^*_x$. (It is the smallest part of u which is still an x-approximation.)

(1) *If u and v are approximations, and we set $t = \mathrm{dom}(u) \cap \mathrm{dom}(v)$ then $u \restriction t = v \restriction t$ and is an approximation.*

Proof: Note that for any $y \in t$, $A^*_y \subseteq t$ so t is R-transitive. Let $Z = \{y \in t \mid u(y) \neq v(y)\}$.

If $Z \neq \varnothing$ let w be an R-minimal element of Z (by the wellfoundedness of R). Then $u \restriction A_w = v \restriction A_w$, hence:

$$u(w) = G(w, u \restriction A_w) = G(w, v \restriction A_w) = v(w).$$

This contradicts the choice of w. So $Z = \varnothing$ and u, v agree on t, the common part of their domains. This finishes (1). Exactly the same argument establishes:

(2) *(Uniqueness)* If F, F_0 are two functions satisfying the theorem then $F = F_0$.

(3) *(Existence) Such an F exists.*

Proof: Let $u \in B \Leftrightarrow \{u \mid u$ is an approximation$\}$. B is in general a proper class of approximations, but this does not matter as long as we are careful. As any two such approximations agree on the common part of their domain, we may define $F = \bigcup B$ and obtain:

(i) F is a function ;

(ii) $\mathrm{dom}(F) = A$.

Proof (ii): Let C be the class of sets $z \in A$ for which there is no z-approximation. So if we suppose for a contradiction that C is non-empty, by Theorem A.15, then it will have an R-minimal element z such that $\forall y \in A_z \exists u(u$ is a y-approximation). But now we let f be the function:

$$\bigcup\{f^y \mid y \in A_z \wedge f^y \text{ is a } y\text{-approximation } \wedge dom(f^y) = \{y\} \cup A^*_y\}.$$

By (1) for a given y such an f^y is unique, and moreover the f^y all agree on the parts of their domains they have in common. Note that the domain of f is R-transitive, being the union of R-transitive sets $\mathrm{dom}(f^y)$ for $y \in z$. Hence $A^*_z \subseteq \mathrm{dom}(f)$ and thus $\{z\} \cup \mathrm{dom}(f)$ is also R-transitive. We can extend f to

$$f^z = f \cup \{\langle z, G(f \restriction A_z)\rangle\}$$

and f^z is then a z-approximation. However we assumed that $z \in C$, contradiction! Hence $C = \varnothing$ and (ii) holds. Q.E.D.

For some applications it is useful to note that the AxPower was not used in the proof of this theorem, and it can be proved in ZF^-. For $\langle A, R\rangle$ a wellfounded relation, we can define a *rank function* $\rho_{\langle A,R\rangle} : A \to \mathrm{On}$ by appealing to the last theorem: $\rho_{\langle A,R\rangle}(x) = \sup\{\rho_{\langle A,R\rangle}(y) + 1 \mid y \in A \wedge yRx\}$. Clearly this satisfies $xRy \to \rho_{\langle A,R\rangle}(x) < \rho_{\langle A,R\rangle}(y)$, and $\rho_{\langle A,R\rangle}(x)$ is onto On if A is a proper class, or an initial segment of On, *i.e.* an ordinal, if $A \in V$.

Exercise A.2 If $\langle A, R\rangle$ a wellfounded set-like relation, $x \in A$, and $\rho_{\langle A,R\rangle}(x) = \alpha$, show that $\forall\beta < \alpha\exists y(y \in A \wedge yR^*x \wedge \rho_{\langle A,R\rangle}(y) = \beta)$.

Exercise A.3 If $\langle A, R\rangle$ a wellfounded set-like relation, show that $\rho_{\langle A,R\rangle}$ is (1-1) if and only if R^* is a total order.

Exercise A.4 (i) If $\langle A, R\rangle$ a wellfounded set-like relation, and $B \subseteq A$, show that $\rho_{\langle B,R\rangle}(x) \leq \rho_{\langle A,R\rangle}(x)$ for any $x \in B$. Show that additionally equality holds if $A^*_x \subseteq B$ where A^*_x is as in the proof of Theorem A.15 above.

(ii) If $\langle A, R\rangle$,$\langle A, S\rangle$ are wellfounded set-like relations, and $S \subseteq R$, show that $\rho_{\langle A,S\rangle}(x) \leq \rho_{\langle A,R\rangle}(x)$ for any $x \in A$.

Bibliography

[1] K. Devlin. *Constructibility*. Perspectives in Mathematical Logic. Springer Verlag, Berlin, Heidelberg, 1984. 37, 72

[2] F.R. Drake. *Set Theory: An introduction to large cardinals*, volume 76 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, 1974. 37

[3] T. Jech. *Set Theory*. Springer Monographs in Mathematics. Springer, Berlin, New York, 3 edition, 2003. 37

[4] K.Kunen. *Set Theory: an introduction to independence proofs*. Studies in Logic. North-Holland, Amsterdam, 1994. 3, 76

Index of Symbols

Index