

Logic

Aidan Backus

December 2019

Contents

1	Elementary recursion theory	2
1.1	Examples of the diaagonal argument	2
1.2	A model for computation	3
1.3	The halting problem	5
1.4	Primitive recursion	6
2	Undecidability in first-order logic	8
2.1	Undecidability of the theory of algorithms	8
2.2	Undecidability of true arithmetic	10
2.3	Goedel's completeness theorem	14
2.4	Motivating Peano arithmetic	16
2.5	Peano arithmetic	16
2.6	Goedel codes in Peano arithmetic	20
2.7	Goedel's first incompleteness theorem	21
2.8	Goedel's second incompleteness theorem	23
3	Kleene's recursion theorem	25
3.1	Kleene's recursion theorem	25
3.2	Rice's theorem	26
4	Semantic properties of arithmetic	28
4.1	Undecidability in Ramsey theory	28
4.2	Overspill	28
4.3	Standard systems	29
4.4	The arithmetical hierarchy	30
4.5	Turing degrees	31
4.6	The arithmetical hierarchy in PA	32
4.7	Submodels of arithmetic	33
4.8	Slaman's theorem	36
5	The Turing degrees	39
5.1	Myhill's theorem	39
5.2	Completeness	41
5.3	Creative sets	42
5.4	Forcing for sets below $0'$	43

5.5	Incomparable reals	45
5.6	Inverting the jump	46
5.7	The axiom of determinacy	47
5.8	The Friedberg-Muchnik theorem and the priority argument	50
6	Elementary set theory	53
6.1	The axioms of ZF without Foundation	53
6.2	Transfinite induction	54
6.3	Ordinals	55
6.4	The axiom of choice	56
6.5	Cardinals	57
7	Constructibility	60
7.1	Definability	60
7.2	The reflection principle	61
7.3	The constructible sets	62
8	Forcing	64
8.1	Generic filters	64
8.2	Constructing the generic extension	65
8.3	Forcing semantics	67
8.4	Cardinal collapse	68
8.5	Forcing $2^{\aleph_0} = \aleph_2$	69
9	Small large cardinals	71
9.1	Regular cardinals	71
9.2	Inaccessible cardinals	72
9.3	Clubs	74
9.4	Ramsey theory	77
9.5	Weakly compact cardinals	80
9.6	The weak compactness theorem	83
10	Measurable cardinals	85
10.1	Properties of measurable cardinals	85
10.2	Combinatorics of ultrafilters	86
10.3	Ultrapowers by measurable cardinals	90
10.4	Measurable cardinals are very large	93
11	One measurable cardinal	95
11.1	Generalizing L	95
11.2	Solovay's theorem and condensation	97
11.3	Iterated ultrapowers	100
11.4	Suitable initial data for iteration	105
11.5	Mice	106
11.6	Applications of mice	108

12 Extender models	110
12.1 The diagram chase	110
12.2 The abstract definition	113
12.3 Reinhardt cardinals	117
12.4 Supercompact cardinals	118
13 The Ultimate L	123
13.1 Averting the dichotomy	124
13.2 Magidor's weak extender model formulation	126
13.3 Woodin's universality theorem	128
13.4 The HOD dichotomy	129
13.5 V is Ultimate L	134

Chapter 1

Elementary recursion theory

Let V be the set of validities; i.e. first-order formulae ψ such that for every structure M , $M \models \psi$.

Theorem 1.1 (completeness). For every first-order formula ψ , $\psi \in V$ iff $\vdash \psi$.

Corollary 1.2. There is an algorithm which enumerates V .

Proof. Iterate over all strings in the language. For each such string check if it is a proof; if it is, yield its conclusion. \square

So it is possible to check that a formula is in V ; but the above procedure does not allow us to check that a formula is not in V .

Theorem 1.3 (Church-Turing). There is no algorithm to determine membership in V .

The proof is easy once we have a rigorous definition of “algorithm.”

Corollary 1.4. There is no complete finite axiomatization of the theory of algorithms.

Proof. Suppose there were finitely many axioms from which we could deduce the theory of algorithms. Then we could decide membership in V , by considering the algorithm for V . \square

Note that we prove that mathematical statements are undecidable in the same way every time.

Example 1.5. Suppose that P is an NP-complete problem. Then the satisfaction relation embeds in P (in fact, this is how we always prove that P is NP-complete!) So there is really only one reason that a problem is “hard” in complexity theory.

1.1 Examples of the diaagonal argument

Example 1.6. For every real x , there is a number which is not algebraic in x . Enumerate the algebraic numbers in x as q_i and let r disagree with q_i at the i th digit. In fact, when Liouville proved that there was a transcendental number, his proof used a difficult Diophantine approximation argument that does not generalize to the case of numbers which are actually algebraic in x .

Example 1.7 (halting problem). There is a set which is not computable. Enumerate the programs as P_i . Let X be the set of i such that $P_i(i)$ returns False. Thus the halting problem is not decidable.

Example 1.8 (Tarski's undefinability theorem). There is an undefinable set. Enumerate the first-order formulae with one free variables as ψ_i and let X be the set of i such that $\mathbb{N} \models \neg\psi_i(i)$. Thus the satisfaction predicate is not definable.

1.2 A model for computation

We now make the above diagonal arguments rigorous.

Let A be an alphabet, so words in A are elements of $A^{<\omega}$, which we will also write as A^* . By a *procedure* in A we mean an algorithm which takes words as inputs and outputs.

Definition 1.9. A *decision procedure* for $W \subseteq A^*$ is a procedure which takes a word w as input and returns the empty word \square if $w \in W$ and a nonempty word if $w \notin W$.

If A is a finite and W has a decision procedure, then so does W^c . Similarly for finite unions and intersections, and set differences. (So the set of sets which admit a decision procedure forms a boolean algebra.)

Definition 1.10. An *enumeration procedure* is a procedure which on \square input returns a sequence of words.

So a set W is enumerable iff there is an enumeration procedure whose image is W , and W has a decision procedure iff W and W^c are enumerable.

Definition 1.11. Any $f : A^* \rightarrow A^*$ is a *computable function* iff there is a procedure which takes $w \in A^*$ and returns $f(w)$.

So a function is computable iff its graph is enumerable iff its graph has a decision procedure.

We now rigorously define “procedure” using the notion of a register machine. Let $A = \{a_1, \dots, a_r\}$. A *register machine* has finitely many *registers* R_1, \dots, R_N , each of which can hold a word in A^* .

Definition 1.12. A *register machine program* is a finite list of instructions of the following forms:

1. Let $R_i = R_i + a_j$.
2. If the last letter of R_i is a_j then let $R_i = R_i - a_j$.
3. If $R_i = \square$ goto line L . Elif the last letter of R_i is a_j goto line L_j .
4. Print R_1 .
5. Halt.

Here $+$ denotes concatenation. We will assume that if $\alpha_1, \dots, \alpha_k$ are the instructions, then α_j says “Halt” if and only if $j = k$.

Note that if we only cared about computable sets rather than enumerable sets we would not need “Print”. This is analogous to the “yield” command in Python and lazy returns in Lisp.

Example 1.13. Let $A = \{a\}$. Consider the register machine program:

1. If $R_2 = \square$ then goto 5, else goto 2
2. Let $R_1 = R_1 + a$
3. If the last letter of R_2 is a then let $R_2 = R_2 - a$
4. If $R_2 = \square$ then goto 5, else goto 2
5. Print R_1
6. Halt

This program takes R_1 and R_2 as input, and replaces R_1 with $R_1 + R_2$ and R_2 with \square . So this program witnesses that addition is computable.

If P is a register machine program, then we write

$$P : w \mapsto \eta$$

if P halts on input w and prints η . We write $P : w \mapsto \infty$ if P does not halt.

Axiom 1.14 (Church-Turing thesis). Assume that A is finite. If there is a procedure to decide $W \subseteq A^*$, then there is a register machine program which also decides W .

So without loss of generality we may assume that any program is a register machine program. Note that the Church-Turing thesis only applies to idealized programs. For example, if Planck’s constant is uncomputable, one could in principle use the digits of Planck’s constant in some machine, but it would not be necessarily computable using a register machine program.

One might also hope to extend the Church-Turing thesis to nondeterministic algorithms. However, it is a theorem that if X is a measurable subset of $[0, 1]$ whose measure is positive, then any function that can be computed by a program which can take elements of $[0, 1]$ as parameters and returns the same as long as those elements lie in X is actually computable in the deterministic sense.

Definition 1.15. A set W is *register-decidable* if there is a register machine program which decides W .

Example 1.16. Let $A = \{a\}$. Consider the register machine program p :

1. If $R_1 = \square$ then goto 2 else goto 3

2. Let $R_2 = R_2 + a$
3. If the last letter of R_1 is a then let $R_1 = R_1 - a$
4. If $R_1 = \square$ then goto 5 else goto 3
5. If $R_2 = \square$ then goto 7 else goto 6
6. Let $R_1 = R_1 + a$
7. Print
8. Halt

Suppose that $R_0 = w$ and $R_1 = \square$ initially. Then this program outputs a if $w = \square$ and \square otherwise. Therefore the set of nonempty words, $\{\square\}^c$, is register-decidable. In particular it is decidable.

1.3 The halting problem

Let P be a register machine program. Then P is a finite sequence of instructions, each of which consists of finitely many symbols. Therefore P can be viewed a finite word in a finite alphabet A .

Let **Halt** be the set of register machine programs P such that P halts on input P .

Theorem 1.17. **Halt** is not register-decidable by any machine in the alphabet A .

Proof. Suppose that **Halt** is register-decidable by some program P in the alphabet A . Then P has a halt instruction, say on line k . Define a program Q by taking the code for P and replacing line k with the instruction “goto k ”, and append line $k+1$ with a halt instruction. In addition, for all lines ℓ in P with a print instruction, replace ℓ in Q with “if $R_1 = \square$ then goto $\ell+1$ else goto $k+1$.”

Let σ be a string. If $P : \sigma \mapsto \eta$ and $\eta \neq \square$, then the change to lines ℓ and $k+1$ causes Q to halt. That is, if P thinks that σ does not halt, then Q halts on input σ . Conversely, if $P : \sigma \mapsto \square$, then Q does not print anything due to the change to line ℓ , and then loops forever due to the change to line k . So if P thinks that σ halts, then Q does not halt on input σ .

We call Q on input Q . If $P : Q \mapsto \eta$ and $\eta \neq \square$, then Q halts when called with input Q , so $P : Q \mapsto \square$, a contradiction. Otherwise, if $P : Q \mapsto \square$, then Q does not halt when called with input Q , so $P : Q \mapsto \eta$ for some $\eta \neq \square$, which is also a contradiction. \square

Let $A \subset B$ be alphabets. Clearly we cannot solve the halting problem in A with a program in the alphabet A , but one could consider choosing B so big that a program in B can decide the halting problem in A . However, this turns out to be impossible, and we can always collapse A to an alphabet of just one letter without affecting the halting problem.

Exercise 1.18. Show that if P is a program in any nonempty alphabet A and $a \in A$ is a symbol, there is a program Q in the alphabet $\{a\}$, such that if $P : Q \mapsto \square$ then Q does not halt, and if $P : Q \mapsto \eta$ for some $\eta \neq \square$ then Q halts. (Hint: Show that $\{a\}$ can be used to describe an arbitrarily large alphabet using Gödel codes.)

Corollary 1.19. `Halt` is not register-decidable by any register machine written in any alphabet.

Corollary 1.20. There is a definable set which is not register-decidable.

1.4 Primitive recursion

We now introduce a stronger notion of recursive function, those which are especially easy to analyze because we have a priori control over their runtime.

Definition 1.21. A loop in a computer program is said to be a *bounded loop* if it can be expressed as a `for` loop, such that the `for` loop indexes over is fixed at the start of the loop and cannot be edited from within the loop.

Definition 1.22. A *primitive recursive function* is a partial recursive function f such that there is a program which computes f , for which we are only allowed the operations $+$, $-$, $=$, $<$, `if`, and bounded loops. The set of primitive recursive functions is denoted **PR**.

Theorem 1.23. Every primitive recursive function is recursive.

Proof. Let f be primitive recursive and suppose that P witnesses that f is primitive recursive. We fix a notion of runtime where the primitive operations $+$, $-$, $=$, $<$, and `if` take one unit of time. Those blocks of code in P which are s lines long and have no bounded loops have runtime at most $s < \infty$. Those blocks of code in P which have s lines of code and N bounded loops, each with at most n_k iterations and internal runtime at most $s_k < \infty$, have runtime at most $s + \sum_{k=1}^N n_k s_k < \infty$. By Noetherian induction, the outermost code block in P has finite runtime. Therefore $f(n) \downarrow$. \square

We now give a primitive recursive function that is not recursive by a diagonal argument. So it is not so easy to use Noetherian induction to bound the runtime of a program.

Definition 1.24. The *Ackermann function* $A : \mathbb{N}^2 \rightarrow \mathbb{N}$ is defined recursively by

$$A(m, n) = \begin{cases} n + 1, & m = 0 \\ A(m - 1, 1), & m > 0, n = 0 \\ A(m - 1, A(m, n - 1)) & (m, n) \neq 0. \end{cases}$$

Example 1.25. Let $\alpha(n)$ denote the least m such that $A(n, n) \leq m$, the *inverse Ackermann function*. Then $\alpha(n) \leq 5$ for any n which would “normally appear” in applied complexity theory (in fact for any n less than the number of particles in the universe). This is because $A(4, 3) = 2^{2^{65536}} - 3$. In fact, the disjoint-set data structure, implemented with path compression and union-by-rank, has amortized runtime $O(\alpha(n))$, hence $O(1)$ for any physically existing machine.

Theorem 1.26. The Ackermann function is not primitive recursive.

Proof. Let \mathcal{A} be the class of all multivariable functions $f : \mathbb{N}^n \rightarrow \mathbb{N}$ such that there is a $t \in \mathbb{N}$, such that for every $x \in \mathbb{N}^n$,

$$f(x_1, \dots, x_n) < A(t, \max_i x_i).$$

Clearly \mathcal{A} contains the constant functions, and is closed under the successor function and projections. If $g : \mathbb{N}^k \rightarrow \mathbb{N}^m$ and $h : \mathbb{N}^m \rightarrow \mathbb{N}$ have $g, h \in \mathcal{A}$, then $g_i(x) < A(r_i, \max x)$, $h(y) < A(s, \max y)$ for some r, s . If $f = h \circ g$ and $g_j(x) = \max_j g(x)$ then

$$f(x) < A(s, g_j(x)) < A(s, A(r_j, x)) < A(s + r_j + 2, x)$$

so setting $t = s + r_j + 2$ we see that $f \in \mathcal{A}$. Therefore \mathcal{A} is closed under function composition.

Now suppose that $g : \mathbb{N}^k \rightarrow \mathbb{N}$, $h : \mathbb{N}^{k+2} \rightarrow \mathbb{N}$, $g, h \in \mathcal{A}$. Suppose that $g(x) < A(r, \max x)$ and $h(y) < A(s, \max y)$. Suppose that $f : \mathbb{N}^k \times \mathbb{N} \rightarrow \mathbb{N}$ is recoverable from g, h using a primitive recursive program. We claim $f \in \mathcal{A}$.

Let $q = 1 + \max(r, s)$ and induct on n to see that

$$f(x, n) < A(q, n + x).$$

If $n = 0$ then $f(x, 0) < A(q, x)$. So

$$f(x, n + 1) = h(x, n, f(x, n)) < A(s, z)$$

where $z = \max(\max x, n, f(x, n))$. Then $z < A(q, n + x)$ by the inductive hypothesis. Thus

$$f(x, n + 1) < A(s, z) < A(s, A(q, n + x)) \leq A(q, n + 1 + x)$$

so $f \in \mathcal{A}$.

Therefore A strictly dominates **PR**, so $A \notin \mathbf{PR}$. \square

The Ackermann function is used to test the ability of compilers to optimize deep recursion, which is difficult because of the limited height of the stack and the inability to tail-recuse, among other problems.

Chapter 2

Undecidability in first-order logic

2.1 Undecidability of the theory of algorithms

Definition 2.1. A *validity* is a sentence φ in a language \mathcal{L} such that for every \mathcal{L} -structure M , $M \models \varphi$.

In other words, φ is a validity iff $\Box\varphi$ (here \Box is the modal operator, not the empty string). For example, “ $\forall x x = x$ ” is a validity. Hilbert asked for an algorithm to determine whether a sentence is a validity. This wasn’t completely unreasonable, because zeroth-order logic has an obvious algorithm to determine validity (namely truth table calculus), though this is NP-complete. However, it was unreasonable.

Theorem 2.2 (Church-Turing). The set of validities is not register-decidable.

We set up the proof. Let us assume that programs are given by register-machine programs in the alphabet $\{a\}$, words in which we identify with natural numbers: (a, a, \dots, a) with k entries is identified with k .

Definition 2.3. Let P be a program with n registers. A *configuration* for P is a vector

$$(S, L, m_1, \dots, m_n) \in \omega^{n+3}$$

such that S denotes the time step that we are on (so $S \mapsto S + 1$ every time we carry out an instruction), L is the label of the instruction to be done next, and m_j is the value of the register R_j at the current time stage. The *initial configuration* with input k is the configuration $(0, 0, k, 0, \dots, 0)$.

The set X of configurations which appear in the evaluation of P with input k clearly satisfies a recursion condition. In fact, the initial configuration is in X . Moreover, for every S , if $(S, L, m_1, \dots, m_n) \in X$ and the instruction at label L is not “Halt”, then $(S + 1, L^*, m_1^*, \dots, m_n^*) \in X$ where L^*, m_1^*, \dots, m_n^* are the values that result at the next time stage.

Definition 2.4. We will say that a set of configurations X is *closed under P* if for every $(S, L, m_1, \dots, m_n) \in X$ and the instruction at label L is not “Halt”, then the vector $(s +$

$1, L^*, m_1^*, \dots, m_n^*)$ defined by the instruction in P at label L with registers m_1, \dots, m_n , where L^* is the label of the next instructions and m_1^*, \dots, m_n^* are the new values of the registers, is also in X .

So a set which is closed under X is a generalization of the set of configurations that appear in the evaluation of P with input k . This gives us the following useful fact:

Lemma 2.5. The program P halts on input k if and only if for every set of configurations $X \subseteq \omega^{k+3}$ containing the initial configuration $(0, 0, k, 0, \dots, 0)$ which is closed under P , there is a configuration in X at label L , such that P has a “Halt” instruction at label L .

We now construct a first-order language \mathcal{L} to discuss algorithms. Let c be a constant symbol, f a unary function symbol, and R an $(n+3)$ -ary relation symbol. We will interpret \mathcal{L} on ω , interpreting c as 0, f the successor function, and R as the set of configurations which appear in the evaluation of a program P with input k , viewed as a subset of ω^{n+3} .

We now let φ_0 be the sentence “ f is injective and c is not in the image of f ”, i.e.

$$\varphi_0 : \forall x_1 \forall x_2 (f(x_1) = f(x_2) \leftrightarrow x_1 = x_2) \wedge \forall x_1 (f(x_1) \neq c).$$

In particular, a model of φ_0 contains a copy of ω , namely the set generated by the interpretations of c and f . Therefore, for any $m \in \omega$, we may let \bar{m} denote the symbol $f(f(\dots f(c)))$ (m copies of f), which corresponds to m in the copy of ω . Given an input string k , let $\varphi_{init}(\alpha)$ be the formula “ α is the initial configuration vector for k ”, i.e. if $\alpha = (\bar{0}, \bar{0}, \bar{k}, \bar{0}, \dots, \bar{0})$ then

$$\varphi_{init}(\alpha) : R(\bar{0}, \bar{0}, \bar{k}, \bar{0}, \dots, \bar{0}).$$

Let φ_α be the formula “if a configuration α is in R then so is the instruction obtained from α ”; i.e.

1. If α is at label L and says “Let $R_i = R_i + a$ ” then

$$\varphi_\alpha : \forall x \forall y_1 \dots \forall y_n R(x, \bar{L}, y) \rightarrow R(f(x), \bar{L+1}, y_1, \dots, y_{i-1}, f(y_i), y_{i+1}, \dots, y_n).$$

2. If α is at label L and says “Let $R_i = R_i - a$ ” then

$$\begin{aligned} \varphi_\alpha : \forall x \forall y_1 \dots \forall y_n R(x, \bar{L}, y) \rightarrow & ((y_i = \bar{0}) \wedge (R(f(x), \bar{L+1}, y))) \\ & \vee ((y_i \neq \bar{0}) \wedge (\exists z (y = f(z)) \wedge R(f(x), \bar{L+1}, y_1, \dots, y_{i-1}, z, y_{i+1}, \dots, y_n))) \end{aligned}$$

3. If α is at label L and says “If $R_i = \square$ then goto L' else goto L_0 ” then

$$\begin{aligned} \varphi_\alpha : \forall x \forall y_1 \dots \forall y_n R(x, \bar{L}, y) \rightarrow & ((y_i = \bar{0}) \wedge R(f(x), \bar{L'}, y)) \\ & \vee (y_i \neq \bar{0}) \wedge (R(f(x), \bar{L_0}, y)). \end{aligned}$$

4. If α is at label L and says “Print” then

$$\varphi_\alpha : \forall x \forall y_1 \dots \forall y_n R(x, \bar{L}, y) \rightarrow R(x, \bar{L+1}, y).$$

5. If α is at label L and says “Halt” then φ_α is not defined.

Finally, we let φ_{halt} mean “ P halts”, i.e. if L is the label of the halting instruction in P then

$$\varphi_{halt} : \exists x \exists y_1 \cdots \exists y_n R(x, \overline{L_{halt}}, y).$$

Proof of the Church-Turing theorem. For every program P and input string k , we define a sentence $\varphi_{P,k}$ such that $(P, k) \mapsto \varphi_{P,k}$ is a computable function and P halts on input k if and only if $\varphi_{P,k}$ is a validity. Then deciding validity in $\{\varphi_{P,k}\}$ ranging over all programs P is equivalent to deciding **Halt**, which is impossible.

In fact we let

$$\varphi_{P,k} : \left(\varphi_0 \wedge \varphi_{init} \wedge \bigwedge_{\alpha} \varphi_{\alpha} \right) \rightarrow \varphi_{halt}.$$

where the conjunction ranges over all α such that P contains α as an instruction and α is not the halting instruction. Clearly $\varphi_{P,k}$ is computable from (P, k) ; we must show that P halts on input k iff $\varphi_{P,k}$ is valid.

Suppose that P diverges on k . Define a structure M on ω by giving c and f their correct interpretations, and letting $R \subseteq \omega^{n+3}$ be the set of $(s, t, m) \in \omega \times \omega \times \omega^n$ such that (s, t, m) is a configuration in the evaluation of P on k . Then M clearly satisfies the antecedent of $\varphi_{P,k}$ but violates its consequent, so $M \models \neg \varphi_{P,k}$. Therefore $\varphi_{P,k}$ is invalid.

Conversely, if P halts on k , then suppose that $M = (M, c^M, f^M, R^M)$ is a model of the antecedent of $\varphi_{P,k}$. We will prove that M is a model of the consequent, so $\varphi_{P,k}$ is valid. In fact, we induct to show that if (s, t, m) is a configuration in the evaluation of P on k , then

$$M \models R(\overline{s}, \overline{t}, \overline{m}).$$

Since $M \models \varphi_{init}$, this is true when $s = t = 0$ and $m = (k, 0, \dots, 0)$. Now suppose that (s, t, m) appears and $M \models R(\overline{s}, \overline{t}, \overline{m})$. Let α be the instruction at label t and let t^* and m^* be the result of the computation at time s . Since $M \models \varphi_\alpha$ it follows that $M \models R(\overline{s+1}, \overline{t^*}, \overline{m^*})$. Since P eventually halts, it follows by induction that $M \models \varphi_{halt}$. \square

Notice that a key property of the theory of algorithms is that for every model M (i.e. set of configurations), there is an embedding of ω in M defined by $n \mapsto f^n(c)$. Since $M \models \varphi_0$ this mapping is actually injective. We can think of M as a (possibly nonstandard) model of arithmetic (since there may be $k \in M$ which was not obtained from n by applying the successor operation f). However, the nonstandard part of M is irrelevant to the validity of $\varphi_{P,k}$.

2.2 Undecidability of true arithmetic

We now prove something stronger than the Church-Turing theorem, namely that the theory of arithmetic is undecidable.

Theorem 2.6. The theory T of ω with the language $(0, 1, +, \times)$ is not decidable.

By definition, $\varphi \in T$ iff $\omega \models \varphi$. We sometimes call T *true arithmetic*. In fact, T is not even recursively axiomatizable.

To show that T is not decidable using the above argument, we need a way to refer to configuration vectors using just natural numbers. Therefore we introduce a Gödel coding. This will be challenging because we cannot refer directly to exponentiation in the language $(0, 1, +, \times)$. In fact the function $(p, i) \mapsto p^i$ is not definable in ω .

Lemma 2.7 (β -lemma). There is a function $\beta : \omega^3 \rightarrow \omega$ such that:

1. For every sequence $(a_0, \dots, a_r) \in \omega^{r+1}$ there exist $t, p \in \omega$ such that for every $i \leq r$, $\beta(t, p, i) = a_i$.
2. β is definable in ω .

In other words, if we view (t, p) as coding the vector a , then $\beta(\cdot, \cdot, i)$ is the projection of a onto the i th factor. Obviously many such β exist, but it is important that β can be easily described. Definability in ω means that there is a formula φ in the language $(0, 1, +, \times)$ with four free variables such that for all t, p, i, a , $\omega \models \varphi(t, p, i, a)$ if and only if $\beta(t, p, i) = a_i$.

The idea of the proof of the β -lemma is to use the p -adic expansion of t . This is not unique, but that is OK.

Before the proof, we note that $<$ is definable in $(0, 1, +, \times)$ by noting that $a < b$ iff

$$\exists c((c \neq 0) \wedge (a + c = b)).$$

Therefore we use $<$ freely.

Proof. Given a_0, \dots, a_r , let p be a prime such that $p > r + 1$ and for each i , $p > a_i$, which is possible by Euclid's theorem on primes. Let

$$t = 1 + a_0p + 2p^2 + a_1p^3 + 3p^4 + a_2p^5 + \dots + (r+1)p^{2r} + a_rp^{2r+1}.$$

Then the coefficient of p^{2k} indicates the number $k + 1$ and the coefficient of p^{2k+1} indicates a_k . In particular, the p -adic expansion of t is unique.

Lemma 2.8. For all a and $i \leq r$, $a = a_i$ if and only if there exist b_0, b_1, b_2 such that:

1. $t = b_0 + b_1((i+1) + ap + b_2p^2)$.
2. $a < p$.
3. $b_0 < b_1$.
4. There exists m such that $b_1 = p^{2m}$.

Proof of sublemma. Suppose $a = a_i$. Then the claim immediately follows from the definition of t, p : b_0 is given by lower powers of p , and then

$$t = b_0 + p^{2i}((i+1) + a_ip + p^2((i+2) + a_{i+1}p + \dots + a_rp^{2r+1-(2i+2)})).$$

Then $b_1 = p^{2i}$ and b_2 consists of the higher powers of p , divided by p^{2i} , namely $b_2 = (i+2) + a_{i+1}p + \dots$. Since $p > a_i$ it follows that $a < p$. The $a_i < p$ and $p > r+1$, so $b_0 < p^{2i}$.

Conversely, suppose that b_0, b_1, b_2 are given and $b_1 = p^{2m}$. Then by the assumptions, b_0 is the remainder of t by dividing by p^{2m} . Since $i + 1 < p$, $i + 1$ is the remainder of dividing $(t - b_0)p^{-2m}$ by p . So $i + 1$ is the $2m$ th coefficient of the p -adic expansion of t . By definition of t , $m = i$. Since $a < p$, a is the $2i + 1$ th coefficient in the p -adic expansion of t . Therefore $a = a_i$. \square

The fourth claim in the sublemma is not obviously definable in ω . But it is definable by the formula

$$\exists x((x \times x = b_1) \wedge P(p, x))$$

where $P(q, y)$ means that q is the unique prime divisor of y . Therefore the sublemma implies that the statement $a = a_i$ can be expressed in $(0, 1, +, \times)$. In fact, if we choose p to be minimal possible, then $(a_0, \dots, a_r) \mapsto (p, t)$ is a well-defined, computable function.

Now let $\beta(u, q, j)$ by the smallest $a \in \omega$ such that there are b_0, b_1, b_2 such that:

1. $u = b_0 + b_1((j + 1) + aq + b_2q^2)$.
2. $a < q$.
3. $b_0 < b_1$.
4. There is an m such that $b_1 = q^{2m}$.
5. q is prime.

If no such b_0, b_1, b_2 exist, then let $\beta(u, q, j) = 0$.

Since $<$ is definable, so is “the smallest ...”, so the sublemma implies that β is definable. \square

Note that if p is composite then $\beta(\cdot, p, \cdot) = 0$.

As we will show, the coding of finite sequences (by any computable means, but in particular by the β -lemma) is sufficient to emulate a register machine within ω . In particular, the theory of arithmetic can decide the halting problem. Trouble comes when we realize that we need to implement recursion in $(0, 1, +, \times)$. In fact, the exponential map $(x, y) \mapsto x^y$ is computable, but cannot be defined in the obvious way because it grows faster than any polynomial.

Given a register machine program P and input $m \in \omega$, we define a first-order formula $\chi_{P,m}$ such that $\omega \models \chi_{P,m}$ iff P halts on input m . Intuitively, $\chi_{P,m}$ says “There is a finite sequence

$$0, m, 0, \dots, 0, L_1, x_{1,0}, \dots, x_{1,n}, L_2, x_{2,0}, \dots, x_{2,n}, L_3, \dots, L_z, x_{z,0}, \dots, x_{z,n}$$

such that for all i , the $(i + 1)$ th block $L_{i+1}, x_{i+1,0}, \dots, x_{i+1,n}$ is obtained from the i th block $L_i, x_{i,0}, \dots, x_{i,n}$ by applying the instruction L_i to the registers $x_{i,0}, \dots, x_{i,n}$ to obtain registers $x_{i+1,0}, \dots, x_{i+1,n}$, and such that L_z is the Halt instruction.” Thus $\omega \models \chi_{P,m}$ iff we can code the configuration sequence which witnesses that P halts on input m in ω (which will be possible provided that such a configuration sequence exists, by the β -lemma).

Proof of Theorem 2.6. Let φ_β be the definition of β . So $\omega \models \varphi_\beta(x, y, z, w)$ iff $\beta(x, y, z) = w$. Let H denote the label for the halting instruction.

Then $\chi_{P,m}$ says that there are t, p, z such that the following conjunction holds:

1. $\varphi_\beta(t, p, 0, 0)$
2. $\varphi_\beta(t, p, 1, \bar{m})$
3. For every $j \in \{2, \dots, n+2\}$, $\varphi_\beta(t, p, \bar{j}, \bar{0})$
4. $\forall i < z \forall L \forall x_0 \dots \forall x_n \forall L' \forall x'_0 \dots \forall x'_n$, if the conjunction
 - (a) $\varphi_\beta(t, p, i \times \overline{n+2}, L)$
 - (b) For every $j \in \{1, \dots, n+1\}$, $\varphi_\beta(t, p, i \times \overline{n+2} + \bar{j}, x_0)$
 - (c) For every $j \in \{0, \dots, n+1\}$, $\varphi_\beta(t, p, (i+1) \times \overline{n+2} + \bar{j}, L')$
5. $\varphi_\beta(t, p, \bar{z} \times \overline{n+2} + \bar{1}, \bar{H})$

The clause ψ in quotes (“ (L', x'_0, \dots, x'_n) is obtained...”) is tedious but easy to define. It is a conjunction of H many clauses, one for each line of P that is not a halting instruction.

We do not construct all conjuncts of ψ but give an example of just one clause. Suppose that at label ℓ we have the instruction “Let $R_g = R_g + a$.” Then the ℓ th conjunct of ψ is

$$(L = \bar{\ell}) \rightarrow \left(L' = \overline{\ell+1} \wedge (x'_g = x_g + 1) \wedge \bigwedge_{j \neq g} (x'_j = x_j + 1) \right).$$

The other conjuncts are defined similarly.

Therefore if we can determine if $\chi_{P,P} \in T$, we can decide whether $P \in \mathbf{Halt}$. \square

Note that $\chi_{P,m}$ is of the form $\exists p \exists t \exists z \psi_{P,m}$ where $\psi_{P,m}$ has bounded quantifiers; p, t control how large we have to check for each quantifier in $\psi_{P,m}$. Thus $\chi_{P,m}$ has a very low quantifier complexity, but even this is undecidable, so we have actually proven something stronger than the claimed theorem:

Corollary 2.9. The set of true Σ_1^0 sentences in ω is not decidable.

Stronger than this, the set of multivariable polynomials p over \mathbb{Z} such that p has an integer zero is not decidable. This theorem does not even have a bounded quantifier, but just a single unbounded \exists .

Corollary 2.10 (representability theorem). Suppose D is a decidable predicate on ω . Then the definition of D has quantifier complexity Σ_1^0 .

To see this, simply use the formula $\chi_{P,m}$ where P is defined to halt iff D returns True. In particular, every computable function $\omega^n \rightarrow \omega$ has a Σ_1^0 definition. The converse is not true, as we already proved.

2.3 Goedel's completeness theorem

Theorem 2.11 (completeness). For every set of formula Γ and every formula φ , if $\Gamma \models \varphi$, then $\Gamma \vdash \varphi$.

Here $\Gamma \models \varphi$ means that for every (M, ν) , if $(M, \nu) \models \Gamma$, then $(M, \nu) \models \varphi$. Meanwhile, $\Gamma \vdash \varphi$ simply means that there is a proof P of φ from Γ , namely a finite sequence of formulae, each of which follows from those before it and elements of Γ by modus ponens, such that the final formula in P is φ .

In particular, if Γ is decidable, then the set of φ such that $\Gamma \vdash \varphi$ is computably enumerable; namely, iterate over all finite sequences of formulae which follow from elements of Γ by modus ponens, and check if each one is a proof.

Definition 2.12. A *theory* is a set of formulae which is closed under the relation \vdash . A *complete theory* is a theory T such that for every sentence φ , either $\varphi \in T$ or $\neg\varphi \in T$.

In other words, if T is complete then for every φ , $T \vdash \varphi$ or $T \vdash \neg\varphi$.

Example 2.13. Let $\text{Th}(M)$ be the set of φ such that $M \models \varphi$. Then $\text{Th}(M)$ is a complete theory.

Now note that every proof has a finite length, and in particular there is an algorithm that decides whether each line in a sequence of formulae follows from those before it, whence the set of proofs is decidable. Therefore we have the following theorem.

Theorem 2.14 (compactness). Let Γ be a set of sentences and φ a formula. If $\Gamma \vdash \varphi$ then there is a finite subset $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \vdash \varphi$.

We start by assuming that the language of Γ has no function symbols. In fact every function is a relation, so any formula that is true about a function is also true about some relation. If F is a function let G_F be the graph of F ; then for any φ ,

$$\varphi(F(x)) \leftrightarrow \exists y(G_F(x, y) \wedge \varphi(y)).$$

We now carry out a modified Henkin construction. Fix a formula φ . Expand the language of Γ by adding a countable set of constant symbols $\{c_n\}_n$ and fix an enumeration $\{\theta_n\}_n$ of the sentences in the expanded language such that for every $j < i$, c_i does not appear in θ_j . (This is possible by choosing a sentence that has no c_n 's, then a sentence that has c_0 , then a sentence with no c_n 's, then a sentence that has c_0 , then a sentence that has c_0 or c_1 , then a sentence that has no c_n 's, et cetera. Then before we get to a sentence that has c_n we have already taken n sentences that had no c_i 's at all.)

Now consider the binary tree of all possible attempts to satisfy all the formulae in Γ and $\neg\varphi$. The root of the tree is $\neg\varphi$. Then include two branches: θ_0 and $\neg\theta_0$. Now c_0 does not appear in θ_0 , so if $\theta_0 = \exists x\psi_0(x)$, we can define c_0 to be the witness to θ_0 . In this case, θ_0 only has one child: $\psi_0(c_0)$.

We inductively assume that we have constructed a finite binary tree G , and that certain nodes have been “terminated.” At each stage we give each nonterminated leaf two children, namely θ_n and $\neg\theta_n$, and if $\theta_n = \exists x\psi_n(x)$, we give θ_n the child $\psi_n(c_n)$. We then also add

$\psi_k(c_n)$ for every k such that $\forall x\psi_k(x)$ has already appeared. Then, we terminate one of the new leaves L if there is no finite structure which satisfies all the quantifier-free sentences which appear in the path from $\neg\varphi$ to L .

In fact, since such sentences are quantifier-free, they only refer to the c_k , and so it suffices to check finitely many sentences on the set of all relational structures defined on the set $\{c_j : j < k\}$, which can be done computably! Thus one could concievably build the infinite path by breadth-first search.

Suppose that M is a model of $\Gamma \cup \{\neg\varphi\}$. Then we construct an infinite path through G : at each θ_n , if $M \models \theta_n$, go down the path θ_n (and conversely for $\neg\theta_n$) and if necessary interpret c_n to be the least witness to $\exists x\psi_n(x)$ after choosing some well-ordering of M (which is possible with just countable choice by the Lowenheim-Skolem theorem).

Conversely, if P is an infinite path through G , we define a structure M_P . Elements of M_P will be the constant symbols $\{c_n\}_n$ modulo the equivalence relation that if a formula that appear in P is of the form $c_i = c_j$ then $c_i \sim c_j$. (If a formula implies $c_i = c_j$ then termination implies that $c_i = c_j$ appears at a later stage. So it suffices to just check formulae of the form $c_i = c_j$.) We interpret the constants c not of the form c_i to simply map to the equivalence class of c_j if $c = c_j$ appears (and it must appear for some j). Then we interpret $R([c_{i_1}], \dots, [c_{i_k}])$ to be true if $R(c_{i_1}, \dots, c_{i_k})$ appears in P .

Lemma 2.15. For every sentence χ , $M \models \chi$ iff $\chi \in P$.

Proof. By induction on quantifier complexity. This is true by definition for sentences of the form $c_i = c_j$. Moreover if a, b are constants and $a = b$ appears in P , then there are some i, j such that $a = c_i$ and $b = c_j$, but then P must contain $c_i = c_j$ whence $M \models (a = b)$, and the converse follows from the same argument in reverse. A similar argument checks relational sentences as well as boolean operations.

Finally we consider existential sentences. Suppose $M \models \exists x\psi(x)$; then we can find a minimal i such that $M \models \psi([c_i])$, so by induction $\psi([c_i]) \in P$ and hence $\theta_i = \exists x\psi(x)$ and $\theta_i \in P$. The converse is similar. \square

We now appeal to weak Koenig's lemma:

Lemma 2.16 (weak Koenig's lemma). Every infinite binary tree has an infinite path.

Proof. The first n levels have at most 2^n nodes, so there must be more than n nodes for every n . \square

The decision tree G is a binary tree, so either G is finite and every path terminates (in which case $\Gamma \cup \{\neg\varphi\}$ is inconsistent, and we have a finite proof of this fact, namely G), or G is infinite and has an infinite path P , which gives rise to a model M_P of $\Gamma \cup \{\neg\varphi\}$.

In fact it is tempting to define $\Gamma \vdash \varphi$ if there is a terminated binary tree G , satisfying the properties above, whose root is $\neg\varphi$. It is computable to check that the tree has terminated, so \vdash is a finitistic object. With this interpretation the above construction is a proof of the completeness theorem. It is also a proof that weak Koenig's lemma is not constructive, since first-order logic is not decidable.

2.4 Motivating Peano arithmetic

Definition 2.17. A theory Γ is *recursively axiomatizable* if there is a subset of Γ which is computable and which generates Γ .

Corollary 2.18. The theory of true arithmetic is not recursively axiomatizable.

Proof. If it was, then we could use the completeness theorem to decide truth in ω . \square

We now prove something much stronger.

Theorem 2.19 (Goedel's first incompleteness theorem). There is a fragment T of true arithmetic such that for every $T^* \supseteq T$ such that if T^* is consistent and computable, then T^* is incomplete.

Theorem 2.20 (Goedel's second incompleteness theorem). For every $T^* \supseteq T$ which is consistent and computable, T^* does not prove $\text{Con } T^*$.

We prove these theorems after setting up the formal theory of incompleteness. Let Φ be a set of sentences in the language $\{0, 1, +, \times\}$.

Definition 2.21. An n -ary relation D on ω is a *representable relation* in Φ if there is a formula φ such that for all $m_0, \dots, m_{n-1} \in \omega$, if $D(m_0, \dots, m_{n-1})$ then $\Phi \vdash \varphi(\overline{m_0}, \dots, \overline{m_{n-1}})$, and otherwise, $\Phi \vdash \neg\varphi(\overline{m_0}, \dots, \overline{m_{n-1}})$. A function $f : \omega^n \rightarrow \omega$ is a *representable function* in Φ if the graph of f is representable in Φ by a formula φ , and for every m_0, \dots, m_{n-1} , $\Phi \vdash \exists!y \varphi(\overline{m_0}, \dots, \overline{m_{n-1}}, y)$.

Here $\exists!$ is “there is a unique”.

Example 2.22. Every computable function and every computable relation is representable in $\text{Th } \mathbb{N}$. This follows from the β -lemma. On the other hand, Φ is inconsistent if and only if every relation and every function is representable in Φ , by a cardinality argument.

If Φ is consistent and computable then only relations and functions representable in Φ must also be computable, since if Φ is computable then we can decide if $\Phi \vdash \varphi$ or not.

2.5 Peano arithmetic

We now give a possible recursive axiomatization of arithmetic.

Definition 2.23. *Peano arithmetic PA* is the theory consisting of the universal closures of $\neg((x+1) = 0)$, $(x+0) = x$, $x \times 0 = 0$, $((x+1) = (y+1)) \rightarrow (x = y)$, $(x+(y+1)) = ((x+y)+1)$, and $(x \times (y+1)) = ((x \times y) + x)$, as well as, for every formula φ with a free variable x , the axiom

$$(\varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(x+1))) \rightarrow \forall x\varphi(x).$$

So **PA** implies the usual, set-theoretic principle of induction for definable sets.

Suppose $M \models \text{PA}$. Then we will show that M admits the relation $<$, and there is a $<$ -initial segment which is isomorphic in $\{0, 1, +, \times\}$ to \mathbb{N} , the *standard part* of M . If M

has further elements beyond its standard part we will refer to it as the *nonstandard part*. Therefore any existential statement in $\text{Th } \mathbb{N}$ is already true in \mathbf{PA} . We will then show that every computable function and computable relation is representable in \mathbf{PA} .

We now prove a bunch of theorems inside \mathbf{PA} .

Theorem 2.24 (PA). $\forall x \forall y \forall z ((x + y) + z) = (x + (y + z))$.

Proof. By induction on z . If $z = 0$ we have $(x + y) + 0 = x + y = x + (y + 0)$. Now suppose $(x + y) + z = x + (y + z)$, then $(x + y) + (z + 1) = ((x + y) + z) + 1 = (x + (y + z)) + 1 = x + ((y + z) + 1) = x + (y + (z + 1))$. Therefore the theorem. \square

Theorem 2.25 (PA). $\forall x \forall y x + y = y + x$.

Proof. We first prove $\forall x (x + 0 = 0 + x)$ by induction on x . $0 + 0 = 0 + 0$ and if $x + 0 = 0 + x$ then $(x + 1) + 0 = x + 1 = (0 + x) + 1 = 0 + (x + 1)$. Therefore the claim.

We now prove $\forall x (x + 1 = 1 + x)$ by induction on x . We already know $0 + 1 = 1 + 0$, and if $x + 1 = 1 + x$ then $(x + 1) + 1 = (1 + x) + 1 = 1 + (x + 1)$, therefore the claim.

Finally we prove the theorem by induction on x . We already know $0 + y = y + 0$. If $x + y = y + x$ then $(x + 1) + y = x + (1 + y) = x + (y + 1) = (x + y) + 1 = (y + x) + 1 = y + (x + 1)$. Therefore the theorem. \square

Henceforth I will omit the proofs of basic facts in \mathbf{PA} unless some trick was involved.

Theorem 2.26 (PA). Multiplication is commutative and associative and distributes over addition. Addition and multiplication are injective.

We now begin the proof of the following theorem:

Theorem 2.27. A certain order relation $<$ is definable in \mathbf{PA} with the following property: Let M be a model of \mathbf{PA} . Then a $<$ -initial segment of M is $(0, 1, +, \times, <)$ -isomorphic to ω .

We first show that if φ is an equality of arithmetic terms which is true in \mathbb{N} (e.g. $\varphi : ((1 + 1 + 1 + 1 + 1) = ((1 + 1) \times (1 + 1 + 1)))$, then $\mathbf{PA} \vdash \varphi$. This requires us to go outside \mathbf{PA} – we are proving a fact about \mathbf{PA} itself, not proving something from \mathbf{PA} .

Lemma 2.28. For any $i, j \in \mathbb{N}$, if $i \neq j$ then $\mathbf{PA} \vdash \bar{i} \neq \bar{j}$.

Proof. Without loss of generality we may assume $i < j$ and then induct on i . If $i = 0$ then there is a k such that $j = k + 1$, and $\mathbf{PA} \vdash \bar{k} + 1 = 0$ so we're done. Otherwise, assume $\mathbf{PA} \vdash \bar{i} - 1 \neq \bar{j} - 1$. Therefore $\mathbf{PA} \vdash \bar{i} - 1 + 1 \neq \bar{j} - 1 + 1$ since \mathbf{PA} proves that addition is injective. But $\bar{i} = \bar{i} - 1 + 1$ by definition and similarly for \bar{j} so $\mathbf{PA} \vdash \bar{i} \neq \bar{j}$. \square

On the other hand obviously $\mathbf{PA} \vdash \bar{i} = \bar{i}$. So \mathbf{PA} gets facts about equality of natural numbers correct.

Lemma 2.29. For all $i, j \in \mathbb{N}$, $\mathbf{PA} \vdash \bar{i} + \bar{j} = \bar{i + j}$.

Proof. Fix \mathbb{N} and induct on j . If $j = 0$ this follows immediately. If $\mathbf{PA} \vdash \bar{i} + \bar{j} = \bar{i + j}$, then $\bar{i} + \bar{j + 1} = \bar{i} + (\bar{j} + 1)$ so by associativity $\mathbf{PA} \vdash \bar{i} + \bar{j + 1} = (\bar{i} + \bar{j}) + 1 = \bar{i + j + 1}$ so the claim follows by induction. \square

Similarly **PA** proves facts about multiplication on \mathbb{N} correctly.

Definition 2.30 (PA). Let $x_1 < x_2$ be the formula

$$\exists y \ x_1 + (y + 1) = x_2.$$

In other words there is a nonzero number $y + 1$ such that $x_1 + y + 1 = x_2$.

Lemma 2.31 (PA). $\forall y \ (y \neq 0) \rightarrow (\exists z \ y = (z + 1))$.

Proof. The formula

$$y = 0 \wedge (\exists z \ y = (z + 1))$$

satisfies the hypotheses of an induction axiom. \square

Lemma 2.32. For every $j \in \mathbb{N}$,

$$\mathbf{PA} \vdash \forall x \ (x < \bar{j}) \rightarrow (x = 0 \vee x = 1 \vee \dots \vee x = \bar{j-1}).$$

Proof. By induction on j . If $j = 0$ we are proving $(x < 0) \rightarrow (0 = 1)$ which follows from $\neg(x < 0)$. In fact if $x < 0$ then $\mathbf{PA} \vdash \exists y \ x + (y + 1) = 0$ which contradicts an axiom.

If **PA** proves the claim for j , suppose $x < \bar{j+1}$. Therefore $\exists y \ x + (y + 1) = \bar{j+1}$, and fixing that y we have $x + y = \bar{j}$. If $y = 0$ then $x = \bar{j}$. Otherwise $\exists z \ y = z + 1$, so by induction $\mathbf{PA} \vdash x = 0 \wedge \dots \wedge x = \bar{j-1}$. \square

Lemma 2.33. For all $i, j \in \mathbb{N}$, $\mathbf{PA} \bar{i} < \bar{j}$, iff $i < j$.

Proof. Suppose $i < j$. Then $j - i \in \mathbb{N}$, $j - i \neq 0$, and because **PA** gets facts about addition correct, $\mathbf{PA} \vdash \bar{i} + (\bar{j-i-1} + 1) = \bar{j}$. In particular $\mathbf{PA} \vdash \exists y \bar{i} + (y + 1) = \bar{j}$.

Conversely, we use Lemma 2.32. \square

Finally to prove Theorem 2.27, we just need to show that in every model M , $(M, <)$ is a chain; then Lemma 2.32 immediately gives a bijection between a $<$ -initial segment and \mathbb{N} , and the other lemmata imply that this bijection preserves $(0, 1, +, \times, <)$.

Lemma 2.34 (PA). $\forall x \forall y \ (x < y) \vee (x = y) \vee (y < x)$.

Proof. By induction. If $x = 0, y = 0$, done. Otherwise if $x = 0$ and $y \neq 0$ then $\exists z \ (z + 1) = y$, so $\exists z \ ((0 + z) + 1) = y$ so $y > x$.

Reasoning inductively, we know $(x < y) \vee (x = y) \vee (y < x)$. If $y < x$ then $\exists z \ (y + (z + 1)) = x$ and then $y < x + 1$. If $y = x$ then $y + (0 + 1) = x + 1$ so $y < x + 1$. If $x < y$ then $\exists t \ (x + (t + 1)) = y$. If $t = 0$ then $y = x + 1$. Else $\exists s \ ((x + 1) + (s + 1)) = y$ so $y > x + 1$. \square

Lemma 2.35 (PA). $\forall x \forall y \ ((x < y) \rightarrow \neg(x = y)) \wedge (x = y) \rightarrow \neg(x < y)$.

Proof. Suppose $x < y$ and $x = y$, then $\exists z \ (x + (z + 1)) = x$ so $z + 1 = 0$ so $0 = 1$. Similarly for the other conjunct. \square

Therefore we have proven Theorem 2.27.

We now state a few more facts about **PA** before discussing its models in greater detail.

Lemma 2.36 (PA). $(y < x + 1) \leftrightarrow ((y < x) \vee (y = x))$.

Theorem 2.37 (least-number principle). For any formula φ ,

$$\mathbf{PA} \vdash (\exists x \varphi(x)) \rightarrow (\exists x \varphi(x) \wedge \forall y < x \neg \varphi(y)).$$

Proof. Assume the consequent fails and work inside **PA**. Then

$$\forall x (\forall y < x \neg \varphi(y)) \rightarrow \neg \varphi(x).$$

Let $\psi(x) = \forall y \leq x \neg \varphi(y)$. Then trivially $\forall y < 0 \neg \varphi(y)$ so $\neg \varphi(0)$ and hence $\psi(0)$. Suppose $\psi(x)$. Then $y \leq x + 1$ iff $y = x + 1$ or $y \leq x$. If $y \leq x$ then $\neg \varphi(x + 1)$. So by induction $\forall x \psi(x)$, and hence $\forall x \neg \varphi(x)$. \square

Recall that a sentence is Σ_1^0 if there is an n and a formula θ with bounded quantifiers (i.e. a Δ_0^0 formula) such that the sentence is

$$\exists x_1 \cdots \exists x_n \theta.$$

Now if $M \subset M'$ are structures in the language of a Σ_1^0 sentence φ , $M \models \varphi$, and M is an initial segment of M' , then $M' \models \varphi$: any witnesses to the truth of θ must appear in M and hence also in M' , since all quantifiers are existential or bounded by existential quantifiers.

Corollary 2.38. Let φ be a Σ_1^0 sentence and suppose that $\mathbb{N} \models \varphi$. Then **PA** $\vdash \varphi$.

Proof. Let M be a model of $\varphi = \exists x_1 \cdots \exists x_n \theta$, so \mathbb{N} is an initial segment of M . Let w_1, \dots, w_n be witnesses to the truth of φ in \mathbb{N} , so

$$\mathbb{N} \models \theta(w_1, \dots, w_n).$$

Since θ has bounded quantifiers we also have $M \models \theta(w_1, \dots, w_n)$ whence the claim. \square

Now we have already proven that there is a Σ_1^0 sentence which is not decidable in a computably axiomatizable theory, in particular not in **PA**. This corollary therefore implies that any such sentence must be false in \mathbb{N} , so while we cannot disprove that sentence from **PA**, we can disprove it by reasoning externally about **PA**.

Corollary 2.39. **PA** does not prove the full induction schema.

Proof. Let M be a nonstandard model. Then $\omega \subset M$ and ω is closed under 0 and successor. \square

Corollary 2.40. ω is not definable from **PA**.

Proof. If it was definable, then ω would be equal to the entire model by the schema of definable induction. \square

In particular, second-order logic allows us to consider the full induction schema rather than just the definable schema, and then reasoning internally, unique model of arithmetic up to isomorphism. But reasoning externally, there are in fact multiple models of second-order arithmetic, because the power set of ω is not uniquely determined (by Cohen's theorem, for example).

Corollary 2.41. Let M be a nonstandard model, and let $m \in M$ be nonstandard. Then there is a $m' < m$ such that m' is nonstandard. In fact, m is contained in a copy of \mathbb{Z} whose elements are nonstandard in M , but m' is not in that copy of \mathbb{Z} .

Proof. Either m or $m + 1$ is even. Then either $m/2$ or $(m + 1)/2$ exists, so let m' be the one that exists. Then $m' \notin \omega$ because if it was then so would be m . Moreover $m, m/2$ are successors of something, and $m - m/2$ is nonstandard, hence the claim about \mathbb{Z} . \square

In fact, if M is a countable nonstandard model, then there are \mathbb{Q} many copies of \mathbb{Z} after ω . This can be shown using the uniqueness of countable dense linear orders without endpoints. However, the fact that M has ordertype $\omega + \mathbb{Q}\mathbb{Z}$ does not uniquely determine the arithmetic of M . On the other hand, a theorem of Friedman implies that every nonstandard model M contains an initial segment which is isomorphic to M .

In spite of this anomaly, most of elementary number theory, at least those facts which can be established without complex analysis, is provable from **PA**, because it is just given by algebraic manipulation.

2.6 Goedel codes in Peano arithmetic

Recall that the β -lemma implies that if **PA** can encode pairs, then **PA** can code arbitrarily long vectors. We now show that in fact **PA** can code pairs.

Definition 2.42. Define the *pairing function*

$$J(p, t) = \frac{(p+t)(p+t+1)}{2} + p.$$

Then J is a bijection $\omega^2 \rightarrow \omega$, because J computes the position of (p, t) in the ordering on ω^2 defined by $(a, b) < (c, d)$ iff $a + b < c + d$ or $(a + b) = (c + d) \wedge a < c$. Moreover, J and the projections of J^{-1} are recursive; and **PA** in fact proves that J is a bijection.

Definition 2.43. For every $n \in \omega$, if $J(p, t) = n$, then we let $(n)_0 = p$ and $(n)_1 = t$.

Given a recursive $D \subseteq \omega^n$, let $\psi(p, t, x)$ be the formula that means that (p, t) encodes, in the sense of the β -lemma, a sequence of configurations to compute whether $x \in D$, and that in fact this computation proves that $x \in D$. Then let $\varphi(x)$ be the formula

$$\begin{aligned} \exists i (\exists p < i \exists t < i \ i = J(p, t) \wedge \psi(p, t, x)) \\ (\wedge \forall j < i \neg(\exists q < j \exists s < j \ j = J(s, q) \wedge \psi(q, s, x))) \end{aligned}$$

so $\varphi(x)$ says there is a *minimal* i such that $(i)_0, (i)_1$ codes, in the sense of the β -lemma, a sequence of configurations to determine whether $x \in D$.

If $m \in D$ then $\varphi(m)$ is true in ω , and $\varphi(m)$ is Σ_0^1 , so **PA** $\vdash \varphi(\bar{m})$. Conversely, if $m \notin D$, then **PA** proves that there is a nonaccepting computation for \bar{m} , coded by j , then **PA** can prove that for every $i \leq j$, i does not code an accepting computation for \bar{m} . Nothing bigger than j satisfies φ so in fact **PA** $\vdash \neg\varphi(\bar{m})$. Therefore we have proven the following result.

Theorem 2.44. If $D \subseteq \omega^n$ is recursive, then D is Σ_0^1 -definable.

To show that a recursive function f is Σ_0^1 -definable, we not only show that the graph of f is Σ_0^1 -definable (which follows from the above) but also must show that **PA** proves that the graph satisfies the vertical line test. Let f be given and let $\varphi(x, y) = \exists w \theta(x, y, w)$ be the definition of the graph $\{(x, y) \in \omega^{n-1} \times \omega : f(x) = y\}$ as given by the pairing function. Then θ has bounded quantifiers and we define $\varphi^*(x, y)$ by

$$\exists w \theta(x, y, w) \wedge \forall z < J(y, w) \neg \theta(x, (z)_0, (z)_1).$$

Thus $\varphi^*(x, y)$ iff y is the least y' such that $f(x) = y'$. So the relation defined by $\varphi^*(x, y)$ must be the graph of a function. By the least number principle, **PA** can then prove that there is a $(y, w) \in \omega^2$ such that $\theta(m, y, w)$ and $J(y, w)$ is minimal possible. That is, $\mathbf{PA} \vdash \exists!y \varphi^*(m, y)$, and it is easy to check that **PA** proves that the unique such y is $f(m)$. Therefore recursive functions are Σ_0^1 -definable.

2.7 Goedel's first incompleteness theorem

Theorem 2.45 (Goedel's first incompleteness theorem). Let Φ be a computable set of sentences in the language $(0, 1, +, \times)$. If Φ allows representation of recursive functions and relations, then Φ is incomplete. In fact, we can explicitly give a sentence φ such that Φ cannot decide φ .

Note that the assumption that the language is $(0, 1, +, \times)$ can be dropped. In fact, the theorem holds for the language of set theory, since the set of Zermelo-Fraenkel axioms without infinity is bi-interpretable with **PA**.

To prove this theorem, we must internalize self-reference in Φ . Fix a computable enumeration of the set of all formulas in the language $(0, 1, +, \times)$, say $\{\varphi_n\}_n$, and assume that this enumeration is a bijection. If φ is a formula, let n_φ be such that $\varphi = \varphi_{n_\varphi}$. We note that since syntactic operations are computable, they correspond to operations on indices. For example, finding m such that $\varphi_{n_1} \rightarrow \varphi_{n_2}$ is φ_m gives a computable map $(n_1, n_2) \mapsto m$.

Theorem 2.46 (Goedel's fixed point theorem). Let Φ be a set of sentences which allows for representations of computable functions and relations. Then for any formula ψ there is a sentence φ such that

$$\Phi \vdash (\varphi \leftrightarrow \psi(\overline{n_\varphi})).$$

Proof. Let ψ be given. Let $F : \omega^2 \rightarrow \omega$ be defined by

$$F(m_1, m_2) = \begin{cases} n_{\varphi_{m_1}(\overline{m_2})} & , \text{ if } \varphi_{m_1} \text{ has one free variable} \\ 0 & \text{else.} \end{cases}$$

So if φ_{m_1} has one free variable then $F(m_1, m_2)$ is the index of $\varphi_{m_1}(\overline{m_2})$. By assumption F is computable, therefore representable in Φ . Let $\theta(x)$ represent the graph of F , so $\Phi \vdash \theta(x)$ iff $F(x_0, x_1) = x_2$. Let $\beta(x) = \forall z \theta(x, x, z) \rightarrow \psi(z)$. Thus $\beta(x)$ says that $\psi(F(x, x))$; in other words, the index of $\varphi_x(\overline{x})$ satisfies ψ . Let

$$\varphi(x) = \beta(n_\beta).$$

Then φ asserts that ψ holds at the index of $\varphi_{n_\beta}(\overline{n_\beta})$.

We first show $\Phi, \varphi \vdash \psi(\overline{n_\varphi})$. In fact,

$$\Phi, \varphi \vdash \theta(n_\beta, n_\beta, n_{\beta(\overline{n_\beta})}) \rightarrow \psi(\overline{n_{\beta(\overline{n_\beta})}}).$$

This is true because if $F(n_\beta, n_\beta) = n_{\beta(\overline{n_\beta})}$ and $\beta(\overline{n_\beta})$ (i.e. φ) then $\psi(n_{\beta(\overline{n_\beta})})$. But by definition of θ , and because $F(n_\beta, n_\beta) = n_{\beta(\overline{n_\beta})}$, we have

$$\Phi, \varphi \vdash \theta(n_\beta, n_\beta, n_{\beta(\overline{n_\beta})}).$$

Therefore the claim.

We finally show $\Phi, \psi(\overline{n_\varphi}) \vdash \varphi$. Now $n_\varphi = n_{\beta(\overline{n_\beta})}$. By definition of θ ,

$$\Phi \vdash \exists!y \theta(\overline{n_\beta}, \overline{n_\beta}, y)$$

but also $F(n_\beta, n_\beta) = n_{\beta(\overline{n_\beta})}$. Therefore

$$\Phi \vdash \forall z \theta(\overline{n_\beta}, \overline{n_\beta}, z) \rightarrow (z = \overline{n_{\beta(\overline{n_\beta})}})$$

or in other words

$$\Phi \vdash \forall z \theta(\overline{n_\beta}, \overline{n_\beta}, z) \rightarrow (z = \overline{n_\varphi}).$$

Hence

$$\Phi, \psi(\overline{n_\varphi}) \vdash \forall z \theta(\overline{n_\beta}, \overline{n_\beta}, z) \rightarrow \psi(z)$$

and $\varphi = \forall z \theta(\overline{n_\beta}, \overline{n_\beta}, z) \rightarrow \psi(z)$. \square

To give an idea of how this theorem will be used, we will construct $\psi(n)$ to mean “There is no proof of φ_n .” The proof seems a little mysterious until we recall the proof of Banach’s contraction fixed point theorem: one iterates the contraction until we have found a fixed point. Similarly we iterate F until we find a fixed-point. We note that in the proof of $\Phi, \psi(\overline{n_\varphi}) \vdash \varphi$, we used the fact that $\Phi \vdash \exists!y \theta(\overline{n_\beta}, \overline{n_\beta}, y)$, which is why we went to the pain of requiring that Φ can actually prove that function representations are single-valued – something that was not easy to do when $\Phi = \mathbf{PA}$. Doing this allowed us to replace the $\forall z$ with $z = n_{\beta(\overline{n_\beta})}$ since that is the only z meeting the antecedent.

Corollary 2.47. Suppose Φ meets the hypotheses of Goedel’s fixed point theorem. Let $\Phi^\vdash \subseteq \omega$ encode the theory generated by Φ , i.e.

$$\Phi^\vdash = \{n_\varphi : \Phi \vdash \varphi\}.$$

If Φ is consistent then Φ^\vdash is not representable in Φ .

Proof. Suppose that Φ^\vdash is representable by ψ . Let φ be such that

$$\Phi \vdash \varphi \leftrightarrow \neg\psi(\overline{n_\varphi});$$

i.e. φ says “I am not provable from Φ .” Then $\Phi \vdash \varphi$ iff $n_\varphi \in \Phi^\vdash$ iff $\Phi \vdash \psi(\overline{n_\varphi})$ iff $\Phi \vdash \neg\varphi$, so Φ is inconsistent. \square

Proof of Goedel’s first incompleteness theorem. Suppose that Φ is complete. Then Φ^\vdash is computable, since Φ is computable, Φ^\vdash is computably enumerable, and we can just run two programs in parallel to check whether $\varphi \in \Phi^\vdash$ or $\neg\varphi \in \Phi^\vdash$ since Φ is complete. So Φ^\vdash is representable in Φ , which by the previous corollary implies that Φ is inconsistent. \square

2.8 Goedel's second incompleteness theorem

Goedel's second incompleteness theorem is much harder to prove, because to do so requires us to fill in lots of details about how certain proof systems behave. We will not do this, but sketch the proof.

To prepare for the proof, fix a recursively axiomatized extension Φ of **PA**. We might as well assume that Φ is consistent (if not, use the principle of explosion). We will formalize the proof of the first incompleteness theorem within Φ . Fix a recursive enumeration of all valid proofs from Φ and let $H(n, m)$ be the recursive predicate which checks whether m encodes the proof of φ_n . Let φ_H be the formula that represents H in Φ . Let $\text{Der}_\Phi(x)$ be the formula $\exists y \varphi_H(x, y)$; intuitively $\text{Der}_\Phi(n)$ means there really is a proof of φ_n .

Fix a formula φ such that

$$\Phi \vdash \varphi \leftrightarrow \neg \text{Der}_\Phi(\overline{n_\varphi}).$$

This formula exists by Goedel's fixed point theorem.

Lemma 2.48. $\Phi \not\vdash \varphi$.

Proof. If $\Phi \vdash \varphi$, then let m encode the proof of φ ; then $H(n_\varphi, m)$. So $\varphi_H(\overline{n_\varphi}, \overline{m})$ holds, so does $\text{Der}_\Phi(\overline{n_\varphi})$, so $\Phi \vdash \neg \varphi$, a contradiction. \square

Now define $\text{Con } \Phi$ to be the sentence $\neg \text{Der}_\Phi(\overline{n_{0=1}})$.

Theorem 2.49 (Goedel's second incompleteness theorem). If Φ is consistent then $\Phi \neg \vdash \text{Con } \Phi$.

By the lemma,

$$\omega \models (\text{Con } \Phi \rightarrow \neg \text{Der}_\Phi(\overline{n_\varphi})).$$

Here we need to work in the model ω to be sure that the definition of a “proof” is what we intend it to be; there are no proofs where the number of formulae appearing in the proof is nonstandard. We discard this assumption below.

We now study properties of Der_Φ .

Lemma 2.50. One has:

1. If $\Phi \vdash \varphi_n$ then $\Phi \vdash \text{Der}_\Phi(n)$.
2. $\Phi \vdash (\text{Der}_\Phi(\overline{n_{\varphi_{m_1} \rightarrow \varphi_{m_2}}}) \rightarrow (\text{Der}_\Phi(\overline{m_1}) \rightarrow \text{Der}_\Phi(\overline{m_2})))$.
3. $\Phi \vdash (\text{Der}_\Phi(\overline{n_\psi}) \wedge \text{Der}_\Phi(\overline{\neg n_\psi})) \rightarrow \text{Der}_\Phi(\overline{n_{0=1}})$.
4. $\Phi \vdash \text{Der}_\Phi(\overline{n}) \rightarrow \text{Der}_\Phi(\overline{n_{\text{Der}_\Phi(\overline{n})}})$

We prove this lemma by inducting on the length of a proof.

Proof of Goedel's second incompleteness theorem. We need to prove that $\Phi \vdash \text{Con } \Phi$. We know that $\Phi \vdash \varphi \leftrightarrow \neg \text{Der}_\Phi(\overline{n_\varphi})$. Then:

$$\begin{aligned}
& \Phi \vdash \text{Der}_\Phi(\overline{n_{\varphi \rightarrow (\text{Der}_\Phi(\overline{n_\varphi}) \rightarrow (0=1))}}) \\
& \Phi \vdash \text{Der}_\Phi(\overline{n_\varphi}) \rightarrow (\text{Der}_\Phi(\overline{n_{\text{Der}_\Phi(\overline{n_\varphi})}}) \rightarrow \text{Der}_\Phi(\overline{n_{0=1}})) \\
& \Phi \vdash \text{Der}_\Phi(\overline{n_\varphi}) \rightarrow \text{Der}_\Phi(\overline{n_{0=1}}) \\
& \Phi \vdash \varphi \leftrightarrow \neg \text{Der}_\Phi(\overline{n_\varphi}) \\
& \Phi \vdash \varphi \rightarrow \neg \text{Der}_\Phi(\overline{n_\varphi}) \\
& \Phi \vdash \text{Der}_\Phi(\overline{n_{\varphi \rightarrow \text{Der}_\Phi(\overline{n_\varphi})}}) \\
& \Phi \vdash \text{Der}_\Phi(\overline{n_\varphi}) \rightarrow \text{Der}_\Phi(\overline{n_{\neg \text{Der}_\Phi(\overline{n_\varphi})}}) \\
& \Phi \vdash \neg \text{Der}_\Phi(\overline{n_{0=1}}) \rightarrow \neg \text{Der}_\Phi(\overline{n_\varphi}) \\
& \Phi \vdash \text{Con } \Phi \rightarrow \varphi.
\end{aligned}$$

So if $\Phi \vdash \text{Con } \Phi$ then $\Phi \vdash \varphi$. We know that if this is true then $\Phi \vdash \neg \text{Con } \Phi$, a contradiction. \square

Chapter 3

Kleene's recursion theorem

3.1 Kleene's recursion theorem

Definition 3.1. Fix a recursive enumeration of the partial recursive functions. Let φ_e be the e th partial recursive function, and let W_e be the image of φ_e .

Theorem 3.2 (Kleene). For every recursive function f there is a fixed point of the map $\varphi_e \mapsto \varphi_{f(e)}$.

Proof. Define

$$\psi_u(x) = \varphi_{\varphi_u(u)}(x)$$

whenever $\varphi_u(u)$ and $\varphi_{\varphi_u(u)}(x)$ are defined. Now let $t(u)$ be the index for ψ_u , i.e.

$$\varphi_{t(u)} = \varphi_{\varphi_u(u)}.$$

Since ψ_u is partial recursive, $t(u)$ must exist. Then $f \circ t$ is recursive, say $f \circ t = \varphi_v$. Then

$$\varphi_{t(v)} = \varphi_{\varphi_v(v)} = \varphi_{f \circ t(v)} = \varphi_{f(t(v))}.$$

So $t(v)$ is a fixed point. □

Kleene's recursion theorem is highly mysterious, essentially the same as the Goedel fixed-point theorem. Most of the time, the fixed point e is just an index for the program that loops forever on all inputs. Intuitively, $t(v)$ says “I want f of myself to talk about me.”

Though its behavior is mysterious, Kleene's recursion theorem is constructive, because the map h which sends e to a fixed point to φ_e , i.e.

$$\varphi_{\varphi_e(h(e))} = \varphi_{h(e)}$$

is even recursive. This follows because the fixed point for φ_e was obtained by an explicit syntactic manipulation.

Corollary 3.3. There is an e such that $W_e = \{e\}$.

Proof. Let f be such that $\varphi_{f(x)}$ enumerates $\{x\}$. Then f is clearly recursive, and by Kleene's recursion theorem, we can find an e such that $\varphi_e = \varphi_{f(e)}$, so $W_e = W_{f(e)} = \{e\}$. □

Corollary 3.4. The set $\{e : W_e \neq \emptyset\}$ is not recursive.

Proof. Let $g(e)$ be the indicator function, and assume that g is recursive. Let $\varphi_{f(e)}(x)$ be such that if $g(e) = 0$ then $\varphi_{f(e)}(x) = 1$ and otherwise leave $\varphi_{f(e)}(x)$ undefined. Since we just need to wait for $\varphi_{f(e)}(x)$ to halt to define $\varphi_{f(e)}(x)$, $\varphi_{f(e)}$ is partial recursive, so $f(e)$ exists and is even recursive. By Kleene's recursion theorem we may assume $\varphi_e = \varphi_{f(e)}$. But then $g(e) = 0$ iff $g(e) = 1$, which is absurd. \square

3.2 Rice's theorem

Definition 3.5. The *index set* for a set of r.e. reals C is the set $P(C) = \{e : W_e \in C\}$.

Example 3.6. \mathbb{N} and \emptyset are recursive index sets. Moreover, $\{e : |W_e| < \infty\}$ and $\{e : W_e \text{ is creative}\}$ are index sets, which, as we will see, are not recursive.

We can think of an index set as a set of programs which compute some semantic property. We can also view the complexity of a semantic property in terms of the complexity of the index set.

Theorem 3.7 (Rice). The only recursive index sets are \mathbb{N} and \emptyset .

In other words, there are no recursive nontrivial semantic properties.

Proof. Suppose P_C is neither \emptyset or \mathbb{N} . Let $W_{e_1} \in C, W_{e_2} \notin C$. Assume P_C is recursive. Let f be the recursive function such that if $e \in P_C$, then $f(e) = e_2$, and if $e \notin P_C$, then $f(e) = e_1$. By Kleene's recursion theorem, there is an e such that $W_e = W_{f(e)}$. Therefore $e \in P_C$ iff $e \notin P_C$, which is impossible. \square

The point that semantic properties are not recursive can be made more strongly. Let D_e be the finite set coded by e .

Theorem 3.8 (Rice-Shapiro). If P_C is an r.e. index set and C is nonempty, then there is a recursive function f such that $C = \{W_e : \exists x (D_{f(x)} \subseteq W_e)\}$.

Proof.

Lemma 3.9. If $A \subseteq B$ and $A \in C$, then $B \in C$.

Proof of lemma. Let f be the recursive function such that $W_{f(e)}$ is enumerated according to the following algorithm: While e has not been enumerated into P_C , enumerate A into $W_{f(e)}$; once e has been enumerated into P_C , then enumerate B into $W_{f(e)}$. Since $A \subseteq B$, nothing done in the first loop interferes with the second loop. So if $e \in P_C$, $W_{f(e)} = B$, and otherwise $W_{f(e)} = A$.

Let e be such that $W_e = W_{f(e)}$. If $e \notin P_C$, then $W_e = A$, so $e \in P_C$, a contradiction. So $e \in P_C$, so $W_e = B$, and so $B \in C$. \square

Lemma 3.10. If $A \in C$ then there is a finite set $D \subseteq A$ such that $D \in C$.

Proof of lemma. Define f to be recursive and such that $W_{f(e)}$ is enumerated by: if e has not been enumerated into P_C , enumerate A into $W_{f(e)}$; once e has been enumerated, halt. By a similar reasoning as in the previous lemma, $e \in P_C$ and $W_e \subseteq A$, yet W_e is finite. \square

Now let g be such that $W_{g(x)} = D_x$ and let f enumerate $\{x : g(x) \in P_C\}$. \square

Chapter 4

Semantic properties of arithmetic

4.1 Undecidability in Ramsey theory

Most examples of statements that are undecided by **PA** can be proven undecidable in less stupid ways than the proof that $\mathbf{PA} \not\vdash \text{Con PA}$. We give an example.

Definition 4.1. Let $m \rightarrow (h)_k^e$ be the sentence which says that for every coloring of the set of all subsets of $\{1, \dots, m\}$ of cardinality e with k colors, there is a $H \subseteq \{1, \dots, m\}$ such that every subset of H of cardinality e is monochromatic, and $\text{card } H \geq k$.

Theorem 4.2 (Ramsey). For every e, k, h , there is a m such that $m \rightarrow (h)_k^e$.

This result was a generalization of the pigeonhole principle.

Definition 4.3. Let $m \rightarrow^* (h)_k^e$ be the sentence which says that for every coloring of the set of all subsets of $\{1, \dots, m\}$ of cardinality e with k colors, there is a $H \subseteq \{1, \dots, m\}$ such that every subset of H of cardinality e is monochromatic, and $\text{card } H \geq \max(k, \min H)$.

Definition 4.4. The *Paris-Harrington principle* is the sentence $\forall e \forall k \forall h \exists m (m \rightarrow^* (h)_k^e)$.

Theorem 4.5 (Paris-Harrington). The Paris-Harrington principle is true in ω and not provable from **PA**.

The idea behind the proof of truth in ω is that $\aleph_0 \rightarrow (\aleph_0)_k^e$ and hence $\aleph_0 \rightarrow^* (\aleph_0)_k^e$, which follows by the existence of nonprincipal ultrafilters on ω ; one then uses Koenig's lemma, or Tychonoff's theorem, to drop this to a statement to finite sets. However, the Paris-Harrington principle implies that **PA** is consistent.

4.2 Overspill

Fix $M \models \mathbf{PA}$ and suppose that M is nonstandard. Recall that ω is an initial segment of M , which is not definable in M .

Definition 4.6. A *cut* $I \subseteq M$ is a nonempty initial segment of M such that for every $x \in I$, $x + 1 \in I$.

In particular, $\omega \subseteq I$. Let \mathbf{PA}^- be the axioms of \mathbf{PA} except for induction; then a model $N \models \mathbf{PA}^-$ is a model of \mathbf{PA} iff N does not have a definable proper cut. Therefore we have the following pair of results.

Theorem 4.7 (overspill). Let $I \subset M$ be a proper cut and $a \in M^k$ a parameter vector. If for every $b \in I$, $M \models \varphi(b, a)$, then there is a $c \notin I$ such that $M \models \varphi(c, a)$.

This is clear; if not, I would be definable by φ .

Theorem 4.8 (underspill). Let $I \subset M$ be a proper cut and $a \in M^k$ a parameter vector. If for every $c \notin I$, $M \models \varphi(c, a)$, then there is a $b \in M$ such that $M \models \forall x \geq b \varphi(x, a)$.

As a consequence, we show that while model theory cannot distinguish between models of \mathbf{PA} , recursion theory can.

Theorem 4.9 (Tennenbaum). Suppose that M is countable and $+, \times$ are computable. Then $M \cong \mathbb{N}$.

To prove Tennenbaum's theorem we will need the following lemma, which shows that the weak Koenig's lemma is not constructive.

Lemma 4.10. There is a computable, infinite binary tree T such that no infinite path through T is computable.

Proof. Let $\{\varphi_i\}_i$ be an enumeration of sentences in the language $\{0, 1, +, \times\}$. If σ is a finite binary sequence, then σ encodes the sentences φ_i such that $\sigma_i = 1$ and $\neg\varphi_i$ such that $\sigma_i = 0$. Let us say that σ is consistent with \mathbf{PA} if there is no proof of $0 = 1$ from the sentences encoded by σ and the first $|\sigma|$ axioms of \mathbf{PA} . Then it is computable to check whether σ is consistent with \mathbf{PA} , so let T be the set of all σ such that σ is consistent with \mathbf{PA} . One can easily check that T is a computable, infinite binary tree.

Let x be a path through T ; then the set of all initial segments of x encodes a complete, consistent extension of \mathbf{PA} , which is not recursive by Goedel's first incompleteness theorem. \square

Proof of Tennenbaum's theorem. Suppose $M \not\cong \mathbb{N}$. Let T^M be the subtree of $2^{<M}$ whose definition agrees with the definition of T (so if M were actually \mathbb{N} then $T^M = T$).

If σ is actually a finite binary sequence then $\sigma \in T$ if $\sigma \in T^M$. Moreover, for every $n \in \mathbb{N}$ there is a $\sigma \in T^M$ such that $|\sigma| = n$. So by overspill, there is a $\sigma^* \in T^M$ such that $|\sigma^*|$ is nonstandard.

If $+, \times$ were computable, then σ^* would be computable, yet restricts to an infinite path through T . \square

4.3 Standard systems

Let M be a nonstandard model.

Definition 4.11. A set $S \subseteq M$ is an M -finite set if there are $p, t \in M$ such that the sequence coded by (p, t) according to the β -lemma encodes S .

Definition 4.12. The *standard system* of M is the set of all $X \in 2^\omega$ such that there is an M -finite set S where $S \cap \omega = X$.

Definition 4.13. A *Scott set* is a nonempty set $S \subseteq 2^\omega$ such that if $Y_1, \dots, Y_k \in S$ and X is computable from Y_1, \dots, Y_k then $X \in S$; and if T is an infinite binary tree computed from Y_1, \dots, Y_k , then T has an infinite path in S .

So in other words, Scott sets are closed under computability and under weak Koenig's lemma.

Theorem 4.14 (Scott). Every standard system is a Scott set. Conversely, for every Scott set S of cardinality $\leq \aleph_1$, there is a nonstandard model M such that S is the standard system of M .

Clearly the restriction on cardinality can be dropped if the continuum hypothesis holds. Whether the restriction can be dropped in general is not known.

4.4 The arithmetical hierarchy

We now consider the complexity of a real number (i.e. a set of natural numbers) in terms of the complexity of the formulae that define it.

Recall that if R is a computable relation, then **PA** interprets R as a relation definable using only bounded quantifiers.

Definition 4.15. Let Σ_0^0 be the set of all formulae (logically equivalent to formulae) with only bounded quantifiers. For every n , define sets by induction:

1. Σ_n^0 is the set of all formulae of the form $\exists x_{i_1} \dots \exists x_{i_k} \psi$ where $k \geq 0$ and $\psi \in \Pi_{n-1}^0$.
2. Π_n^0 is the set of all formulae of the form $\neg\psi$ where $\psi \in \Sigma_n^0$.
3. $\Delta_n^0 = \Sigma_n^0 \cap \Pi_n^0$.

In particular, $\Sigma_0^0 = \Pi_0^0 = \Delta_0^0$ is the set of formulae defining recursive relations. We will abuse notation and say that a real is Σ_n^0 if it is definable by a Σ_n^0 formula. So a recursive real is Σ_0^0 .

Lemma 4.16. For every Σ_n^0 formula φ and bounded quantifier $\forall x < \tau$, there is a Σ_n^0 formula ψ such that

$$\omega \models (\forall x < \tau \varphi) \leftrightarrow \psi.$$

Proof. We induct on n . The case $n = 0$ is obvious. So suppose the claim holds for $n - 1$. Let

$$\varphi = \exists y_1 \dots \exists y_k \chi$$

where χ is Π_{n-1}^0 .

The formula $\forall x < \tau \varphi$ holds iff there is a y which codes a sequence of length $k\tau$ such that

$$\forall x < \tau \psi(x, y_{xk}, y_{xk+1}, \dots, y_{xk+k-1}).$$

Now this formula, say θ , has a bounded quantifier followed by a Π_{n-1}^0 formula. By induction, θ is logically equivalent to $\exists y \theta'$ where θ' is Π_{n-1}^0 . Now let $\psi = \exists y \theta'$. \square

In particular, if φ is Σ_0^n , then $\forall x < \tau \varphi$ is also Σ_0^n . We also have the following theorem, left to an exercise.

Theorem 4.17. Suppose that φ is Σ_0^n . Then there is a Σ_0^0 formula ψ such that φ is logically equivalent to

$$\exists x_1 \forall x_2 \exists x_3 \cdots Q_n x_n \psi$$

where $Q_n = \exists$ if n is odd and $Q_n = \forall$ if n is even.

Definition 4.18. If θ is a formula, let $[\theta]$ be the *Goedel number* of θ , the number which codes θ .

Definition 4.19. Let S_n^* be the set of all (θ, a) such that a codes a vector \vec{a} , $\mathbb{N} \models (\theta, \vec{a})$, and θ is Σ_n^0 . Let S_n be the set of all $([\theta], a)$ with this property. We call S_n the *universal Σ_n^0 set*.

Theorem 4.20. The universal Σ_n^0 set is Σ_n^0 .

Proof. By a previous theorem, for every $\theta \in \Sigma_n^0$ there is a θ' with bounded quantifiers such that

$$\theta = \exists x_1 \forall x_2 \exists x_3 \cdots Q_n x_n \theta'$$

and the map $[\theta] \mapsto [\theta']$ is recursive. Let γ represent that map; then γ is Σ_1^0 . Then $([\theta], a) \in S_n$ iff

$$\exists x_1 \forall x_2 \cdots Q_n x_n \theta'(\vec{x}, \vec{a})$$

which happens iff $\exists x_1 \forall x_2 \cdots Q_n x_n$ such that $([\theta'], y) \in S_0$ where $\vec{y} = (\vec{x}, \vec{a})$, which happens iff $\exists w$ such that $\gamma([\theta], w)$ and $\exists x_1 \forall x_2 \cdots Q_n x_n$ such that $(w, y) \in S_0$ and $\vec{y} = (\vec{x}, \vec{a})$. The statement $\vec{y} = (\vec{x}, \vec{a})$ is bounded. Since $\gamma([\theta], w)$ is Σ_1^0 , the statement “ $\exists w$ such that $\gamma([\theta], w)$ and $\exists x_1 \forall x_2 \cdots Q_n x_n$ such that $(w, y) \in S_0$ and $\vec{y} = (\vec{x}, \vec{a})$ ” is Σ_n^0 . \square

We now show that S_n is not Σ_{n-1}^0 .

Theorem 4.21 (Kleene). S_n is not Π_n^0 .

Proof. We use the diagonal argument. Suppose S_n was defined by a Π_n^0 formula φ . Define $\psi(x)$ to be the formula $\neg\neg\varphi(x, [(x)])$, which is Σ_n^0 . Then $\psi([\psi])$ iff $\neg\varphi([\psi], [([\psi])])$. But φ defines S_n , so $\psi([\psi])$ iff $\varphi([\psi], [([\psi])])$, a contradiction. \square

4.5 Turing degrees

Definition 4.22. Let X, Y be reals. Let $X \leq_T Y$ mean that Y is computable relative to X .

One easily checks that \leq_T is a preorder. Let $X \equiv_T Y$ mean that $X \leq_T Y$ and $Y \leq_T X$; then \equiv_T is an equivalence relation. The equivalence classes under \equiv_T are known as *Turing degrees*; then \leq_T drops to a partial order on the Turing degrees.

Given Turing degrees X, Y let $X \oplus Y$ be the supremum of $\{X, Y\}$ under \leq_T . If we choose representatives, then in fact $X \oplus Y$ is the equivalence class of $\{2n : n \in X\} \cup \{2n+1 : n \in Y\}$.

Definition 4.23. Given a real X , let X' be the *Turing jump* of X , the set of m such that the m th program relative to X halts on empty input. Let $X^{(n)}$ denote the n th iterate of the Turing jump of X .

By a straight diagonalization we see that $X <_T X'$. Of course X' can compute X , and for the opposite, we consider that the problem of computing X from X' is the halting problem. Thus given X , we obtain an infinite chain of Turing degrees above X . We let 0 denote the minimal Turing degree. Let \mathcal{D} denote the set of Turing degrees. The equivalence class of any set is countable, so \mathcal{D} has cardinality 2^{\aleph_0} .

Theorem 4.24 (Post). For every n and every real X , X is Σ_{n+1}^0 iff X is r.e. relative to $0^{(n)}$.

Proof. Proof by induction. We already proved the case $n = 0$. So assume that the theorem is true for n .

Suppose that X is r.e. relative to $0^{(n+1)}$. Since $0^{(n+1)}$ is r.e. relative to $0^{(n)}$, by induction, $0^{(n+1)}$ is Σ_{n+1}^0 . Let φ define $0^{(n+1)}$, so φ is Σ_{n+1}^0 . Let P be the program which enumerates X . Then $m \in X$ iff there is a c such that c codes a computation of P with output m such that whenever $x \in 0^{(n+1)}$, $\varphi(x)$, and else $\neg\varphi(x)$. This is Σ_{n+2}^0 since an “if and only if” statement involving φ is Δ_{n+2}^0 , c is unbounded, and the predicate “codes a computation” is Σ_0^0 .

Conversely, if X is Σ_{n+2}^0 , suppose that $m \in X$ iff $\exists w \varphi(w, m)$ and φ is Π_{n+1}^0 . Then φ defines a set of pairs which is co-r.e. relative to $0^{(n)}$. Let $P = \{(w, y) : \omega \models \varphi(w, y)\}$. Then $P \leq_T 0^{(n+1)}$, so $0^{(n+1)}$ can enumerate $m \in X$ by searching for a w such that $(w, m) \in P$. \square

Now if P is a preorder of cardinality at most 2^{\aleph_0} such that for every $p \in P$, $\{q \in P : q > p\}$ is countable, Sacks asked if there was an embedding $P \rightarrow \mathcal{D}$. This is still an open question.

4.6 The arithmetical hierarchy in **PA**

We now formalize the above results inside **PA**. Note that here the superscript 0 in the symbol Σ_n^0 , which is meant to indicate that we are working in first-order logic, is irrelevant, since **PA** is already first-order. So we drop the superscripts for the time being.

Recall that if a is a number we let \vec{a} be the vector coded by a .

Theorem 4.25. For every $n \geq 1$ and $Q \in \{\Sigma, \Pi, \Delta\}$ there is a formula Sat_{Q_n} which is Q_n such that for every Q_n formulae φ ,

$$\mathbf{PA} \vdash \forall a \varphi(\vec{a}) \leftrightarrow \text{Sat}_{Q_n}([\varphi], a).$$

The proof is in Chapter 9 of Kaye’s book “Models of Peano Arithmetic”, but it is quite tedious. We sketch the ideas of the proof:

1. The vector \vec{a} can be coded by a number a using the β -lemma and the J -pairing function.
2. There is a definable bijection between logical symbols and numbers, so a bijection between formulae and vectors. We will call the number a such that \vec{a} maps to a given formula ψ the “code” of ψ , and write $a = [\psi]$.
3. The formula-building operations are definable operations on codes.
4. The relation “ x_1, \dots, x_n are codes for symbols that make up a formula coded by y ” is a definable relation $R(x_1, \dots, x_n, y)$.

5. If x is the Goedel number of a term τ with n free variables and \vec{y} has n entries, let $V(x, y)$ be the evaluation of $\tau(\vec{y})$.
6. There is a formula γ such that

$$\mathbf{PA} \vdash \forall y_1 \dots \forall y_n \forall y ((\forall i < n (\vec{y}_i = y_i)) \rightarrow (\tau(\vec{y}) = \gamma([\tau], y))).$$

So \mathbf{PA} proves that γ correctly represents V . This takes infinitely many sentences to state, one for each τ .

7. We now are in the position to define satisfaction for Δ_0 formulae. Let $S(s, x)$ mean that:
 - (a) s is the code for a Δ_0 formula,
 - (b) x is the code for a sequence of triples (i, z, w) where i is a formula in the sequence coded by s and w is the truth value of that formula when it is evaluated at z , and
 - (c) For each k less than the length of the sequence coded by x , if \vec{x}_k is $([\exists y_\ell < \tau\theta], z, w)$ then for each $u < V([\tau], z)$ there is a $j < k$ such that \vec{x}_j is $([\theta], z^*, w)$, $z_\ell^* = u$, and w is the truth value.
8. Let $|x|$ be the length of \vec{x} , which is a first-order term in x and let 1 be the code for the truth value “True.”

9. Let

$$\text{Sat}_{\Sigma_0}(x, y) = \exists s \exists t S(s, t) \wedge \exists j < |s| (\vec{s}_j = x) \wedge \exists k < |t| (\vec{t}_k = (j, y, 1)).$$

Here x codes the formula and y codes the free variables.

10. Sat_{Σ_0} is clearly Σ_1 , but either there is a witness that x is satisfied or there is a witness that x is not satisfied, so in fact $\mathbf{PA} \vdash \text{Sat}_{\Sigma_0} \in \Delta_1$.
11. By induction, $\mathbf{PA} \vdash \text{Sat}_{\Sigma_0}([\varphi], y) \leftrightarrow \varphi(\vec{y})$.
12. One then inductively defines Sat_{Σ_n} inductively by adding existential quantifiers to the front of $\text{Sat}_{\Sigma_{n-1}}$.

4.7 Submodels of arithmetic

We prepare for the proof that \mathbf{PA} does not have a finite axiomatization.

Definition 4.26. Let M be a first-order structure and $A \subseteq M$. Let $K^n(M, A)$ be the substructure of all $x \in M$ such that x is definable from parameters in A by a Σ_n formula.

Let $N \preceq_{\Sigma_n} M$ mean that a Σ_n formula holds in N iff it holds in M . So N is an elementary substructure if $N \preceq_{\Sigma_n} M$ for all n .

Theorem 4.27. If M is a nonstandard model of arithmetic then $K^n(M, A) \preceq_{\Sigma_n} M$.

To prove this theorem, we need the theory of the following axiom.

Axiom 4.28 (bounding scheme). For every φ we have

$$\forall a(\forall x < a \exists w \varphi(x, w, a) \rightarrow \exists b \forall x < a \exists w < b \varphi(x, w, a)).$$

In other words, every definable function on a finite set has a bounded image.

We let B denote the bounding scheme. Let $B\Sigma_n$ be the bounding scheme for Σ_n formulae; let $I\Sigma_n$ be the induction scheme for Σ_n formulae.

Theorem 4.29. In $\mathbf{PA}^- + I\Sigma_0$ we have $I\Sigma_n$ if and only if we have III_n . In fact, $\mathbf{PA}^- + I\Sigma_0 + B$ iff \mathbf{PA} .

In $\mathbf{PA}^- + I\Sigma_0$ we have $I\Sigma_n \implies B\Sigma_n \implies I\Sigma_{n-1}$.

Proof that $I\Sigma_n$ implies III_n . Suppose $n \geq 1$ and $\mathbf{PA}^- + I\Sigma_n$. Let φ be Π_n and suppose $\forall x(\varphi(x) \rightarrow \varphi(x+1))$. Suppose $\neg\varphi(a)$, which is Σ_n . Let $\psi(x) = (x > a) \vee \neg\varphi(a-x)$. Then $\psi(0) \wedge \forall x(\psi(x) \rightarrow \psi(x+1))$. So $\forall x\psi(x)$, so $\neg\varphi(0)$. Therefore induction holds for φ . \square

The proof that III_n implies $I\Sigma_n$ is similar.

Proof that $I\Sigma_0 + B\Sigma_{n+1}$ implies $I\Sigma_n$. We prove this by induction on n . The base case is given. Assume $I\Sigma_n$ and let $\varphi(x) = \exists w \theta(x, w)$ is Σ_{n+1} . Assume $\varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(x+1))$. We will show that $\forall a \forall x < a \varphi(x)$.

Suppose not, so there is an a such that a is larger than a counterexample to φ . Then

$$\forall x < a \exists w(\theta(x, w) \vee (\neg\varphi(x) \wedge (w = 0))).$$

By B , there is a b such that

$$\forall x < a \exists w < b(\theta(x, w) \vee (\neg\varphi(x) \wedge (w = 0))).$$

Let

$$\varphi'(x) = \exists w < b(\theta(x, w) \vee x \geq a).$$

Then $\varphi'(x)$ is Π_n and $\varphi'(0) \wedge \forall x(\varphi'(x) \rightarrow \varphi'(x+1))$. Since $I\Sigma_n$ implies III_n , $\forall x \varphi'(x)$, contradicting the definition of a . \square

Proof that $I\Sigma_n$ implies $B\Sigma_n$. It suffices to prove $B\Pi_n$. Suppose that $\varphi \in \Pi_n$ and

$$\forall x < a \exists w \varphi(x, w, a).$$

Let

$$\psi(z) = \exists b \forall x < z \exists w < b \varphi(x, w, a) \vee z > a.$$

Then $\psi(0)$. If $\psi(z)$ is true, then if $z \geq a$, $z+1 > a$ so $\psi(z+1)$. Otherwise, $z < a$ so $\exists w \varphi(z, w, a)$. Let b' be the largest b that works for $\psi(z)$ and $w+1$ where $\varphi(z, w, a)$. So

$$\forall x < z+1 \exists w < b' \varphi(x, w, a)$$

so b' is a witness to $\psi(z+1)$. Therefore $I\Sigma_{n+1}$ implies $\forall z \psi(z)$. But $\psi(a)$ implies B for φ . \square

The next thing we need is the following relativized variant of the Tarski-Vaught test.

Lemma 4.30. Assume M_1 is a substructure of M_2 . Then $M_1 \preceq_{\Sigma_n} M_2$ iff for every $m_1, \dots, m_n \in M_1$ and Σ_n formula $\exists x\varphi$ such that $M_2 \models \exists x\varphi$, there is a $m \in M_1$ such that $M_2 \models \varphi(m, m_1, \dots, m_n)$.

This follows immediately from the usual proof of the Tarski-Vaught test. It can be interpreted to mean that every nonempty $\Sigma_n(M_1)$ -subset meets Σ_n .

Proof of Theorem 4.27. One first checks $K^n(M, A)$ is a substructure. In fact if x_1, x_2 are Σ_n -definable, say by $\theta_{x_1}, \theta_{x_2}$, then $\exists y_1 \exists y_2 \theta_{x_1}(y_1) \wedge \theta_{x_2}(y_2) \wedge (z = y_1 + y_2)$ is a definition (with free variable z) of $x_1 + x_2$, which is Σ_n . Similarly for multiplication and A .

Now suppose $\exists w\varphi(x, w)$ is a Σ_n formula and $M \models \exists x \exists w \varphi(x, w, b_1, \dots, b_m)$ where the $b_i \in K^n(M, A)$, say $\psi_i(b_i)$. Then

$$M \models \exists x \exists w \exists \vec{y} \left(\bigwedge_i \psi_i(y_i) \wedge \varphi(x, w, \vec{y}) \wedge \forall x^* < x \neg \exists w^* \leq w \varphi(x^*, w^*, \vec{y}) \right);$$

by the definable well-ordering principle. The sequence (x, w, \vec{y}) is unique and the formula is in Σ_n , so (x, w, \vec{y}) consists of elements of $K^n(M, A)$. Thus the claim follows by the relativized Tarski-Vaught test. \square

Theorem 4.31. For every nonstandard model M and n , there is a $K \preceq_{\Sigma_n} M$ such that $K \not\models \mathbf{PA}$.

Proof. Let $a \in M$ be nonstandard, $K_0 = \{a\}$. Let K_{n+1} be the M -finite set containing M_i and all z such that there is an $x < a$ and a y such that y codes a sequence of m elements from M_i , z is unique, $\text{Sat}_{\Sigma_n}(x, (\vec{y}, z))$. So K_{n+1} consists of the things that M thinks can be definable from the first a Σ_n formulae and parameters from K_n . This is why M thinks that K_{n+1} is finite. Let K be the limit of the K_n so $K^n(M, K) = K$, so $K \preceq_{\Sigma_n} M$.

Then K_n is definable in K (relative to finite subsets of K), since it is definable in M and K and M agree on existence and uniqueness for solutions to Sat_{Σ_n} . Therefore $n \in \mathbb{N}$ iff K thinks there is a code for K_n , so \mathbb{N} is definable in K . (In M you could run this construction for any M -finite length, so if m is nonstandard then K_m exists. But this is impossible in K , since K consists of only the things we constructed at a finite stage.) Therefore K does not model \mathbf{PA} . \square

Theorem 4.32 (Ryll-Nardzewski). \mathbf{PA} does not have a consistent, finitely axiomatizable extension.

Proof. If T is such a theory, let n be so large that every axiom of T is Σ_n . This is possible since T is finitely axiomatizable. Let $M \models T$ be nonstandard and let $K \preceq_{\Sigma_n} M$ not model \mathbf{PA} . Then $K \models T$, a contradiction. \square

Originally we developed the theory of Sat_{Σ_n} to show that the arithmetic hierarchy doesn't collapse at any finite stage. However, the Ryll-Nardzewski theorem is a nice application of this theory that isn't just a statement about the arithmetic hierarchy.

4.8 Slaman's theorem

Axiom 4.33 ($I\Delta_n$). For all Σ_n formulas φ, ψ ,

$$\forall x(\varphi(x) \leftrightarrow \neg\psi(x)) \rightarrow (\varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(x+1))) \rightarrow \forall x\varphi(x).$$

Axiom 4.34 (EXP). $\forall x\forall y\exists z(x^y = z)$.

Theorem 4.35 (Slaman). Work in $\mathbf{PA}^- + I\Sigma_0 + EXP$. For all $n \geq 1$, $I\Delta_n$ iff $B\Sigma_n$.

The clause $x^y = z$ is a definable, Σ_0 relation since exponentiation is recursive. So EXP is a Π_2 formula, and in fact follows from $I\Sigma_1$. Thus the very similar theorem that says that if $\mathbf{PA}^- + I\Sigma_1$ implies that for all $n \geq 2$, $I\Delta_n \Leftrightarrow B\Sigma_n$ is accurate. It is an open problem to remove the assumption on EXP in the case $n = 1$.

The idea behind the trick in the proof of $B\Sigma_n \implies I\Delta_n$ is that if we knew the supremum $M(N)$ of the runtimes of all halting programs in the first N programs, we can decide $0' \cap \{0, \dots, N-1\}$ by running each program until time $M(N)$ and seeing if it has halted yet.

Proof of $B\Sigma_n \implies I\Delta_n$. Let φ, ψ be Σ_n and suppose that φ and $\neg\psi$ define the set J from parameters p . We must show that J satisfies $I\Delta_n$.

Suppose $\varphi = \exists y\varphi_0$ and similarly for ψ . Suppose that a is strictly above an element of J^c . Since $\varphi \vee \psi$ is always true,

$$M \models \forall x < a \exists y\varphi_0(x, y, p) \vee \psi_0(x, y, p)$$

so by $B\Sigma_n$, there is a $b \in M$ such that

$$M \models \forall x < a \exists y < b \varphi_0(x, y, p) \vee \psi_0(x, y, p)$$

Then, under a , J is defined by

$$\exists y < b \varphi_0(x, y, p).$$

But $B\Sigma_n$ allows us to absorb the bounded quantifier into φ_0 , so J is defined by a Π_{n-1} formula. Now $B\Sigma_n \implies I\Pi_{n-1}$, so if J satisfies the hypotheses of the principle of induction, a does not exist, a contradiction. Therefore J cannot be a counterexample of $I\Delta_n$. \square

Now recall that a function is cofinal if its image is unbounded. To show that $I\Delta_n \implies B\Sigma_n$, we first construct a function that explicitly demonstrates the failure of $B\Sigma_1$, if it is actually false.

Lemma 4.36. Let $M \models \mathbf{PA}^- + I\Delta_1 + EXP + \neg B\Sigma_1$. Then there is a $a \in M$ and $f : [0, a) \rightarrow M$ such that f is injective, f is cofinal, and f is Σ_0 relative to M .

Proof. Let $a \in M$, $p \in M$, and $\varphi = \exists w\varphi_0$, where φ_0 is Σ_0 , and

$$M \models \forall x < a \exists y \exists w \varphi_0(x, y, w, p)$$

and

$$M \models \forall s \exists x < a \exists y < s \exists w < s \neg \varphi_0(x, y, w, p).$$

This is what it means for $M \not\models B\Sigma_1$.

Given $x < a$, let $f(x)$ be the code of (x, s_x) where s_x is the least s such that

$$M \models \exists y < s \exists w < s \varphi_0(x, y, w, p).$$

The existence of s follows from

$$M \models \forall x < a \exists y \exists w \varphi_0(x, y, w, p)$$

and the existence of s_x follows from $I\Delta_1$.

Now x is determined by the first coordinate of $f(x)$ so f is injective, and because

$$M \models \forall s \exists x < a \exists y < s \exists w < s \neg \varphi_0(x, y, w, p),$$

f is cofinal. Moreover, the definition of f is Σ_0 . □

Lemma 4.37. Let $M \models \mathbf{PA}^- + I\Delta_1 + EXP + \neg B\Sigma_1$. Then there is a $a \in M$, a nonprincipal Σ_1 -cut $I \subset [0, a)$, and a $g : I \rightarrow M$ such that g is Σ_1 from M , and:

1. For every $i \in I$, $g(i)$ is the code for a sequence $\{m_j : j < i\}$ such that $j \mapsto m_j$ is injective and $m_j < a$.
2. For every $i_1 < i_2 \in I$, $g(i_2)$ is an end-extension of $g(i_1)$.
3. For every $m < a$, there is an $i \in I$ such that m appears in $g(i)$.

So the sequence $g(i)$ is an enumeration of i many numbers under a , and $g(\omega)$ is an enumeration of the numbers under a . Recall that a cut is nonprincipal if it has no greatest element.

Proof. Let (f, a) be given by the previous lemma. Then $[0, a)$ is M -finite, so is M -r.e. Then m is enumerated into this set at $f(m)$. So let g satisfy the claimed properties, and let the new entry m in the sequence $g(s)$ at stage s be the m such that $f(m) = s$, which is possible since f is injective. Now let $I = \text{dom } g$. In order to enumerate $s + 1$, we must have already enumerated s , so I is an initial segment of M .

If $i > a$ and $i \in I$ then we have enumerated a many numbers less than a , yet $\mathbf{PA}^- + I\Delta_1 + EXP$ proves the Σ_0 -pigeonhole principle, so this is absurd. So $I \subset [0, a)$.

Since f is cofinal and I has the same ordertype as $f([0, a))$ it follows that I has no greatest element. So I is a proper nonprincipal cut, and is Σ_1 since it is the domain of the Σ_1 function g .

Finally, if $m < a$ and $i \in I$ large enough, m appears in $g(i)$, since $m \in \text{dom } f$. □

The cut I above is a counterexample to $I\Sigma_1$ but we actually want it to be a counterexample to $I\Delta_1$. Moreover, M thinks that m_i has no finite end extension, by the following lemma.

Lemma 4.38. Let a and m_i be as in the previous lemma. Let $c \in M$, n a sequence of length c of elements $< a$. Then either c is not an upper bound for I or there is an $i \in I$ such that $n_i \neq m_i$.

Proof. Suppose c is an upper bound on I . If $m_i = n_i$ for all $i \in I$, let

$$J = \{j < c : \forall i < j \neg(n_i = n_j)\}.$$

Then J is Σ_0 , and if we can show that J is a cut, then J will be a counterexample to $I\Delta_1$, which is a contradiction.

Given $i \in I$, $g(i)$ is an injective sequence, so $I \subseteq J$. But if $x < a$ then $x = m_i$ for some $i \in I$, so it follows that $J \subseteq I$. But I is a proper cut. \square

Proof of $I\Delta_1 + EXP \implies B\Sigma_1$. Let $M \models \mathbf{PA}^- + I\Delta_1 + EXP + \neg B\Sigma_1$. Let I, g, m be as above.

Let $c_0 = 0$ and $d_0 = a^a$. Let $c_i = \sum_{j < i} m_j a^{-(j+1)}$ and $d_i = c_i + a^{a-i}$. Then $[c_{i+1}, d_{i+1}] \subseteq [c_i, d_i]$. Viewing elements of M as in base a , the i th interval specifies the i th digit of the number in $\bigcap_i [c_i, d_i]$. Let $J = \{x : \exists i (x \leq c_i)\}$ and $K = \{x : \exists i (x \geq d_i)\}$. Then J, K are Σ_1 , since the intervals are. Moreover J is closed downwards, K is closed upwards, and J has no greatest element (because neither does I , so that the c_i is eventually a strictly increasing sequence). Clearly $J \cap K$ is empty.

We claim $M = J \cup K$. If not, there is a $n \in \bigcap_i [c_i, d_i]$. Then the base- a representation of n is given by the entries of m . Therefore n can be computed from m , so by $I\Delta_1$, the entries of n are coded by M , which contradicts a previous lemma.

So the decomposition $M = J + K$ implies that J is a Δ_1 -cut, so by $I\Delta_1$, $J = M$, a contradiction since K is nonempty. \square

Conjecture 4.39 (Paris-Wilkie). $\mathbf{PA}^- + I\Sigma_0 + \neg B\Sigma_1 \vdash EXP$.

This problem has been open for three decades. If the Paris-Wilkie conjecture were true we could remove the assumption on EXP from Slaman's theorem.

Chapter 5

The Turing degrees

5.1 Myhill's theorem

Definition 5.1. Let X, Y be reals. Say that $Y \leq_m X$, i.e. Y is *many-one reducible* to X , if there is a recursive function h such that $h(X) = Y$ and $h^{-1}(Y) = X$. Say that $Y \leq_1 X$, i.e. Y is *one-one reducible* to X , if there is an injective recursive function h such that $h(X) = Y$ and $h^{-1}(Y) = X$.

Intuitively a many-one reduction h is a recursive homomorphism $X \rightarrow Y$.

Definition 5.2. Two reals X, Y are *recursively isomorphic* if there is a recursive bijection $h : \omega \rightarrow \omega$ such that $h(X) = Y$ and $h^{-1}(Y) = X$.

We now prove a recursive version of the Schroeder-Bernstein theorem.

Theorem 5.3 (Myhill). If $X \leq_1 Y$ and $Y \leq_1 X$ then X, Y are recursively isomorphic.

Proof. Suppose $f(X) = Y$ and $g(Y) = X$. We construct a bijection h in stages.

At stage s , if $s \notin \text{dom } h_s$, look at the sequences $x_0 = s, y_n = f(x_n), x_{n+1} = h_s^{-1}(y_n)$. Let $h_{s+1}(s)$ be the first y_i such that $y_i \notin h_s(\text{dom } h_s)$. A similar argument with h_s and h_s^{-1} swapped, and f and g swapped, allows us to also guarantee $s \in h_{s+1}(\text{dom } h_{s+1})$.

Clearly h is a bijection $\omega \rightarrow \omega$. By induction on s we see that $h(X) = Y$ and $h^{-1}(Y) = X$. \square

Corollary 5.4. Sat_{Σ_1} is recursively isomorphic to $0'$.

Proof. Here we view Sat_{Σ_1} as $\{([\theta], x)\}$ and $0' = \{(e, m)\}$ where e is a program and m is an input. Given a program e , choose θ so that $\varphi_e(m)$ exists whenever $\mathbb{N} \models \theta(x)$. This gives a reduction $\text{Sat}_{\Sigma_1} \leq 1 0'$. Arguing similarly, we have a reduction $0' \leq \text{Sat}_{\Sigma_1}$. \square

Definition 5.5. A r.e. real X is *1-complete* if for all r.e. sets W , $W \leq_1 X$. Similarly, X is *m -complete* if for all r.e. sets X , $W \leq_m X$.

By Myhill's theorem, all 1-complete sets are recursively isomorphic.

Example 5.6. Sat_{Σ_1} is 1-complete. In fact, given W , there is a Σ_1 definition θ of W , but $\theta(x)$ is equivalent to $([\theta], n) \in \text{Sat}_{\Sigma_1}$. Thus $n \mapsto ([\theta], n)$ is a 1-1 reduction.

So the 1-complete sets are exactly those which are isomorphic to Sat_{Σ_1} .

Theorem 5.7 (Matiyasevich-Robinson-Davis-Putnam). The set of polynomials with solutions in \mathbb{Z}^n is 1-complete.

The proof of this is beyond the scope of these notes but it is worth noting. It shows that Hilbert's tenth problem, which asked for an algorithm to find solutions to polynomial equations or decide that they are unsolvable, has no answer.

Example 5.8. $\text{Fin} = \{e : |W_e| < \infty\}$ is Σ_2 -complete.

Proof. $|W_e| < \infty$ iff $\exists b \forall n \forall s (n \in W_e[s] \implies n < b)$. This is clearly Σ_2 .

Let θ be bounded and $\text{Sat}_{\Sigma_2}(e, n) = \exists x \forall y \theta(e, n, x, y)$. Let f be recursive and enumerate $W_{f(e, n)}$ such that at stage s , if $k < s$, $k \notin W_{f(e)}[s]$ and $\forall x < k \exists y < s \neg \theta(e, n, x, y)$, enumerate k into $W_{f(e, n)}$. Then $W_{f(e, n)} = \{k : \forall x < k \exists y \neg \theta(e, n, x, y)\}$. So $W_{f(e, n)}$ is finite iff $\exists x \forall y \theta(e, n, x, y)$. So f reduces Sat_{Σ_2} to Fin . \square

Corollary 5.9. $\text{Tot} = \{e : W_e = \mathbb{N}\}$ is Π_2 -complete.

Example 5.10. Being a recursive convergent sequence is Π_3 -complete. More precisely, consider the recursive functions $\varphi_e : A \rightarrow \mathbb{Q}$ where $A \subseteq \mathbb{N}$. Let C be the set of e such that φ_e is total and $\lim_{n \rightarrow \infty} \varphi_e(n)$ exists. Then C is Π_3 -complete.

Proof. Whether $e \in \text{Tot}$ is Π_2 . The limit exists iff the sequence $\varphi_e(n)$ is Cauchy, which is Π_3 .

To reduce Sat_{Π_3} to C , let θ be bounded and $\text{Sat}_{\Sigma_3}(e, n) = \forall x \exists y \forall z \theta(e, n, x, y, z)$. At stage s , let $x < s$. Let $y(x, s)$ be the least $y < s$ such that $\forall z < s \theta(e, n, x, y, z)$ if one such exists, or $y = s$ otherwise. So y is the “least candidate for a witness that is viable at stage s .” Let $p(x, s)$ be the cardinality of $\{t < s : y(x, t) \neq y(x, t+1)\}$. Then $\text{Sat}_{\Pi_3}(e, n)$ iff $\forall x \exists y$ such that $\forall z \theta(e, n, x, y, z)$, which would mean that $y(x, s)$ and $p(x, s)$ are Cauchy in s , since they converge to the witness y and the number of times we changed our mind about the witness y , respectively.

Define f so that

$$\varphi_{f(e, n)}(s) = \sum_{x=1}^s \frac{(-1)^{p(x, s)}}{4^x}.$$

Then f is recursive, and if for every x $p(x, s)$ has a limit, $\varphi_{f(e, n)}(s)$ converges. But if not, there is an x_0 which is least possible and $p(x_0, s)$ is not Cauchy. Let

$$a = \lim_{s \rightarrow \infty} \sum_{x=1}^{x_0-1} \frac{(-1)^{p(x, s)}}{4^x}.$$

If $p(x, s)$ is even then $\varphi_{f(e)}(s) \geq a + 2/3 \cdot 4^{x_0}$ but otherwise $\varphi_{f(e)}(s) \leq a - 2/3 \cdot 4^{x_0}$. So f reduces Sat_{Π_3} to C . \square

So the representation of convergence as Cauchyness is optimal in quantifier complexity.

Corollary 5.11. The sets of cofinite, or recursive, r.e. sets, are Σ_3 -complete.

5.2 Completeness

Theorem 5.12. For every r.e. real A , A is m -complete iff A is 1-complete.

Proof. Let $\{D_i\}_i$ be a recursive enumeration of the finite subsets of ω . Let f be a m -reduction from Sat_{Σ_1} to A . We must find a 1-reduction from Sat_{Σ_1} to A .

Lemma 5.13. There is a recursive function g such that if $\emptyset \subset D_x \subset A$, then $g(x) \in A \setminus D_x$, and if $\emptyset \subset D_x \subset A^c$, then $g(x) \in A^c \setminus D_x$.

Proof. Let $\theta_x(y) = (f(y, y) \in D_x) \vee \exists z(z \in D_x \cap A)$. This definition makes sense because D_x is a finite subset of ω . Note if $D_x \cap A$ is nonempty then $\forall y \theta_x(y)$. The function $x \mapsto [\theta_x]$ is recursive since $x \mapsto D_x$ is. If $f([\theta_x], [\theta_x]) \in D_x$, let $g(x) = f([\theta_x], [\theta_x])$. Otherwise, let $g(x)$ be the first element enumerated into $A \setminus D_x$. Since $\text{Sat}_{\Sigma_1} \leq_m A$, A is infinite, so g is well-defined, and $g(x) \notin D_x$.

Assume that D_x is nonempty. If $D_x \subset A$, then $\mathbb{N} \models \theta_x([\theta_x])$ since $\mathbb{N} \models \forall y \theta_x(y)$. Therefore $([\theta_x], [\theta_x]) \in \text{Sat}_{\Sigma_1}$, so $f([\theta_x], [\theta_x]) \in A$. So $g(x) \in A$. Otherwise, if $D_x \subset A^c$, either $f([\theta_x], [\theta_x]) \in D_x$ or not. If so, then $\mathbb{N} \models \theta_x([\theta_x])$, so $f([\theta_x], [\theta_x]) \in A$, which contradicts the assumption. So $f([\theta_x], [\theta_x]) \notin D_x^c$ and hence $g(x) = f([\theta_x], [\theta_x])$, so $g(x) \in A^c \setminus D_x$. \square

Now assume we have an injective reduction $h : \text{Sat}_{\Sigma_1} \leq_1 A$ defined for all $z < x$. Let $y_0 = f(x)$. If $y_0 \notin \{h(z) : z < x\}$, let $h(x) = y_0$. Otherwise, induct. Let x_i be defined by $D_{x_i} = \{y_j : j \leq i\}$, and let $y_{i+1} = g(x_i)$. If $y_{i+1} \notin \{h(z) : z < x\}$, let $h(x) = y_{i+1}$. Else, continue the induction..

This defines an injective recursive function h . In fact $h(x)$ is y_i where $i \leq x$, so $h(x)$ must exist. Moreover $x \in \text{Sat}_{\Sigma_1}$ iff $y_i \in A$, by definition of g . \square

The technical lemma above is quite remarkable because A is not recursive, yet the recursive function g can decide whether $x \in A$ provided that $D_x \neq \emptyset$.

Now we will give an r.e. set which is neither recursive nor m -complete.

Definition 5.14. An r.e. set A is *simple* if A^c is infinite and A^c has no infinite r.e. subset.

Lemma 5.15. If A is simple then A is neither recursive nor m -complete.

Proof. Since A is not co-r.e., A is not recursive. If A were m -complete, then by the above theory, A would be isomorphic to Sat_{Σ_1} . But $\text{Sat}_{\Sigma_1}^c$ has an infinite recursive subset. \square

Theorem 5.16 (Post). There is a simple set.

Proof. Let $\{W_e\}_e$ be a recursive enumeration of the r.e. sets. We say that $x \in W_k[s]$ if x is enumerated into W_k by time s .

At time s , say that W_k requires *attention* if $k \leq s$ and there is a $x > 2k$ such that $x \in W_k[s]$ and no element of $W_k[s]$ has been enumerated into A before time s . If any r.e. set requires attention, let k be the minimal possible index where W_k requires attention, enumerate the least $x > 2k$ such that $x \in W_k[s]$ into A .

By definition, A is r.e. There are at most k many elements of A below $2k$, so A^c is infinite. If W_k is infinite, then there is an $x > 2k$ and s such that $x \in W_k[s]$. If no element of $W_k[s]$ has been added to A by time s , then W_k required attention at time s . So some element of W_k will have been enumerated by time $s + k + 1$. Therefore $A \cap W_k$ is nonempty. Therefore A^c does not contain W_k . \square

So there is a set A such that $0 <_m A <_m 0'$. But this not does imply that there is a set A such that $0 <_T A <_T 0'$. Post asked whether this was true.

5.3 Creative sets

Recall that W_x is the set enumerated by the x th program; that is, W_x is the image of φ_x . We may assume that φ_x is total and injective by “speeding up” an enumeration, providing that W_x is infinite.

Definition 5.17. A real P is a *productive set* if there is a partial recursive function h such that if $W_x \subseteq P$ then $x \in \text{dom } h$ and $h(x) \in P \setminus W_x$. We call h a *production*.

In other words, productive sets are far from r.e. So in particular, productive sets cannot be recursive.

Definition 5.18. A real A is a *creative set* if A is r.e. and A^c is productive.

Therefore creative sets are not recursive. They also cannot be simple: A^c is productive and hence has infinite r.e. subsets. The intuition is that the production of A^c is a recursive way to show that A is not recursive, which is much stronger than simply not being recursive. According to Post, the interesting parts of recursion theory are in co-r.e. sets. One would have to be quite creative to understand the complement of a creative set.

Example 5.19. Let $S = \{x : x \in W_x\}$. Then S is creative.

Proof. The statement $x \in W_x$ can be expressed by saying $\exists s \exists y (\varphi_{s,x}(x) = y)$ which is Σ_1 .

We must show that S^c is productive. Let $h(x) = x$. Then if $x \in W_x$, W_x is not a subset of S^c since $x \in S$. If $x \notin W_x$, then $x \in S^c \setminus W_x$. So h is a production. \square

Theorem 5.20 (Myhill). A real A is creative iff A is m -complete.

Proof. Let A be r.e.

First suppose A is m -complete. Then the creative set S above satisfies $S \leq_m A$. Let f be recursive such that $n \in S$ iff $f(n) \in A$. Given W_x , let

$$W_{g(x)} = \{n : f(n) \in W_x\}.$$

So if $W_x \subseteq A^c$, then $W_{g(x)} \subseteq S^c$. Since the identity is a production for S^c , if $W_x \subseteq A^c$ then $W_{g(x)} \subseteq S^c$, so $g(x) \in S^c \setminus W_{g(x)}$, so $f(g(x)) \in A^c \setminus W_x$. Therefore $f \circ g$ is a production for A^c . Therefore A is creative.

If A is creative, and h is a production for A^c , and B is r.e., given e, y let $W_{f(e,y)} = \{h(e)\}$ if $y \in B$ or \emptyset otherwise. Here $\varphi_{f(e,y)}$ is defined by running the enumeration of B until y is enumerated, then halting and returning $h(e)$. Use the Kleene recursion theorem to find e_y such that $W_{f(e_y,y)} = W_{e_y}$. The function $y \mapsto e_y$ is recursive, and either $W_{e_y} = \emptyset$ or $h(e_y) \in W_{e_y}$.

If $W_{e_y} = \emptyset$ then $W_{e_y} \subseteq A^c$ so $h(e_y) \in A^c$. Otherwise, W_{e_y} is not contained in A^c so $h(e_y) \in A$. Thus $y \in B$ iff $h(e_y) \in A$, so $B \leq_m A$. \square

Thus A is creative iff A is m -complete iff A is 1-complete iff A is isomorphic to Sat_{Σ_1} .

Definition 5.21. An r.e. set A is *Turing complete* if $\text{Sat}_{\Sigma_1} \leq_T A$.

Theorem 5.22. There is a simple set which is Turing complete.

Proof. Fix an enumeration of Sat_{Σ_1} which enumerates exactly one number at each stage and enumerates each number exactly once. Let $\text{Sat}_{\Sigma_1}[s]$ denote the first s elements of Sat_{Σ_1} when enumerated this way. Assume $0 \notin \text{Sat}_{\sigma_1}$.

We will find a real A such that A^c is infinite, if W_k is infinite then $W_k \cap A$ is nonempty, and if n is enumerated into Sat_{Σ_1} at s then there is an $x \in [n^2, (n+1)^2]$ such that x is enumerated into A at stage s .

At stage s , we declare that k requires attention if $k \leq s$ and there is an $x > (k+1)^2$ such that $x \in W_k[s]$ and no element of $W_k[s]$ has been enumerated into A before s . For the least k which requires attention, enumerate its respective x into A . This guarantees that $W_k \cap A$ is nonempty. Let n be the number enumerated into Sat_{Σ_1} at stage s . Enumerate the least element of $[n^2, (n+1)^2]$ which is not in A into A .

By definition A is r.e. We enumerated at most $2n$ elements under $(n+1)^2$ so A is coinfinite. By attention, $W_k \cap A$ is nonempty. Moreover $[n^2, (n+1)^2]$ has $2n+1$ elements and at most $2n$ belong to A . Thus if n enters Sat_{Σ_1} at stage s , then there is an element of $[n^2, (n+1)^2]$ not yet in A which enters A at stage s .

Given n we compute whether $n \in \text{Sat}_{\Sigma_1}$. Run the enumeration of A until the first stage s_n where $A \cap [n^2, (n+1)^2]$ has been totally enumerated. At this point no more elements of $[n^2, (n+1)^2]$ can be added to A , which tells us that we could not enumerate n into Sat_{Σ_1} at a later stage. So $n \in \text{Sat}_{\Sigma_1}$ iff $n \in \text{Sat}_{\Sigma_1}[s_n]$, and the latter is finite. \square

So in particular there are Turing complete sets which are not many-one complete, and so Turing completeness is a weaker notion.

5.4 Forcing for sets below $0'$

Throughout this section, if A is a real, we view it as an infinite binary sequence and let $A|w$ denote its truncation to its first w entries. We let $\Sigma_1(A)$ be the set of all reals defined by $\exists w \theta(n, w, A|w)$ where θ is bounded. We define $Q_n(A)$ similarly. Then a real is $\Delta_1(0')$ iff it is $\Delta_2(0)$ iff it is recursive in $0'$. To do this, we need to use a special type of forcing that was already known to Kleene and Post.

Definition 5.23. A *forcing condition* is a finite binary sequence. We say $p \geq q$ if p is an initial segment of the forcing condition q .

Definition 5.24. Let D be a set of forcing conditions. Then D is *closed downwards* if $p \in D$, $p \geq q$, implies $q \in D$; D is *dense* if for every $p \in 2^{<\omega}$ there is a $q \in D$ such that $p \geq q$.

Being closed downwards means that we are closed under taking stronger conditions, and being dense means that for every condition we can find a weaker condition in it.

Definition 5.25. If θ is a bounded sentence and p is a forcing condition, then $p \Vdash \theta(G)$, that is, p forces $\theta(G)$, if the length of p is greater than the bounds in θ and $\theta(p)$ is true.

Intuitively $p \Vdash \theta(G)$ if we know that $\theta(q)$ will be true for $p \geq q$ such that q is arbitrarily long.

Definition 5.26. 1. $p \Vdash \exists x\theta(x, G)$ if there is an n such that $p \Vdash \theta(n, G)$.
2. $p \Vdash \forall x\theta(x, G)$ if for every $q \leq p$, $q \Vdash \exists x\neg\theta(x, G)$.
3. p decides ψ if $p \Vdash \psi$ or $p \Vdash \neg\psi$.

If ψ is a $\Sigma_1(G)$ -sentence for some real G , then $p \Vdash \psi$ is a recursive statement about p and $p \Vdash \neg\psi$ is a Π_1 property of p . To see this, note that to check whether $p \Vdash \exists x\theta(x, G)$ is decided by checking whether $p \Vdash \theta(n, G)$ for all $n \leq |p|$.

Lemma 5.27. If ψ is $\Sigma_1(G)$, then the set of forcing conditions which decide ψ is dense.

Proof. If $p \Vdash \theta(G)$ and θ is bounded, then for any $q \leq p$ we have $q \Vdash \theta(G)$, since the quantifiers in θ only referred to numbers under $|p| \leq |q|$. In particular, $p \Vdash \theta(G)$ iff for every $q \leq p$, $q \Vdash \theta(G)$. Now that we know the claim is true for $\Delta_1(G)$ sentences it is easy to see that it is true for $\Sigma_1(G)$ sentences. \square

Definition 5.28. A real G is 1-generic if for every Σ_1 set of conditions $S \subseteq 2^{<\omega}$, either there is an ℓ such that $G|\ell \in S$, or there is no $q \leq G|\ell$ such that $q \in S$.

Therefore at any finite stage, G has decided whether G contains an element of S . In other words, if ψ is Σ_1 , there is an ℓ such that $G|\ell$ decides ψ .

We now show that forcing is equivalent to truth for generic reals.

Theorem 5.29. Suppose that G is a 1-generic real and $\psi(G)$ is a Σ_1 -sentence. Then $\psi(G)$ is true iff there is an ℓ such that $G|\ell \Vdash \psi$.

Proof. Suppose $\psi(G) = \exists x\theta(x, G|x)$ where θ is bounded. Let $S = \{p : p \Vdash \psi(G)\}$. Thus S is defined by $\exists x(p \Vdash \theta(x, G))$, so S is r.e. Since G is generic, there is an ℓ such that $G|\ell \in S$ or $G|\ell$ has no extension in S .

If $G|\ell \in S$ then $G|\ell \Vdash \psi(G)$ and $\mathbb{N} \models \psi(G)$.

Otherwise, $G|\ell \Vdash \neg\psi(G)$, and if $\exists x\theta(G)$ were true, then there would be an x such that $\theta(x, G|x)$ holds, and $G|\max(x, \ell)$ would extend $G|\ell$ in S , a contradiction. Therefore $\neg\psi(G)$ is true. \square

Theorem 5.30. There is a 1-generic real which is recursive in $0'$.

Proof. Let $\{S_n\}$ be an enumeration of all the Σ_1 subsets of $2^{<\omega}$. Let p_0 be the empty condition.

Let p_{n+1} be the lexicographically least $q \leq p_n$ such that $q \in S_n$, if one exists, or p_n otherwise. Computing p_{n+1} from p_n is recursive in $0'$ since it is asking an existential question.

Let $G = \lim_n p_n$. Since $A_k = \{p : |p| \geq k\}$ is dense, and $G \cap A_k$ for all k , G is infinite. For every S_n , either G meets it or cannot enter into it by extension, so G is 1-generic. \square

This was just the proof of the Baire category theorem. To see this, let $\{N_n\}$ be an enumeration of a countable set of nowhere dense sets of reals. Let $U_0 = \mathbb{R}$. Choose U_n to be an open set outside N_n and contained in U_{n-1} . Then the U_n have nonempty intersection U . A generic real lies inside U .

In fact, the above theorem immediately follows from the Baire category theorem when applied to the topology of the Cantor set 2^ω . The topology of 2^ω is generated by the cylinders

$$B_\sigma = \{x : \exists n(x|n = \sigma)\}.$$

Let $S \subseteq 2^{<\omega}$ be Σ_1 . Then for every generic real G , either $\exists \ell(G|\ell \in S)$ or $\exists \ell \forall \tau(G|\ell \geq \tau \rightarrow \tau \notin S)$, and the condition $\exists \ell(G|\ell \in S)$ or $\exists \ell \forall \tau(G|\ell < \tau \rightarrow \tau \notin S)$ is dense in the Cantor sets. Since there are only countably many Σ_1 sets the claim immediately follows.

Theorem 5.31. If G is 1-generic then G is not computable and does not compute $0'$.

Proof. Suppose $f : \mathbb{N} \rightarrow 2$ is total and recursive. Let $S = \{p : \exists n < |p|(p(n) \neq f(n))\}$. Then S is Σ_1 and dense, so there is an ℓ such that $G|\ell \in S$. So there is an n such that $G(n) \neq f(n)$, so G is not computable.

Now suppose G compute $0'$. Then $N \setminus 0'$ is $\Sigma_1(G)$, say defined by $\psi = \exists x \theta$. Let S be the set of p such that $\exists n < |p| \psi(n, p)$. Then S is Σ_1 and G has no initial segment in S . Let ℓ be such that $G|\ell$ has no extension in S . Then for any n , $n \notin 0'$ iff $\exists q \leq G|\ell$ such that $\exists x < |q| \theta(n, x, q)$.

$$S = \{p : \exists n \in 0' \exists x < |p| \theta(n, x, p)\}.$$

□

Intuitively, recursive sets cannot achieve everything that is possible because they never disagree from themselves, so are not generic.

5.5 Incomparable reals

We now use forcing to prove the following theorem of Kleene and Post.

Theorem 5.32 (Kleene-Post). There are reals $A, B \leq_T 0'$ which are Turing incomparable.

Proof. Let G be 1-generic and recursive in $0'$. Let $A = \{n : 2n \in G\}$ and $B = \{n : 2n + 1 \in G\}$. By symmetry, it suffices to show that $A \not\leq_T B$, and in fact we will show $A \notin \Sigma_1(B)$.

Suppose that $\psi = \exists x \theta$ is Σ_1 . For p a condition, let p_0 and p_1 be the even and odd parts of p , $p_j(n) = p(2n + j)$. Let

$$S = \{p : \exists n(p_0 \Vdash \psi(n, G) \wedge p_1(n) = 0)\}.$$

Then S is recursive. If G meets S , say $G|\ell \in S$, then there is an $n \notin B$ such that $\psi(n, A)$. So $\psi(A) \neq B$.

Otherwise, there is an ℓ such that no extension of $G|\ell$ is in S . We claim that $\psi(A)$ is finite. To see this, let $n > \ell$ and x be given. If q_0 is an extension of p_0 such that $|q_0| > n$ and $\theta(n, x, q_0)$, extend p_1 to q_1 such that $|q_1| = |p_1|$ and $p_1(n) = 0$. So q extends $G|\ell$ and $q \in S$, which is impossible. Therefore $|\psi(A)| \leq n$.

But, for any k , it is generic that $|\{n : p_1(n)\}| > k$, so B is infinite. Therefore $\psi(A) \neq B$. □

The original Kleene-Post theorem was a bit stronger than this: for any countable poset P , there is an embedding of P into the set of Turing degrees below $0'$. As a corollary, the existential theory of the Turing degrees below $0'$ is decidable.

Definition 5.33. A *locally countable poset* is a poset P such that for every x , $\text{card}\{y : y < x\} = \aleph_0$.

Clearly the Turing degrees are locally countable. An open problem asks whether this is a universal property of the Turing degrees.

Conjecture 5.34. Suppose P is a locally countable poset such that $\text{card } P = 2^{\aleph_0}$. Then there is an embedding of P into the Turing degrees.

Theorem 5.35. If $\aleph_1 = 2^{\aleph_0}$ then the above conjecture is true.

5.6 Inverting the jump

Suppose that G is any real. Then $0' \leq G'$, so the Turing jump maps the set of all reals into the set of reals above $0'$. The converse is also true.

Theorem 5.36 (Friedberg jump inversion). If $0' \leq_T X$, then there is a 1-generic real G such that $G' \equiv_T X$.

Proof. Let $\{S_n\}_n$ be an enumeration of the Σ_1 -definable subsets of $2^{<\omega}$. Let p_0 be the empty condition. If there is a q which extends an element of S_n , let p_{n+1}^* be the lexicographically least q . Otherwise let $p_{n+1}^* = p_n$. Then let $p_{n+1} = p_{n+1}^* + X(n)$, where $+$ is concatenation. Then G is 1-generic.

We now show that $G' \equiv_T G \oplus 0' \equiv_T X$. By assumption X can decide $0'$, and then X can decide whether a q exists (and what it is), so X can decide p_{n+1} and hence G . Therefore $G \oplus 0' \leq_T X$. On the other hand, $0'$ can compute p_n^* and hence $G \oplus 0'$ can compute X .

Clearly $G' \geq_T G \oplus 0'$. Conversely, we show that $G \oplus 0'$ computes G' . Let ψ be Σ_1 . Since G is 1-generic, so there is an ℓ such that $G|\ell$ decides $\psi(G)$. What $G|\ell$ decides is computable from $0'$. Therefore $0'$ can compute the existential theory of G , hence can compute G' . \square

Theorem 5.37 (Posner-Robinson). For every nonrecursive real X , there is a 1-generic real G such that $G' \equiv_T G \oplus X$.

Proof. By replacing X with X^c if necessary, we can assume that X is not r.e. Let $\{\psi_n\}_n$ be an enumeration of the $\Sigma_1(G)$ formulae.

Let p_0 be the empty sequence. Given p_n , let k be the least number such that $k \in X$ and $p_n + \bar{k} + 0 \Vdash \neg\psi_n(G)$, or $k \notin X$ and there is a q such that $p_n + \bar{k} + 0 \geq q$ and $q \Vdash \psi_n(G)$. Here \bar{k} consists of k many 1's and $+$ is concatenation. In the first case, let $p_{n+1} = p_n + \bar{k} + 0 + X(n)$, and otherwise take q' to be the lexicographically least q and take $p_{n+1} = q' + X(n)$. Let $G = \lim_{n \rightarrow \infty} p_n$.

We need to show that for every n , k exists. Suppose not, so p_n exists but not p_{n+1} . Then for every k , $k \in X$ iff there is a q such that $p_n + \bar{k} + 0 \geq q$ and $q \Vdash \psi_n(G)$. Since \Vdash for Σ_1 sentences is Σ_1 , X now has a Σ_1 definition, namely

$$\exists q((p_n + \bar{k} + 0 \geq q) \wedge q \Vdash \psi_n(G)).$$

This is a contradiction.

To see $G \oplus X \leq_T G'$, note that G' knows what our computation steps were, so can compute p_n . But if we know every p_n , we can compute G and X .

Conversely, we will show that $G \oplus X$ can compute the p_n and hence G' . Suppose we know p_n . To compute p_{n+1} , we determine the length k of consecutive 1's and then checking whether $k \in X$. Knowing k and X allows us to compute the initial segment p_{n+1} of X . Then $\psi_n(G)$ is true iff p_{n+1} thinks it is true. \square

According to Ted Slaman, coming up with the coding trick used in the Posner-Robinson theorem probably took a long time, but once one comes up with it it's obvious that it will work. Slaman learned this theorem in the most poorly motivated talk he had heard in a long time, when he was a graduate student. However, he frequently found the theorem useful in the years following.

The Posner-Robinson theorem also easily relativizes.

Corollary 5.38. Let B, X be real numbers such that $X \not\leq_T B$. Then there is a 1-generic G such that $B \leq_T G$ and $G \oplus X = G'$.

The proof is exactly the same; just whenever we had a computation in the proof of the Posner-Robinson theorem, allow the use of B as a blackbox oracle.

5.7 The axiom of determinacy

The Posner-Robinson theorem has applications to the axiom of determinacy.

Throughout, let ω^ω be the Baire space, whose points we will call reals (rather than points in Cantor space, as above.) We consider the *standard Gale-Stewart game* G_A on ω^ω with victory set $A \subseteq \omega^\omega$: Alice and Bob alternate between playing natural numbers: Alice plays a_1 , Bob plays a_2 , Alice plays a_3 , etc. Then Alice wins iff the sequence is in the victory set A .

Since ω^ω is a product space, its topology is generated by cylinders of the form U_σ for each $\sigma \in \omega^{<\omega}$.

Definition 5.39. A *strategy* for Alice (Bob) is a function that maps finite sequences of even (odd) length to natural numbers. The strategy *wins* the game G_A if it gives an algorithm which guarantees that no matter what Bob (Alice) plays, the infinite sequence generated is in A .

Example 5.40. If $A = \omega^\omega$ then any strategy for Alice wins. Similarly if $A = \{x : x(0) = 0\}$, where the strategy “Always play 0” wins for Alice. If $A = \emptyset$ then any strategy for Bob wins.

Definition 5.41. If there is a winning strategy for victory set A , we say that G_A is a *determined game*.

Theorem 5.42 (Gale-Stewart). Every open game is determined. In particular, closed games or games of finite length are determined.

Proof. Suppose we are playing G_A , A is open and Alice does not have a winning strategy. Assume that we have played a_1, \dots, a_{2n} and now Alice plays a_{2n+1} . Since Alice does not have a winning strategy, there is a a_{2n+2} such that there is a sequence beginning with a_1, \dots, a_{2n+2} which is not in A . So let Bob play a_{2n+2} .

Suppose that Bob does not win following the above strategy. In particular, since Bob did not win, $\{a_n\}_n \in A$, so there is a restriction σ of $\{a_n\}_n$ such that $\{a_n\}_n \in U_\sigma$. So Bob lost at the finite time $|\sigma|$, which is impossible since Alice didn't have a winning strategy.

Games of finite length are exactly those of the form G_{U_σ} for some σ , so are open. The complement of a determined game is determined. \square

The Gale-Stewart theorem is really just another restatement of the Baire category theorem.

Theorem 5.43. There is a victory set that is not determined.

Proof. First, for any strategy σ , the set of possible outputs if a player uses strategy σ has cardinality 2^{\aleph_0} . Similarly there are 2^{\aleph_0} many strategies. So let $<$ be a well-ordering of ω^ω and $<_s$ a well-ordering of the set of strategies, such that $<$ and $<_s$ have the same ordertype.

We now diagonalize against the set of all strategies. If $\alpha < 2^{\aleph_0}$, we have specified membership in A or its complement for fewer than 2^{\aleph_0} many elements of ω^ω . Let σ_α be the α th strategy under ordering $<_s$. If σ_α is a strategy for player P , let a_α be the $<$ -least output such that we have not determined whether a_α is in A ; if $P = 1$ declare $a_\alpha \in A^c$, otherwise declare $a_\alpha \in A$.

If a was not decided for any a_α , let $a \in A$. Then neither player has a winning strategy. \square

Note that this theorem uses the axiom of choice in an extremely strong way.

Axiom 5.44 (determinacy). For every set $A \subseteq \omega^\omega$, G_A is a determined game.

This axiom is provably false from ZFC, but it is still very interesting, because of the following theorem.

Definition 5.45. A set $A \subseteq \omega^\omega$ is a *projective set* if there is a closed set $C \subseteq \omega^\omega$ such that A can be constructed from C using only the operations of complementation and projection $\omega^\omega \times \omega^\omega \rightarrow \omega^\omega$.

Here we are using the fact that $\omega^\omega \times \omega^\omega = \omega^\omega$.

Definition 5.46. A *Woodin cardinal* is a cardinal λ such that for every $f : \lambda \rightarrow \lambda$, there is

1. a cardinal $\kappa < \lambda$ such that f restricts to a function $\kappa \rightarrow \kappa$,
2. a transitive class $M \subseteq V$,
3. and an elementary embedding $j : V \rightarrow M$ such that $\text{crt } j = \kappa$

such that j restricts to a function $V_{j(f)(\kappa)} \rightarrow M$.

Theorem 5.47 (Martin-Steel). Every Borel game is determined, and if there are infinitely many Woodin cardinals, then every projective game is determined.

The assumption that there are infinitely many Woodin cardinals seems ridiculously strong. But the proof of this for Borel sets is not much better: it uses arbitrarily large iterates of the power set operation, and this part of the proof is provably unavoidable.

Theorem 5.48 (Friedman-Woodin). If every projective set is determined then there is a model of ZFC with infinitely many Woodin cardinals. Moreover, there is a model of Zermelo theory with a Borel game which is not determined.

Here *Zermelo theory* is set theory without the axiom schema of replacement (which in particular does not allow one to iterate the power set operation for arbitrarily long.)

Determined sets have very nice properties. This includes applications to measure theory and the Turing degrees.

Theorem 5.49. Every determined set is Lebesgue measurable.

Here we identify subsets of 2^ω with subsets of $[0, 1]$ according to the mapping

$$a \mapsto \sum_{n=0}^{\infty} \frac{a_n}{2^{n+1}}.$$

Thus it makes sense to ask if a subset of \mathbb{R} is determined.

Proof. Let μ denote Lebesgue measure and μ^* denote Lebesgue outer measure.

We need the following lemma.

Lemma 5.50. Assume that S is a determined set and for every measurable $Z \subseteq S$, Z is null. Then S is null.

Once we have proven this lemma, the claim immediately falls out. Indeed, for any set E , there is a G_δ -set $F \supseteq E$ such that $\mu^*(F) = \mu^*(E)$; in particular, for every measurable subset Z of $F \setminus E$, Z is null. Now by the Gale-Stewart theorem, F is determined; if E is also determined, $F \setminus E$ is determined. So by lemma, every measurable subset of $F \setminus E$ is null, so $F \setminus E$ is null. Therefore E is measurable.

To prove the lemma, let $\varepsilon > 0$; we claim $\mu^*(S) < \varepsilon$. For any n let K_n be the countable set of all sets G which are finite unions of open intervals with rational endpoints, and $m(G) < \varepsilon/2^{2n+1}$, and let $G_{n,k}$ enumerate K_n .

We define a game that Alice wins if she can play a point $a \in S$ such that $a \notin \bigcup_n G_{n,k}^n$, where Bob plays b . Since everything is determined, to show that Bob has a winning strategy, it suffices to show that Alice does not.

Let σ be a winning strategy for Alice, which gives a continuous function $\omega^\omega \rightarrow \mathbb{R}$. Then σ is continuous, so $Z = \sigma(\omega^\omega)$ is measurable, and so by assumption Z is null. So there are open sets $H_n \in K_n$ which cover Z . Therefore Bob wins if he plays optimally against σ , a contradiction.

Since Bob has a winning strategy, he can cover S by open sets $H_n \in K_n$, and $\bigcup_n \mu(H_n) < \varepsilon$. Therefore S is null. \square

Theorem 5.51 (Martin cone theorem). Let \mathcal{D} be the set of Turing degrees, and let A be a cofinal subset of \mathcal{D} . Suppose that the set $\{x \in \omega^\omega : [x]_T \in A\}$ is determined. If for every $x \in A$ and every $y \equiv_T x$, $y \in A$, then there is a b such that for all a , if $b \leq_T a$ then $a \in A$.

Proof. Let P be the determined set.

Suppose Alice has a winning strategy for G_P . Now a strategy consists of functions $\omega^{2n} \rightarrow \omega$ for all $n \in \omega$, each of which can be coded as a point $x_n \in \omega^\omega$. Clearly the x_n form a Cauchy sequence, so they converge to a point $x \in \omega^\omega$, Alice's *master strategy*. Let $b = [x]_T$ be the Turing degree of Alice's master strategy.

Suppose that Bob counters Alice's master strategy by playing a sequence y , and $a = [y]_T$. If $a \geq_T b$ then the sequence that Alice and Bob generate is of degree a , so $a \in A$.

Now suppose Bob has a winning strategy. Running the same argument as above, we conclude that if $a \geq_T b$ then $a \in A^c$. Since \mathcal{D} does not have a greatest element (since we can always take the Turing jump), and A is cofinal, we find an element $a \in A \cap A^c$, a contradiction. \square

Definition 5.52 (Martin). A function $f : \omega^\omega \rightarrow \omega^\omega$ is *degree-invariant* if for any x, y , if $x \equiv_T y$ then $f(x) \equiv_T y$. If f, g are degree-invariant, we say that $f \geq_M g$ if there is a y such that for any $x \geq_T y$, $f(x) \geq_T g(x)$.

Example 5.53. The Turing jump operation is degree invariant, as is the constant function, the identity function, and iterates of the Turing jump.

Conjecture 5.54 (Martin). Let I be the set of degree-invariant functions. If $f \in I$ and $f \not\geq_M \text{id}$ then there is a B such that for all $A \geq_T B$ $f(A) \equiv_T B$. Moreover, $I' = \{f \in I : f \geq_M \text{id}\}$ is well-ordered by \geq_M (modulo Martin equivalence \equiv_M). For all $f \in I'$, the successor f' of f is given by

$$f'(x) = f(x)'.$$

Thus Martin's conjecture essentially rules out the existence of \geq_M -incomparable functions.

One can use the Posner-Robinson theorem to prove a special case of Martin's conjecture.

Theorem 5.55 (Slaman-Steel). Assume that $f : \omega^\omega \rightarrow \omega^\omega$ is determined. Let $Jx = x'$ be the Turing jump. Assume f preserves \geq_T and $f >_M \text{id}$. Then $f \geq_M J$.

Proof. We must show that for all B there is an $A \geq_T B$ such that $f(A) \geq_M B$. By assumption, we can take B large enough that $f(B) >_T B$.

By the Posner-Robinson theorem, there is a $G \geq_T B$ such that $f(B) \oplus G \equiv_T G'$. But then $f(G) \geq_T G$ since $G \geq_T B$ and f increases there. Thus $f(G) \geq_T f(B)$. So

$$G' \equiv_T f(B) \oplus G \leq_T f(G).$$

Thus $JG \geq_T f(G)$. \square

5.8 The Friedberg-Muchnik theorem and the priority argument

Theorem 5.56 (Friedberg-Muchnik). There are Σ_1 reals A, B which are Turing incomparable.

If X is a real number that we are enumerating, let $X[s]$ denote the set of reals that have been enumerated at stage s . Let $\Phi_i(X)$ denote the i th Turing functional (i.e. program) which uses X as an oracle. We will construct A, B so that for any i , $\Phi_i(A) \neq B$ and $\Phi_i(B) \neq A$. We let R_i^A mean that $\Phi_i(A) \neq B$ and similarly for R_i^B . Let $\Phi_i(n, X)[s]$ denote the result of running the i th program with oracle X and input n for at most s stages and returning only numbers less than s .

Suppose that we want to satisfy just one requirement, namely R_i^A , at stage s . We execute the following strategy:

1. Let n be the stage the strategy is activated in. Suppose $n \notin B$ and go to Step 2.
2. If $\Phi_i(n, A)[s-1] = 0$, i.e. $\Phi_i(A)$ thinks that n does not belong in B at stage $s-1$, set $s_i^A = s$ and go to Step 3. Otherwise, do not allow n to enter B at stage s .
3. Enumerate n into B . Restrain any number less than $s_i^A - 1$ from entering A at stage s . Expect that n is much smaller than s .

At stage $s+1$ we initialize the above strategy in the Step that we ended it at stage s at.

Now this is not an algorithm to enumerate A, B , but if we can prove this existence of an algorithm to enumerate A, B such that for every s , R_i^A is satisfied at stage s , then $B \not\leq_T A$, regardless of what that algorithm is. In fact, either as $s \rightarrow \infty$ we stay at Step 2, so $B(n) = 0$ and $\Phi_i(n, A) \neq 0$, so $\Phi_i(A) \neq B$, or we end at Step 3, in which case at stage s_i^A , n enters B and hence $B(n) = 1$. But if we ended at Step 3, then $\Phi_i(n, A) = 0$, since the computation that sent us to Step 3 is repeated over and over forever. So $\Phi_i(A) \neq B$.

Unfortunately, it is not true that we can require R_i^A and R_i^B at every stage of computation. If we cannot carry out some strategy, we say that it is an *injured strategy*. Our goal is to show that every strategy is only injured finitely many times. Consider the enumeration

$$R_0^A, R_0^B, R_1^A, R_1^B, R_2^A, \dots$$

We will say that a requirement S has *higher priority* than S' if S appears earlier in the enumeration than S' . We will not allow a higher priority strategy injure a lower priority strategy. This proof technique is known as the *finite injury method*.

At stage 0, let $A[0] = B[0] = \emptyset$ and have no strategies active.

At stage s , we are given the sequence of strategies that were active at the end of stage $s-1$, the record of the numbers they selected, the restraints they posed in Step 3, and the values of $A[s]$ and $B[s]$.

Declare that the strategy for R_i^X *requires attention* if it is not active, or it is in Step 2 with number n and $\Phi_i(n, X)[s-1] = 0$. Let R be the highest priority argument that requires attention. If R is not currently active, activate it and follow its instructions in Steps 1, 2. Choose n and assign it to R . Otherwise, if R is already active, follow its instructions as it moves to Step 3. Enumerate the n associated to R into the appropriate set.

At the end of stage s , deactivate, or *injure*, all strategies of lower priority to R . If they were injured, they will be restarted at a later stage.

Lemma 5.57. For each requirement R , there is a strategy for R that is never deactivated.

Proof. By induction on priority. R_0^A had highest priority, so it is never injured.

Suppose that the claim holds for all requirements of higher priority than R , and assume R^- is the requirement of highest priority such is of lower priority than R . Let s_0 be the final stage where R^- was injured. After this, no requirement of higher priority than R^- can require attention. But then R^- can only require attention at most twice after s_0 : at $s_0 + 1$ when R^- reactivates, and possibly a later stage when it moves from Step 2 to Step 3.

Therefore there is a s_1 such that R^- does not require attention after s_1 . Since no other requirements of higher priority to R can require attention after stage s_1 , after s_1 , R never deactivates. \square

Lemma 5.58. For each requirement R , there is a stage s such that a strategy for R is activated at s and the construction implements all constraints that the strategy imposes after s .

Proof. Let t be the last stage for which there was no active strategy for R . No strategy of higher priority will require attention for any stage after t , and R requires attention at stage $s = t + 1$.

Since R is of highest priority among those requiring attention at stage s , the construction follows R 's instructions at stage s . R records $n = s$. Since strategies can only enumerate numbers less than the current stage, $n \notin A[s]$ and $n \notin B[s]$. Since activating R deactivates all lower-priority strategies, no such strategy enumerates n . Since strategies of higher priority ignore n entirely, they also do not enumerate n .

If R remains at Step 2 forever, then the lemma is verified. Otherwise, suppose that R requires attention at stage $s_1 > s$. Then R will be the highest priority strategy at s_1 , and will enumerate n into the appropriate set, and deactivate all strategies of lower priority. These strategies choose new values for their parameters at later stages, and these parameters will exceed $s_1 > n$, so they will not contradict the requirements R imposed. The higher priority strategies will never again require attention, so they too will not contradict the requirements imposed by R . Therefore the construction implements the requirements of R . \square

Therefore the strategies are satisfied at cofinitely many stages and hence A, B are as desired.

One can extend the above techniques to require that A, B are, for example, simple. We would have to add additional strategies that also require that A, B are simple.

Conjecture 5.59 (Sacks). There is an e such that for all A :

1. $A <_T W_e^A <_T A'$
2. For all B , if $A \equiv_T B$ then $W_e^A \equiv_T W_e^B$.

This would give a counterexample to Martin's conjecture. Note that the first condition requires that e is a solution to Post's problem relative to all A , and the second condition requires that W_e is degree-invariant.

Sacks was the Ph.D. adviser to Ted Slaman, and according to Slaman, Sacks believed that if his conjecture were true he would find the solution extremely interesting. But if Martin's conjecture were true, Sacks would not find the solution very interesting. Sacks had an ideal of recursion theory that was based on elaborate constructions like the finite-injury method, while Martin had an ideal of recursion theory as a field with a beautiful global structure.

Chapter 6

Elementary set theory

6.1 The axioms of ZF without Foundation

By an *LST-formula* we mean a first-order formula in the language of set theory, namely the language consisting of a single binary relation symbol \in .

The axioms are:

Axiom of ZFC 1 (extensionality). A set is determined by its elements.

Axiom of ZFC 2 (empty set). There is a set with no elements, denoted \emptyset .

Axiom of ZFC 3 (pairing). For any two sets x, y , the set $\{x, y\}$ exists.

Axiom of ZFC 4 (union). For any set x , the set $\bigcup x$ exists.

Axiom of ZFC 5 (power set). For any set x , the set $\mathcal{P}x$ exists.

By extensionality, all the above sets are unique. Moreover, the class V_ω of hereditarily finite sets is a model of the above theory. So we need to introduce a new axiom to escape V_ω . This will be the first “small large-cardinal axiom.”

Axiom of ZFC 6 (infinity). There is a set x such that $\emptyset \in x$ and such that for each $y \in x$, $y \cup \{y\} \in x$.

Then ω is the intersection of all such sets. (We will call these sets *inductive*.)

Now $V_{\omega+\omega}$ is a model of the above theory, but does not contain \aleph_ω . So we introduce the following schema of “small large-cardinal axioms.”

Axiom of ZFC 7 (replacement schema). For every LST-formula φ , we have

$$\forall \vec{p} \forall x (\forall y \in x \exists! z \varphi(y, z, \vec{p}) \rightarrow \exists y \forall z (z \in y \leftrightarrow \exists w \in x \varphi(y, z, \vec{p}))).$$

6.2 Transfinite induction

Fix a binary relation R on a class X .

Definition 6.1. The R -extension of $x \in X$ is

$$\text{ext}_R(x) = \{x^* \in X : x^* Rx\}.$$

R is *transitive* if for every x_1, x_2, x_3 , $x_1 Rx_2$ and $x_2 Rx_3$ implies $x_1 Rx_3$. The *transitive closure* of R , $\text{TC}(R)$, is the smallest binary relation which is transitive and contains R .

The R -predecessors of x are

$$\text{pred}_R(x) = \text{ext}_{\text{TC}(R)}(x).$$

$Y \subseteq X$ is *R-transitive* if for every $y \in Y$, $\text{pred}_R(y) \subseteq Y$ (i.e. Y is closed under pred).

Definition 6.2. R is *well-founded* if for every set $Y \subseteq X$ there is a R -minimal element, and if for every $x \in X$, $\text{ext}_R(x)$ is a set.

Axiom of ZFC 8 (foundation). \in is well-founded.

We will always assume foundation in what follows. Foundation, along with the above axioms, comprise the Zermelo-Frankel axiom system. This is still not ZFC because we have no introduced choice.

Note that if X is actually a set, then R is well-founded iff every subset has a minimal element, and if (X, R) is a chain, then this happens iff R is a well-order. In fact we take this as a definition.

Definition 6.3. R is *strict* if for every $x \in X$, $\neg(x Rx)$.

R is *linear* if every $x_1, x_2 \in X$, $x_1 Rx_2$ or $x_2 Rx_1$.

Definition 6.4. R is a *well-ordering* if R is strict, linear, and well-founded.

Axiom 6.5 (well-ordering). Every set admits a well-ordering.

We will *not* assume this axiom.

Now we have set up transfinite induction and recursion.

Theorem 6.6 (transfinite induction). Let R be well-founded and $Y \subseteq X$. If for every $x \in X$ such that $\text{pred}_R(x) \subseteq Y$, $x \in Y$, then $Y = X$.

We think of x as the “successor” of its predecessors $\text{pred}_R(x)$. So this is analogous to strong induction.

Proof. Assume $x \in X \setminus Y$. Then we can take x to be minimal since R is well-founded. Then if $y \in \text{pred}_R(x)$, $y \notin X \setminus Y$ so $y \in Y$. So $x \in Y$, a contradiction. \square

Theorem 6.7 (transfinite recursion). Let R be well-founded and $G : X \times V \rightarrow V$ is a class function. There is a unique class function $F : X \rightarrow V$ such that for every $x \in X$,

$$F(x) = G(x, F|_{\text{pred}_R(x)}).$$

Proof. Uniqueness follows immediately by transfinite induction. For existence, let x be minimal and put $f_0(x) = G(x, \emptyset)$ to start the induction.

Say that $f : D \rightarrow X$ is *good* if $D \subseteq X$, for every $x \in D$, $\text{pred}_R(x) \subseteq D$, and $f(x) = G(x, f|_{\text{pred}_R(x)})$. Clearly f_0 is good. By uniqueness, if $f_1 : D_1 \rightarrow X$ and $f_2 : D_2 \rightarrow X$ are good, then $f_1 \cap f_2 : D_1 \cap D_2 \rightarrow X$ is defined. By transfinite induction with f_0 as the base case, the union of all D such that there is a good f is X itself. So by “gluing” the good functions, F is uniquely defined. \square

6.3 Ordinals

Let R be a well-ordering of a class X . For $x \in X$, we let I_x^R be the initial segment $\text{pred}_R(x)$ with its induced well-ordering.

By an easy transfinite induction, we have the following theorem.

Theorem 6.8. Let S be a well-ordering of a class Y . Then either Y is an initial segment of X or X is an initial segment of Y .

Definition 6.9. A *transitive set* X is one for which $x \in X$ implies $x \subset X$.

Lemma 6.10. If X is transitive then so are $X \cup \{X\}$, $\bigcup X$, and $\mathcal{P}X$.

Obviously \emptyset is transitive.

Definition 6.11. An *ordinal* is a transitive set α such that \in is a well-ordering of α .

We denote the class of ordinals **Ord**. Now if $\alpha \in \mathbf{Ord}$ and $\beta \in \alpha$, $\beta \in \mathbf{Ord}$. So α consists of the ordinals under α . In particular $\beta \in \alpha$ iff $\beta \subset \alpha$. So it follows that **Ord** is well-ordered by \in .

Theorem 6.12. **Ord** is a proper class.

Proof. Suppose not. Then \in is a well-ordering of **Ord**, so **Ord** is an ordinal. So $\mathbf{Ord} \in \mathbf{Ord}$, contradicting foundation. \square

Now we see that the only well-ordered sets are the ordinals.

Theorem 6.13. Let (X, R) be a well-ordered set. Then there is an ordinal α and an isomorphism $X \rightarrow \alpha$.

Definition 6.14. The ordinal α is called the *ordertype* of (X, R) .

To prove this we need a lemma.

Lemma 6.15. For any $x \in X$ there is an ordinal α such that $I_x^R \cong \alpha$.

Proof. Apply replacement from I_x^R into **Ord**. \square

Proof of theorem. Take the set β of all ordinals α such that for some $x \in X$, $I_x^R \cong \alpha$. Then clearly β is an ordinal. \square

Now we define arithmetic on ordinals by transfinite recursion. Namely, we put $\alpha + 0 = \alpha$, $\alpha \cdot 0 = \alpha$, and $\alpha^0 = 1$. We then put $\alpha + (\beta + 1) = (\alpha + \beta) + 1$, $\alpha(\beta + 1) = \alpha\beta + \alpha$, and $\alpha^{\beta+1} = \alpha^\beta\alpha$. Then we take unions at limit stages. We let ω denote the smallest infinite ordinal, which exists by the axiom of infinity.

Theorem 6.16 (division algorithm). Let α and $\beta > 0$ be ordinals. There are unique γ_1 and $\gamma_2 < \beta$ such that

$$\alpha = \beta\gamma_1 + \gamma_2.$$

The proof is the same as for the classical division algorithm.

Theorem 6.17 (Cantor normal form). If $\alpha > 0$ is an ordinal then we can uniquely write

$$\alpha = \omega^{\beta_1}\kappa_1 + \cdots + \omega^{\beta_n}\kappa_n$$

where $\alpha > \beta_1 > \cdots > \beta_n \in \mathbf{Ord}$ and $\kappa_1, \dots, \kappa_n, n \in \omega$.

The proof is by greedy transfinite induction.

Definition 6.18. The *cumulative hierarchy* is defined by transfinite recursion as:

1. $V_0 = \emptyset$.
2. $V_{\alpha+1} = \mathcal{P}V_\alpha$.
3. $V_\lambda = \bigcup_{\alpha \in \lambda} V_\alpha$ for λ a limit ordinal.

Finally V is the proper class $V = \bigcup_{\alpha \in \mathbf{Ord}} V_\alpha$.

By transfinite induction, every V_α is transitive. Since $V_\alpha \in V_\beta$ for $\beta > \alpha$, transitivity implies that the V_α form a chain with respect to \subset .

By foundation, \in is a well-founded relation on the proper class of all sets. Thus we have the following theorem.

Theorem 6.19. For any set x , $x \in V$.

Proof. Suppose not. Then take x to be an \in -minimal set which does not appear in V , which is possible since \in is well-founded. Then $x \in V$, a contradiction. \square

Definition 6.20. The *rank* α of a set x is the least ordinal such that $x \in V_{\alpha+1}$.

6.4 The axiom of choice

We now study the most famous of the axioms of ZFC.

Definition 6.21. A *choice function* on a set X is a function $f : X \rightarrow \bigcup X$ such that for every $x \in X$, $f(x) \in x$.

Axiom of ZFC 9 (choice). Every set admits a choice function.

$\text{ZF} + \text{choice} = \text{ZFC}$. However we will *not* assume choice until a later stage. Here's another famous axiom.

Axiom 6.22 (Zorn). If (X, \leq) is a nonempty poset such that every subchain of X has an upper bound, then X has a maximal element.

Theorem 6.23. Choice, well-ordering, and Zorn are equivalent.

Proof. First we prove that choice implies well-ordering. Given X there is a choice function on $\mathcal{P}X \setminus \{\emptyset\}$. Now let $x_1 \in f(X)$, $x_2 \in f(X \setminus \{x_1\})$, \dots . This gives an bijective function $\alpha \rightarrow X$, $\beta \mapsto x_\beta$, where $\alpha \in \text{Ord}$, and must stop after transfinitely many steps because if not then $\alpha = \text{Ord}$, a contradiction.

Now assume well-ordering. To prove Zorn we let (X, \leq) be a poset such that every subchain of X has an upper bound. By well-ordering there is a bijection $f : \alpha \rightarrow X$ for some $\alpha \in \text{Ord}$. This gives an injection $g : \beta \rightarrow X$ for some $\beta \in \alpha$ by choosing $g(0) = f(0)$ and always letting $g(\beta)$ be the f -minimal element in the current chain. This process stops after transfinitely many steps (or else $\beta = \text{Ord}$), and then $g(\beta)$ is maximal.

To prove choice from Zorn, let \mathcal{F} be the set of all partial choice functions on X . By Zorn, \mathcal{F} has a maximal element F , and it is easy to see that F is a (total) choice function on X . \square

6.5 Cardinals

We still are working in ZF, i.e. still not assuming choice.

Definition 6.24. If κ is an ordinal such that for every ordinal α which is in bijection with κ , $\alpha \geq \kappa$, then κ is a *cardinal*, and we write $\kappa \in \text{Card}$.

Definition 6.25. Assume the axiom of choice. The *cardinality* of X , $\text{card } X$, is the unique $\kappa \in \text{Card}$ such that X and κ are in bijection.

Lemma 6.26. If A is a set of cardinals then $\bigcup A$ is a cardinal.

Proof. Clearly $\bigcup A = \sup A$, so we just have to show that $\sup A \in A$. Assume not. So there is an ordinal $\alpha < \bigcup A$ and a bijection $f : \alpha \rightarrow \bigcup A$. Then there is a $\kappa \in \bigcup A$ such that $\alpha < \kappa$ and an injection $\alpha \rightarrow \kappa$ obtained by restricting f , a contradiction. \square

We let $\aleph_0 = \omega$, $\aleph_{\alpha+1}$ be the smallest cardinal larger than \aleph_α , and let $\aleph_\gamma = \bigcup_{\alpha < \gamma} \aleph_\gamma$ for γ a limit ordinal, which is a set by replacement. In particular, \aleph_ω exists; this is the sense in which replacement is a “large cardinal axiom.” Since Ord is well-ordered, for every cardinal κ there is an ordinal α such that $\kappa = \aleph_\alpha$.

Similarly, we define $\beth_0 = \omega$, $\beth_{\alpha+1} = \text{card } \mathcal{P}\beth_\alpha$, $\beth_\gamma = \bigcup_{\alpha < \gamma} \beth_\gamma$.

Lemma 6.27. If $\kappa \geq \aleph_0$ then there is a $\alpha \in \text{Ord}$ such that $\kappa = \aleph_\alpha$.

Proof. Let α be minimal among those for which $\aleph_\alpha \geq \kappa$. Assume that $\aleph_\alpha > \kappa$. If α is a limit ordinal, then there is a $\beta < \alpha$ such that $\aleph_\beta > \kappa$, so we might as well assume $\alpha = \gamma + 1$ for some γ . Then we have $\aleph_\gamma < \kappa < \aleph_\alpha$, a contradiction. \square

We now consider the two axioms that will motivate much of the rest of what we do. They are independent of ZFC, but proving this is highly nontrivial.

Axiom 6.28 (continuum hypothesis). $\text{card } \mathbb{R} = \aleph_1$.

Axiom 6.29 (generalized continuum hypothesis). For every α , $\aleph_\alpha = \beth_\alpha$.

Obviously GCH implies CH. In the absence of choice these are a little silly, because we have no reason to believe that $\text{card } \mathbb{R}$ is even well-defined.

We now set up cardinal arithmetic. We define $\kappa + \lambda = \text{card}(\kappa \sqcup \lambda)$ and $\kappa\lambda = \text{card}(\kappa \times \lambda)$. This ends up being completely trivial.

Theorem 6.30. If κ is infinite, then

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

Proof. We have

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

The only step in this that isn't trivial is $\max(\kappa, \lambda)^2 = \max(\kappa, \lambda)$. But this is true of any infinite set. \square

On the other hand, exponentiation is highly nontrivial. We let κ^λ be the cardinality of the set of all functions $\lambda \rightarrow \kappa$.

Let κ be an infinite cardinal. Then $\kappa^\kappa = 2^\kappa$, so $\kappa^\kappa > \kappa$ by the diagonal argument. This (and one more theorem) is pretty much all we can prove about cardinal exponentiation. Just about everything can be proven independence using forcing.

More generally, if $2 \leq \kappa \leq \lambda$ and $\lambda \geq \aleph_0$, then $\kappa^\lambda = 2^\lambda$. So we might as well assume $\kappa > \lambda$ unless we are studying 2^λ .

Definition 6.31. A function $f : \alpha \rightarrow \kappa$ is *cofinal* if $f(\alpha)$ is unbounded in κ . The *cofinality* of κ , $\text{cof } \kappa$, is the least α such that there is a cofinal $f : \alpha \rightarrow \kappa$.

Obviously $\text{cof } \kappa \leq \kappa$.

Definition 6.32. A *regular cardinal* is a $\kappa \in \mathbf{Card}$ such that $\text{cof } \kappa = \kappa$. Otherwise, κ is a *singular cardinal*.

It is easy to see that $\text{cof } \aleph_0 = \text{cof } \aleph_\omega = \aleph_0$. So \aleph_0 is a regular cardinal, while \aleph_ω is singular.

Lemma 6.33. Let $\kappa \in \mathbf{Card}$. There is a strictly increasing cofinal function $\text{cof } \kappa \rightarrow \kappa$.

Proof. There is a cofinal function $f : \text{cof } \kappa \rightarrow \kappa$. Now let $g : \text{cof } \kappa \rightarrow \kappa$ be given by

$$\alpha \mapsto \max(f(\alpha), 1 + \sup_{\beta < \alpha} g(\beta)).$$

Then g is strictly increasing and cofinal. \square

Lemma 6.34. Suppose $f : \kappa \rightarrow \lambda$ is cofinal and nondecreasing. Then $\text{cof } \kappa = \text{cof } \lambda$.

Proof. For $\text{cof } \lambda \leq \text{cof } \kappa$, let $g : \text{cof } \kappa \rightarrow \kappa$ be cofinal, then $f \circ g$ is cofinal. For the converse, let $g : \text{cof } \lambda \rightarrow \lambda$ be cofinal and strictly increasing (possible by the above lemma). We let $h : \text{cof } \lambda \rightarrow \kappa$ be given by sending α to the least β such that $f(\alpha) > g(\beta)$, which is possible since f is cofinal. Then h is cofinal since g is. \square

In particular, $\text{cof cof } \kappa = \text{cof } \kappa$. It's not hard to show that $\text{cof } \kappa$ is a cardinal, so $\text{cof } \kappa$ is a regular cardinal.

Theorem 6.35. Assume choice. If κ is a successor cardinal then κ is regular.

Proof. Assume κ is the successor of λ . If κ is singular then there is a cofinal function $f : \lambda \rightarrow \kappa$. For each $\alpha \in \lambda$ we have $f(\alpha) < \kappa$, so $\text{card } f(\alpha) \leq \lambda$, which is well-defined by choice. So there is a surjective function $g_\alpha : \lambda \rightarrow f(\alpha)$. Now let $g(\beta) = \bigcup_\alpha g_\alpha(\beta)$. Then g is surjective and maps onto κ , a contradiction. \square

Now we can prove SOMETHING about cardinal exponentiation.

Theorem 6.36. If $\kappa \geq \aleph_0$ then $\kappa^{\text{cof } \kappa} > \kappa$.

Proof. Assume $f : \text{cof } \kappa \rightarrow \kappa$ is cofinal and $\kappa^{\text{cof } \kappa} = \kappa$. Enumerate $\kappa^{\text{cof } \kappa}$ as g_α , for $\alpha < \kappa$. So define $g : \text{cof } \kappa \rightarrow \kappa$ by sending β to the least element of κ such that $g_\alpha(\beta)$ does not have $\alpha < f(\beta)$. Then $g(\beta) \neq g_\alpha(\beta)$ for any $\alpha < f(\beta)$. Since f is cofinal, $g \neq g_\alpha$ for any α . So we have diagonalized. \square

We now arrive at a large cardinal axiom.

Definition 6.37. A *weakly inaccessible cardinal* κ is a regular limit cardinal such that $\kappa > \aleph_0$. An *inaccessible cardinal* is a weakly inaccessible cardinal κ such that if $\lambda < \kappa$ then $2^\lambda < \kappa$.

If κ is weakly inaccessible, then $\aleph_\kappa = \kappa$. But κ is much larger than the least \aleph -fixed point, since the limit λ of $\aleph_0, \aleph_{\aleph_0}, \aleph_{\aleph_{\aleph_0}}, \dots$ is an \aleph -fixed point, yet $\lambda \leq \kappa$ and $\text{cof } \lambda = \aleph_0$, so $\lambda < \kappa$. Moreover, if κ is inaccessible and choice holds, then V_κ is closed under every set-theoretic operation, so (V_κ, \in) is a model of ZFC. (Weakly inaccessible cardinals already give models of ZFC minus power set, since inaccessibility gives replacement.) So their consistency strength is much stronger than that of ZFC, since they prove the consistency of ZFC.

However, it is often useful (e.g. in the foundations of category theory) to assume the following large cardinal axiom.

Axiom 6.38 (Grothendieck). There is a proper class of inaccessible cardinals.

Chapter 7

Constructibility

7.1 Definability

We introduce a notion of quantifier complexity for LST.

Definition 7.1. A formula has *bounded quantifiers* if every quantification only ranges over sets.

For example, $\forall y \in x(y \neq y)$ has bounded quantifiers (and defines \emptyset).

Definition 7.2. A formula φ is Σ_0 and Π_0 if it has bounded quantifiers. It is Σ_{n+1} if it can be expressed as $\exists x_1 \exists x_2 \dots \exists x_k \psi$ for some ψ which is Π_n . It is Π_n if it can be expressed as $\neg\psi$ for some ψ which is Σ_n . It is Δ_n if it is both Σ_n and Π_n .

Theorem 7.3. Suppose $M \subseteq N \subseteq V$ is a chain of transitive models. Suppose ψ is a formula and $a \in M^n$. Then:

1. If $\psi \in \Delta_0$ then $M \models \psi(a)$ iff $N \models \psi(a)$.
2. If $\psi \in \Sigma_1$ and $M \models \psi(a)$ then $N \models \psi(a)$.
3. If $\psi \in \Pi_1$ and $N \models \psi(a)$ then $M \models \psi(a)$.

Proof. In the Δ_0 case, if $a_1, a_2 \in M$ then $M \models a_1 \in a_2$ iff $a_1 \in a_2$ iff $N \models a_1 \in a_2$. Similarly for $a_1 = a_2$. Something similar happens here for Σ_1 and Π_1 but obviously it only works in one direction.

Now suppose that the theorem holds for φ and ψ . Then obviously the theorem holds for $\neg\psi$ and $\varphi \wedge \psi$. Finally, we check the theorem for \exists . Note that $M \models \exists x \in a_1 \psi(x, a)$ iff there is a $b \in a_1$ such that $b \in M$ and $M \models \psi(b, a)$. Similarly for N . Here upward absoluteness follows because $M \subseteq N$. For downward absoluteness, we know $b \in a_1$, $b \in N$, $a_1 \in M$ and must show $b \in M$. This follows because M is transitive. \square

Definition 7.4. Let M be a set. Let Def^M , the *definable power set* of M , be the set of $X \subseteq M$ which are definable with parameters taken from M .

The definable power set exists, by the axiom of power set. Everything in M is definable from M ; take $x = m$ to define $m \in M$. So $M \subseteq \text{Def}^M \subseteq \mathcal{P}M$. If M is finite then $\text{Def}^M = \mathcal{P}M$; if M is infinite then $\text{card } \text{Def}^M = \aleph_0 \text{ card } M = \text{card } M$.

7.2 The reflection principle

Definition 7.5. Suppose κ is a regular uncountable cardinal. A set $S \subseteq \kappa$ is a *club set* if for every $\alpha < \kappa$ there is a $\beta \in S$ such that $\beta > \alpha$ and if $\alpha = \sup(S \cap \alpha)$ then $\alpha \in S$.

So S is closed and unbounded in the topology of κ .

Recall that $M \preceq N$ means that M is an elementary substructure of N .

Theorem 7.6. Let κ be a regular uncountable cardinal and suppose that we have a chain of LST-models $M_\alpha = (M_\alpha, E_\alpha)$, $\alpha < \kappa$, so $M_\alpha \subseteq M_\beta$ whenever $\alpha < \beta$. Suppose that for every $\alpha < \kappa$, $\text{card } M_\alpha < \kappa$, and for every limit $\beta < \kappa$, $M_\beta = \bigcup_{\alpha < \beta} M_\alpha$.

Let $M = \bigcup_{\alpha < \kappa} M_\alpha$, $E = \bigcup_{\alpha < \kappa} E_\alpha$, so $M = (M, E)$ is the injective limit of the M_α . Let $S \subseteq \kappa$ be the set of α such that $M_\alpha \preceq M$. Then S is a club set.

In particular, for any $\alpha, \beta \in S$, if $\alpha < \beta$ then $M_\alpha \preceq M_\beta$.

Proof. To see that S is closed, let β be a limit point of S . We must show $M_\beta \preceq M$, so let ψ be a formula, $b \in M_\beta^n$. Then we can find a $\alpha < \beta$ such that $b \in M_\alpha^n$ and $M_\alpha \preceq M$, since β is a limit point.

Suppose that there is a $a \in M$ such that $M \models \psi(a, b)$. By the Tarski-Vaught test, there is a $a_\alpha \in M_\alpha$ such that $M_\alpha \models \psi(a_\alpha, b)$. Since M_β is a limit, we can take α so large as to guarantee $M_\beta \models \psi(a_\alpha, b)$. Therefore $M_\beta \preceq M$ by the Tarski-Vaught test. Therefore S is closed.

Given ψ , let $f_\psi : \kappa \rightarrow \kappa$ send α to the least β such that for all $b \in M_\alpha^n$, if there is a $a \in M$ such that $M \models \psi(a, b)$ then there is a $a_\beta \in M_\beta$ such that $M_\beta \models \psi(a_\beta, b)$.

To see that f_ψ is well-defined, fix ψ and α and suppose that no such β exists. There are $|M_\alpha|^n = |M_\alpha| < \kappa$ choices of b . Let β_b be a witness to the failure of the theorem for b . Then $b \mapsto \beta_b$ is cofinal $M_\alpha^n \rightarrow \kappa$, so κ is singular.

Let $g : \kappa \rightarrow \kappa$ be defined by

$$g(\alpha) = \sup_\psi f_\psi(\alpha).$$

The mapping $\psi \mapsto f_\psi(\alpha)$ is cofinal $\omega \rightarrow \kappa$ if it is unbounded; since κ is regular, it follows that the mapping is bounded, so g is well-defined.

Suppose that g restricts to a function $\alpha \rightarrow \alpha$. By definition of g and the Tarski-Vaught test, for every $\beta < \alpha$ and $b \in M_\beta^n$, if there is a $a \in M$ such that $M \models \psi(a, b)$ then there is a $a' \in M_{g(\beta)}$ such that for some $\gamma \leq g(\beta)$, $M_\gamma \models \psi(a', b)$, hence $M_\alpha \preceq M$.

Now g is an increasing function, so for any α_0 the ordinal

$$\alpha = \sup_{n < \omega} g^{\circ n}(\alpha_0)$$

is closed under g and is larger than α_0 . Taking α_0 arbitrarily large we see that α is arbitrarily large, but $\alpha \in S$, so S is unbounded.

The last paragraph is obvious. □

Theorem 7.7 (reflection). The class of ordinals α such that $V_\alpha \prec V$ is a club.

Proof. The V_α , $\alpha \in \text{Ord}$, meet the hypotheses of the above theorem (this is obvious if ZFC is consistent, and otherwise it follows by the principle of explosion). □

So it is impossible to distinguish V ! Anything that we can say about it was already true for a club class of submodels.

7.3 The constructible sets

We now consider those sets which are constructible from the ordinals.

Definition 7.8. Let $L_0 = 0$, $L_{\alpha+1} = \text{Def}^{L_\alpha}$, and $L_\gamma = \bigcup_{\alpha < \gamma} L_\alpha$ for γ a limit ordinal. Let $L = \bigcup_{\alpha \in \text{Ord}} L_\alpha$. If $x \in L$, we say that x is a *constructible set*.

By induction, one can easily check that L_α is a transitive set and $L_\alpha \subset L_\beta$ for $\alpha < \beta$. Moreover, we have $L_\alpha \subseteq V_\alpha$, and every finite subset of L_α lies in $L_{\alpha+1}$. Moreover, for $\alpha \leq \omega$, we have $L_\alpha = V_\alpha$. But, on the other hand, if $\alpha \geq \omega$ then $\text{card } L_\alpha = \aleph_0(\text{card } \alpha) = \text{card } \alpha$ so $\text{card } L_\alpha = \text{card } \alpha$, which usually (but not always) implies $L_\alpha \neq V_\alpha$.

Theorem 7.9. L is a model of ZF.

Proof. Since L is a transitive class, extensionality and foundation hold. Clearly $\emptyset, \omega \in L$ so empty set and infinity hold. Pairing, union, replacement, and power set follow from the definition of definability. The only nontrivial axiom is replacement. Suppose that for every $A, w \in L$ and every $x \in A$ there is a unique $y \in L$ such that $\phi^L(x, y, A, w)$. Let α be the sup, taken over the ranks (in L) of all $y \in L$ such that there is an $x \in A$ with $\phi^L(x, y, A, w)$.

Let $Y = L_{\alpha+1}$; then $Y \in L$ and for every $y \in Y$ such that there is a $x \in A$ such that $\phi^L(x, y, A, w)$, $y \in Y$. Therefore $y \in L$, so L is a model of replacement. \square

Axiom 7.10 ($V = L$). Every set is constructible.

In other words, $V_\alpha = L_\alpha$ for every $\alpha \in \text{Ord}$.

Since the notion of “definability” is absolute, and **Ord** is absolute to L , it follows that L_α is absolute to L . In particular, L proves that for every $x \in V$, $x \in L$, so $L \models (V = L)$.

Theorem 7.11. If ZF is consistent, then $V = L$ is consistent.

Proof. If V exists, then so does L , and L is a model of $V = L$. \square

In fact, L is in some sense the minimal inner model of ZF.

Lemma 7.12. Let M be a transitive proper class, and suppose M is a model of ZF. Then **Ord** is absolute to M , and **Ord** is a subclass of M .

Proof. The rank of a set x is defined by recursion on x . Therefore rank is absolute to M . If $\alpha \in \text{Ord}$, M is not contained in V_α since M is a proper class. So there is a $x \in M \setminus V_\alpha$ whose rank β has $\beta \geq \alpha$; therefore $\beta \in M$, and since M is transitive, $\beta \in M$. So $\text{Ord}^M = \text{Ord}$. \square

Theorem 7.13. Let M be a transitive proper class, and suppose M is a model of ZF. Then L is absolute to M , and L is a subclass of M .

Proof. Since **Ord** is absolute to M , so is L_α . Therefore $L^M = L$. \square

We now show that L satisfies an especially strong form of the axiom of choice.

Axiom 7.14 (global choice). There is a well-ordering of V .

In fact, the axiom of global choice is equivalent to the existence of a class bijection $V \rightarrow \mathbf{Ord}$. Clearly such a bijection gives a well-ordering, and conversely, if there is a well-ordering $<$ of V , then every ordinal must embed into (V, \in) , which is only possible if $(V, \in) \cong \mathbf{Ord}$ by definition of \mathbf{Ord} . The axiom of global choice implies the axiom of choice, since it restricts to a well-ordering of any set.

Theorem 7.15. L is a model of global choice, hence of ZFC. Stronger, its well-ordering is definable.

Proof. We will define a well-ordering $<_\alpha$ of L_α as follows. Let $<_0 = 0$ and let the well-ordering of a limit ordinal be the limit of the well-orderings. Let $<_\alpha^n$ denote the lexicographic ordering on L_α^n . Fix a definable enumeration E_n of the set of definable n -ary relations on R .

For $x \in L_{\alpha+1}$ let n_x denote the least n such that there is a $s \in L_\alpha^n$ and a definable $(n+1)$ -ary relation R on L_α such that $x = \{y \in L_\alpha : (s, y) \in R\}$. Let s_x be the $<_\alpha^{n_x}$ th least witness to the definition of n_x . Let m be the least index of R in E_{n+1} .

For $x, y \in L_{\alpha+1}$ define $x <_\alpha y$ iff $x, y \in L_\alpha$ and $x <_\alpha y$, $x \in L_\alpha$ and $y \notin L_\alpha$, or $x, y \notin L_\alpha$ and either $n_x < n_y$, or if they are equal then $s_x <_\alpha^{n_x} s_y$, or if they are also equal then $m_x < m_y$.

If $n_x = n_y$, $s_x = s_y$, and $m_x = m_y$, then x and y are defined by the same relation so $x = y$ by extensionality. By induction, $<_\alpha$ is a well-ordering of L_α .

Now write $L = L_0 \cup (L_1 \setminus L_0) \cup (L_2 \setminus L_1) \cup \dots$ and order the α th entry in the above disjoint union decomposition using $<_\alpha$. \square

Corollary 7.16. ZF with the axiom of global choice, and hence ZFC, is consistent if ZF is.

Corollary 7.17. If ZF is consistent, then it is consistent that \mathbb{R} has a definable well-ordering.

However, we cannot prove that such a well-ordering is actually a well-ordering of \mathbb{R} .

Theorem 7.18. If $V = L$, then for every κ , $2^\kappa = \kappa^+$.

We prove this in the chapter on one measurable cardinal; this follows from Solovay's theorem. The proof is exactly the same except with U removed.

Chapter 8

Forcing

The idea behind forcing is that given a model M of ZFC, we want to adjoin a “generic” set $x \notin M$ so that the smallest model $M[x]$ containing M and x has a certain property p . In this chapter we’ll go through this procedure with $p = (2^{\aleph_0} = \aleph_2)$.

8.1 Generic filters

Let M be a transitive model of ZFC. (This might not exist, since the existence of transitive models is strictly stronger than the existence of models, which implies that ZFC is consistent by soundness, but we can remove the transitivity assumption later.) It will be frequently be useful to assume that M is countable, which is always possible up to elementary equivalence by the Lowenheim-Skolem theorem.

Let $\mathbb{P} \in M$ be a poset, and let $1_{\mathbb{P}} \in \mathbb{P}$ be the maximum of \mathbb{P} .

Definition 8.1. An element $x \in \mathbb{P}$ is called a \mathbb{P} -*condition*, and a subset $D \subseteq \mathbb{P}$ is *dense* if for every $p \in \mathbb{P}$ there is a $d \in D$ with $d \leq p$.

Proper dense subsets exist, because any downward slice of \mathbb{P} is dense.

Definition 8.2. If $x, y \in \mathbb{P}$ are such that there is $z \in \mathbb{P}$ with $z \leq x$ and $z \leq y$, then x and y are *compatible conditions* and we write $x \parallel y$. Otherwise, we write $x \perp y$.

If $X \subseteq \mathbb{P}$ is such that every pair in X is incompatible then we say that X is an *antichain*.

Definition 8.3. Let $G \subseteq \mathbb{P}$ be a filter. We say that G is an M -*generic filter* if for every $D \in M$ which is dense in \mathbb{P} , $D \cap G$ is nonempty.

Since we think of $x, y \in \mathbb{P}$ as conditions, the ordering \leq is an ordering by logical strength. Indeed, $1_{\mathbb{P}}$ is the maximal and hence weakest condition (since G is a filter, $1_{\mathbb{P}} \in G - G$ is nonempty since \mathbb{P} is dense). If $x \leq y$ then x is a weaker condition than y (since $y \in G$ implies $x \in G$). Since G is a filter, any pair in G is compatible.

Notice that since \mathbb{P} is in general infinite, its power set $\mathcal{P}\mathbb{P}$ is not absolute. That is, if N is a model of ZFC, $(\mathcal{P}\mathbb{P})^M$ might disagree with $(\mathcal{P}\mathbb{P})^N$. That is why it makes sense to define genericity with respect to a particular model M .

Lemma 8.4. For every dense set $D \subseteq \mathbb{P}$ there is a maximal antichain in D .

Proof. Zorn's lemma. □

Theorem 8.5 (Rasiowa-Sikorski). If M is countable, then there is an M -generic filter in \mathbb{P} .

This turns out to be the same as the Baire category theorem (or, equivalently, the axiom of dependent choice). This motivates the terminology “dense” and “generic.”

Proof. Since M is countable, the set \mathcal{D} of dense sets in \mathbb{P} is countable, say $\mathcal{D} = \{D_n : n \in \omega\}$. Let $p_1 \in D_1$. We can inductively choose $p_{n+1} \leq p_n$ such that $p_{n+1} \in D_{n+1}$ since D_{n+1} is dense. Now let G be the smallest filter containing the sequence of p_n . Clearly G meets every D_n . □

Definition 8.6. \mathbb{P} is *splitting* for every $x \in \mathbb{P}$, there are $y, z \leq x$ with $y \perp z$.

That is, for every condition, there are two incompatible “possible futures.”

Lemma 8.7. If \mathbb{P} is splitting and $F \subseteq \mathbb{P}$ is a filter, then the ideal F^c is dense.

Proof. If $x \in F$ then there are $y, z \leq x$ which are incompatible. But every pair in F is compatible, so one of them lies in F^c . Therefore F^c is dense. □

Theorem 8.8. If \mathbb{P} is splitting and $G \subseteq \mathbb{P}$ is a M -generic filter then $G \notin M$.

Proof. Suppose not. Then by the lemma, G^c is dense, but since $G \in M$, $G^c \in M$. Since G is M -generic, G meets every dense set, so $G \cap G^c$ is nonempty, which is absurd. □

Example 8.9. Let \mathbb{P} be the infinite binary tree, which is splitting. Then a filter is a branch, and a generic set is an infinite path. So the generic sets are exactly the infinite binary sequences 2^ω .

8.2 Constructing the generic extension

Recall that M is a transitive model of ZF, and $\mathbb{P} \in M$ is a poset with a maximum $1_{\mathbb{P}} \in \mathbb{P}$ and an filter $G \subseteq \mathbb{P}$. We will eventually assume that \mathbb{P} is splitting and G is M -generic, so that $G \notin M$, but not yet.

We now construct a class $\text{Name}(\mathbb{P})$ from \mathbb{P} , which admits an injection $V \rightarrow \text{Name}$, by transfinite recursion. First, $\text{Name}_0(\mathbb{P}) = \emptyset$. If $\text{Name}_\alpha(\mathbb{P})$ is defined then

$$\text{Name}_{\alpha+1}(\mathbb{P}) = \mathcal{P}(\text{Name}_\alpha(\mathbb{P}) \times \mathbb{P}).$$

Finally we take unions at limit stages, and let $\text{Name}(\mathbb{P}) = \bigcup_\lambda \text{Name}_\lambda(\mathbb{P})$.

Definition 8.10. A \mathbb{P} -name is an element of $\text{Name}(\mathbb{P})$. The *rank* of a \mathbb{P} -name τ is the least λ such that $\tau \in \text{Name}_\lambda(\mathbb{P})$.

We define the *G -interpretation* τ^G of a \mathbb{P} -name τ by transfinite recursion. Namely, τ^G is the set of all interpretations σ^G such that the \mathbb{P} -name σ has $(\sigma, p) \in \tau$ with $p \in G$.

We think of the set of interpretations

$$X = \{\tau^G : \tau \in \text{Name}_\lambda(\mathbb{P})\}$$

as the “expansion of V_λ by G ”, so “ $X = V_\lambda[G]$ ”.

Example 8.11. We always have

$$\text{Name}_2(\mathbb{P}) = \mathcal{P}(\{\emptyset\} \times \mathbb{P}) = \{(\emptyset, p) : p \in \mathbb{P}\}.$$

So $\text{Name}_2(\mathbb{P})$ can be identified with $\mathcal{P}\mathbb{P}$. Therefore the interpretation set

$$X = \{\tau^G : \tau \in \text{Name}_2(\mathbb{P})\}$$

has each element

$$\tau^G = \{\emptyset : (\emptyset, p) \in \tau, p \in G\}$$

which means $\tau^G = \{\emptyset\}$ iff there is a $p \in G$ (and $\tau^G = \emptyset$ otherwise). So $X = V_2$. In fact this is true for any Name_n and any G .

Now we can define the extension.

Definition 8.12. The *generic extension*

$$M[G] = \{\tau^G : \tau \in \text{Name}^M(\mathbb{P})\}.$$

Theorem 8.13. $M \subseteq M[G]$ and $G \in M[G]$. Moreover, $M[G]$ is transitive.

Notice that while $\text{Name}^M(\mathbb{P}) \subseteq M$, the interpretations of the \mathbb{P} -names may not live in M . For example, suppose that G actually is M -generic and \mathbb{P} is splitting. Then $G \notin M$, even though $G \in M[G]$. So we truly have “adjoined G to $M[G]$.”

Proof. We need to give an injection $M \rightarrow M[G]$. So we define the \mathbb{P} -names

$$\hat{x} = \{(\hat{y}, 1_{\mathbb{P}}) : y \in x\}$$

by transfinite recursion. The map $x \mapsto \hat{x}$ is then an injection $M \rightarrow \text{Name}^M(\mathbb{P})$. Finally, we define

$$\dot{G} = \{(\hat{p}, p) : p \in \mathbb{P}\}.$$

Since these are \mathbb{P} -names, we consider their interpretations. First,

$$\hat{x}^G = \{y : (\hat{y}, 1_{\mathbb{P}}) \in \hat{x}\}$$

by transfinite induction. So we have $M \subseteq M[G]$. Moreover,

$$\dot{G}^G = \{\sigma^G : (\sigma, p) \in \dot{G}, p \in G\} = \{\hat{p}^G : (\hat{p}, p), p \in \mathbb{P}\} = \{p : p \in G\} = G.$$

So indeed, $M \cup \{G\} \subseteq M[G]$.

Now we show that $M[G]$ is transitive. Let $x \in M[G]$ and $y \in x$. Then there is a $\sigma \in \text{Name}^M(\mathbb{P})$ such that $\sigma^G = x$. So $y = \tau^G$ for some $\tau \in \text{Name}^M(\mathbb{P})$. Then

$$\tau \in \sigma \in M.$$

Since M is transitive, $\tau \in M$. Therefore $y \in M[G]$, so $M[G]$ is transitive. \square

Theorem 8.14. $M[G]$ is a model of extensionality, empty set, infinity, pairing, foundation, and union. If M is a model of choice then so is $M[G]$.

So $M[G]$ doesn't quite model ZF, but will happen when we assume that G is actually M -generic.

Proof. Extensionality and foundation follow because $M[G]$ is transitive; empty set and infinity follow because $M \subseteq M[G]$.

For pairing, assume $x = \sigma^G$ and $y = \tau^G$. Then $\{x, y\}^{M[G]} = \{(\sigma, 1_{\mathbb{P}}), (\tau, 1_{\mathbb{P}})\}^G$ so we're good. Union is similar, and same with choice – just lift infinite products from $M[G]$ to $\text{Name}^M(\mathbb{P})$. \square

8.3 Forcing semantics

We now define a relation \Vdash . The idea is that if $p \in \mathbb{P}$ is a condition, then $p \Vdash \varphi(\sigma_1, \dots, \sigma_n)$ iff $M[G] \models \varphi(\sigma_1^G, \dots, \sigma_n^G)$.

Definition 8.15. Let M be a transitive model of ZF and let $\mathbb{P} \in M$ be a poset. For $p \in \mathbb{P}$ and φ a LST-formula (possibly with variables), write $p \Vdash \varphi$ to mean:

1. If $\varphi = (\tau_1 = \tau_2)$, then for each $(\sigma_1, q_1) \in \tau_1$, let $D_{\sigma_1}^{q_1}$ be the set of all $r \in \mathbb{P}$ such that if $r \leq q_1$, then there is a $(\sigma_2, q_2) \in \tau_2$ such that $r \leq q_2$ and such that $r \Vdash (\sigma_1 = \sigma_2)$, and let $D_{\sigma_2}^{q_2}$ be defined similarly. Then $D_{\sigma_i}^{q_i}$ are dense below p for $i \in \{1, 2\}$.
2. If $\varphi = (\tau_1 \in \tau_2)$, then let D be the set of all $q \in \mathbb{P}$ such that there is $(\tau, r) \in \tau_2$ such that $q \leq r$ and $q \Vdash (\tau = \tau_1)$. Then D is dense below p .
3. If $\varphi = (\exists x \psi(x))$ then let D be the set of all $q \in \mathbb{P}$ such that there is τ with $q \Vdash \psi(\tau)$. Then D is dense below p .
4. If $\varphi = \psi \vee \chi$, then $p \Vdash \psi$ and $p \Vdash \chi$.
5. If $\varphi = \neg\psi$, then $p \nVdash \psi$.

If $p \Vdash \varphi$, then we say that p forces φ .

Notice that $D_{\sigma_1}^{q_1}$ encodes $\tau_1 \subseteq \tau_2$. Similarly for the other dense sets we defined.

Also notice that even though \Vdash is a relation in V , \Vdash is definable from M , since all we have referred to are \mathbb{P} -names, which already lie in M .

We now show that \Vdash behaves “modally” as we desire. First, a very tedious induction that we skip shows the following.

Lemma 8.16. Let \Vdash be as above. Then:

1. If $p \Vdash \varphi$ and $q \leq p$ then $q \Vdash \varphi$.
2. If the set of $q \in \mathbb{P}$ such that $q \Vdash \varphi$ is dense below p then $p \Vdash \varphi$.
3. If $p \nVdash \varphi$ then there is a $p' \leq p$ such that $p' \Vdash \neg\varphi$.

Interpreting \Vdash modally, we think of $p \Vdash \varphi$ to mean that p implies that φ will be eventually true, and $q \leq p$ to mean that q is a possible future state reachable from p . Then, the first condition means that if it is known that φ will eventually be true, then it will continue to be known that φ will be eventually be true. The second condition means that if in every possible future, we will learn that φ will be eventually true, then φ is already known to be eventually true in every possible future. The final condition means that if φ is not known to be true then there is a possible future where it is false.

Theorem 8.17. Suppose M is a transitive model of ZF, $\mathbb{P} \in M$ is a poset, $G \subseteq \mathbb{P}$ is an M -generic filter, and $n \in \omega$. For each LST-formula φ with n free variables and each $\sigma_1, \dots, \sigma_n \in \text{Name}^M(\mathbb{P})$,

$$M[G] \models \varphi(\sigma_1^G, \dots, \sigma_n^G)$$

if and only if there is a condition $p \in G$ such that

$$p \Vdash \varphi(\sigma_1, \dots, \sigma_n).$$

Corollary 8.18. Let M and G be as above. Then $M[G]$ is a model of ZF, and if M is a model of ZFC then so is $M[G]$.

Proof. We just need to check comprehension and replacement, and we'll just do comprehension since replacement is similar. For convenience we suppress parameters. Assume $\sigma^G \in M[G]$ and φ is an LST-formula with a free variable. We need to show that the set

$$A = \{x \in \sigma^G : \varphi(x)\} \in M[G].$$

This can be rewritten as, for some $\rho \in \text{Name}^M(\mathbb{P})$,

$$A = \{\rho^G \in \sigma^G : \exists p \in G (p \Vdash \rho \in \sigma \wedge \varphi(\rho))\}.$$

Now let

$$\tau = \{(\rho, p) \in M \times \mathbb{P} : p \Vdash \varphi(\rho)\}.$$

Then obviously $\tau^G = A$ and $\tau \in V$. Since M is a model of comprehension and \Vdash is definable in M , it follows that $\tau \in \text{Name}^M(\mathbb{P})$. So $A \in M[G]$. \square

8.4 Cardinal collapse

We set up some machinery for later.

Definition 8.19. Let C be an uncountable set of finite sets. Then C is a Δ -system if there is an $R \in C$ such that for every distinct pair $x, y \in C$, $x \cap y = R$.

Lemma 8.20. If C is an uncountable set of finite sets, then there is a Δ -system $C' \subseteq C$.

Proof. By the infinite pigeonhole principle, we can find $n \in \omega$ such that uncountably many sets in C have cardinality n , and we replace C with the set of all sets in C with cardinality n . Obviously if $n = 1$ then $R = \emptyset$ and we're done.

Assume that the lemma is true for every $k < n$. If there is an $c \in C$ contained in uncountably many sets, then add c to C' and apply the inductive hypothesis. Otherwise, every $c \in C$ only appears in countably many sets. Now do the usual Baire category theorem trick. \square

Definition 8.21. A poset \mathbb{P} has the *countable chain condition* (or is *ccc*) if every antichain in \mathbb{P} is countable.

Theorem 8.22 (possible values). Let M be a transitive model of ZFC and $\mathbb{P} \in M$ be a poset with an M -generic filter G . Let $X, Y \in M$ and let $f \in M[G]$ be a function $X \rightarrow Y$.

If M is a model of “ \mathbb{P} is ccc,” then there is a function $F \in M$ such that for each $x \in X$, $f(x) \in F(x)$ and M is a model of “ $F(x)$ is countable”.

Let $x \in X$. Notice that we do not assume $f \in M$, so M does not know the value $f(x)$, even though $f(x) \in M$. We think of $F(x)$ as the “possible values of f that M allows”, so M does have some control over its generic extension $M[G]$, ruling out but all but countably many possibilities.

Proof. Let $\dot{f} \in \text{Name}^M(\mathbb{P})$ be such that $\dot{f}^G = f$. Then there is a $p \in \mathbb{P}$ which forces “there exists a function $\dot{f} : \dot{X} \rightarrow \dot{Y}$ ”.

Since M thinks \mathbb{P} is ccc and thinks AC is true, for each $x \in X$ we can use Zorn’s lemma to find a maximal set $A(x)$ of incompatible conditions $q \leq p$ such that for each $q \in A(x)$ there is a $y \in Y$ such that q forces “ $\dot{f}(\dot{x}) = \dot{y}$,” and $A(x)$ is countable.

Each of the $q \in A(x)$ forces $f(x)$ to admit a certain value, say y_q . So we define

$$F(x) = \{y \in Y : \exists q \leq p (y = y_q)\}.$$

By maximality of $A(x)$, $F(x)$ must hit every possible value of y in every possible forcing extension. \square

Definition 8.23. Let M be a transitive model of ZFC. A poset $\mathbb{P} \in M$ *preserves cardinals* if for every M -generic filter $G \subseteq \mathbb{P}$ and each ordinal λ , M is a model of “ $\lambda \in \mathbf{Card}$ ” if and only if $M[G]$ is a model of “ $\lambda \in \mathbf{Card}$ ”.

Theorem 8.24. Let M be a transitive model of ZFC, $\mathbb{P} \in M$ a poset, and M is a model of “ \mathbb{P} is ccc”. Then \mathbb{P} preserves cardinals.

Proof. Taking limits, it suffices to check that \mathbb{P} preserves *regular cardinals*; namely M is a model of “ λ is a regular cardinal” if and only if $M[G]$ is a model of “ λ is a regular cardinal.”

Let $\lambda \geq \aleph_2^M$ be an M -regular cardinal, and suppose that λ is not $M[G]$ -regular. Then there is a function $f \in M[G]$, $\bar{\lambda} \rightarrow \lambda$ cofinal, for some $\bar{\lambda} < \lambda$. By the possible values argument, there is a function $F \in M$, $\bar{\lambda} \rightarrow 2^\lambda$, such that $f(\alpha) \in F(\alpha)$ and

$$\text{card}^M F(\alpha) \leq \aleph_0^M < \lambda.$$

Define a map $\bar{\lambda} \rightarrow \lambda$ by $g(\alpha) = \sup F(\alpha)$. Then M is a model of “ g is cofinal”, even though λ is regular in M , a contradiction. \square

8.5 Forcing $2^{\aleph_0} = \aleph_2$

Our goal is to prove the following theorem.

Theorem 8.25 (Cohen). There is a model $M[G]$ of ZFC such that $M[G] \models \neg \text{CH}$.

We have two problems forcing $2^{\aleph_0} = \aleph_2$. First, we need to find an M and a \mathbb{P} such that for each M -generic filter $G \subseteq \mathbb{P}$, $M[G]$ has \aleph_2 more real numbers than M (and by constructing M from L , we can arrange for $M[G]$ to have $\aleph_2 + \aleph_1 = \aleph_2$ many reals). Second, we need to show that $\aleph_2^M = \aleph_2^{M[G]}$.

Definition 8.26. For $x \in V$, $\text{Add}(x)$ is the set of partial functions $x \rightarrow 2$, ordered by reverse inclusion.

So $p \leq q$ iff $\text{dom}(p) \supseteq \text{dom}(q)$, and $1_{\text{Add}(x)}$ is the empty function.

Lemma 8.27. For each $\kappa \in \text{Card}$, $\text{Add}(\kappa \times \omega)$ is ccc.

Proof. If not, let $\{p_\alpha\}$ be an antichain of length \aleph_1 , and let $C = \{\text{dom}(p_\alpha)\}$. Then $\text{card } C = \aleph_1$, yet C consists only of finite sets, so there is a Δ -system in C with root R . Let $B = \{p_\alpha : \text{dom}(p_\alpha) \in R\}$. Then each $p_\alpha \in B$ determines a total function $R \rightarrow \{0, 1\}$ and R is finite, so $\mathcal{P}R$ is finite, yet there are supposed to be uncountably many p_α . \square

Theorem 8.28. Let M be a countable, transitive model of ZFC + CH, and let

$$\mathbb{P} = (\text{Add}(\aleph_2 \times \omega))^M.$$

Let G be an M -generic filter. Then $M[G]$ is a model of ZFC + “ $2^{\aleph_0} = \aleph_2$ ”.

Notice that \mathbb{P} is splitting. Moreover, such a G exists by the Baire category theorem, since M is transitive, so $G \notin M$.

Proof. By thinking of G as a “measure” on $\aleph_2 \times \omega$, we get a function $\hat{G} : \aleph_2 \times \omega \rightarrow 2$. For each $\alpha \in \aleph_2$, we define

$$G_\alpha = \{n \in \omega : \hat{G}(\alpha, n) = 0\}.$$

Given $\alpha, \beta \in \aleph_2$, the set

$$D = \{q \in \mathbb{P} : \exists n \in \omega (q(\alpha, n) \neq q(\beta, n))\}$$

is dense, so D meets G . Therefore the map $\alpha \mapsto G_\alpha$ is injective. The G_α interpret to give new elements of $\mathcal{P}\omega = 2^{\aleph_0}$ in $M[G]$. So $\text{card } M[G] = \aleph_2^M$, and it remains to show $\aleph_2^M = \aleph_2^{M[G]}$. This happens because \mathbb{P} is ccc in M . \square

Chapter 9

Small large cardinals

Definition 9.1. A *regular cardinal* is an infinite cardinal κ such that for any set Λ of cardinals such that $\text{card } \Lambda < \kappa$ and for every $\lambda \in \Lambda$, $\lambda < \kappa$,

$$\sum_{\lambda \in \Lambda} \lambda < \kappa.$$

A *singular cardinal* is an infinite cardinal which is irregular.

Informally, a large cardinal is a regular cardinal which is a witness to some existential axiom Φ such that Φ proves that ZFC is consistent. So the following definition is the smallest example of a large cardinal.

Definition 9.2. A *worldly cardinal* is a regular cardinal κ such that V_κ , the set of all sets of rank $< \kappa$, is a model of ZFC.

9.1 Regular cardinals

For f a function and $X \subseteq \text{dom } f$ a set, let

$$f'' X = \{f(x) : x \in X\}$$

be the image of X under f .

Definition 9.3. Let α be an ordinal. We say that $X \subseteq \alpha$ is *cofinal* if for every $\beta < \alpha$ there is a $\gamma \in X$ such that $\beta < \gamma < \alpha$. If $f : \beta \rightarrow \alpha$ is a function and $f'' \beta$ is cofinal, we say that f itself is *cofinal*.

Definition 9.4. Let α be an ordinal. The *cofinality* of α , $\text{cof } \alpha$, is the least $\beta < \alpha$ such that there is a cofinal function $\beta \rightarrow \alpha$.

We recall that if Λ is a set of infinite cardinals, or $\text{card } \Lambda \geq \aleph_0$, we have

$$\sum_{\lambda \in \Lambda} \lambda = \sup_{\lambda \in \Lambda} \lambda,$$

by triviality of cardinal arithmetic. In particular, an infinite cardinal κ is regular iff $\text{cof } \kappa = \kappa$. As a consequence, since $\text{cof cof} = \text{cof}$, $\text{cof } \alpha$ is a regular cardinal for any α .

Example 9.5. \aleph_0 is a regular cardinal, since any finite sum of finite numbers is finite. On the other hand, \aleph_ω is a singular cardinal, since

$$\aleph_\omega = \sup_{n < \omega} \aleph_n.$$

In fact, $\text{cof } \aleph_\omega = \omega$. This actually motivated the axiom schema of replacement (the “F” in ZFC), since it seems like singular cardinals should exist, yet this is not provable in Zermelo theory alone.

Lemma 9.6. Let α be an ordinal. Then $\text{cof } \alpha$ is the least cardinal κ such that there is a partition of α into κ many sets of cardinality strictly less than $\text{card } \alpha$.

Proof. Replacing α with $\text{card } \alpha$, we may assume that α is a cardinal. Then notice that the existence of a partition of α , $(S_\beta : \beta < \alpha)$ is the same thing as asserting that

$$\sum_{\beta < \alpha} \text{card } S_\beta = \alpha.$$

Thus we may apply the above definition of regularity. \square

9.2 Inaccessible cardinals

One way we could “access” a cardinal from those below it is by summation; this is impossible for regular cardinals. The other way we might hope to is by applying the power set operation.

Definition 9.7. An *inaccessible cardinal* is a regular cardinal κ such that if $\gamma < \kappa$, then $2^\gamma < \kappa$.

Recall that V_κ is the set of all sets of rank $< \kappa$; i.e., those sets of cardinality $< \kappa$ whose elements are sets of rank $< \kappa$.

Example 9.8. V_ω is the set of all hereditarily finite sets and is roughly equivalent to Peano arithmetic. $V_{\omega+k}$ is the set of all sets that appear in $k+1$ th order arithmetic. $V_{\omega+\omega}$ contains all sets that appear in “ordinary mathematics.” (I think $V_{\omega+\omega}$ is already a model of Borel determinacy since it contains the ω th power set of ω , but I’m not sure.)

We now show that inaccessible cardinals can prove the consistency of ZFC, and hence are much stronger than it. In fact, inaccessibles dwarf even worldly cardinals.

Definition 9.9. A *Grothendieck universe* U is a transitive uncountable set which is closed under pairing and power set, and closed under unions of size γ , where γ is the cardinality of an element of U .

Lemma 9.10. Every Grothendieck universe U gives a model of ZFC of the form (U, \in) , and contains V_ω .

Proof. Obviously U satisfies extensionality, foundation, pairing, power set, and union. Since U is transitive it satisfies empty set and choice. Since U is transitive and satisfies pairing, it contains products and hence contains all functions $x \rightarrow y$ where $x, y \in U$. Therefore U satisfies replacement.

Since U satisfies empty set and pairing, $1 \in U$ and hence $n \in U$ for any n . Therefore $\omega \subseteq U$, yet U is uncountable and transitive, so contains all subsets of a lesser cardinality than itself. In particular, $\omega \in U$, so U satisfies infinity. Since U is transitive, it follows that U contains V_ω . \square

Corollary 9.11. If there is a Grothendieck universe, then ZFC is consistent. In particular, ZFC cannot prove that Grothendieck universes exist.

Proof. Completeness and incompleteness. \square

Example 9.12. If we did not assume Grothendieck universes were uncountable then $V_{-1} = \emptyset$ and V_ω would be Grothendieck universes. In fact, V_ω is a model of ZFC minus the axiom of infinity.

By Maddy's principle of whimsical identity, i.e. the philosophical (not mathematical!) axiom that no ordinal can be uniquely determined by any *whimsical identities*, i.e. properties not forced on it by its definition, this implies that there is a ordinal α such that V_α is a Grothendieck universe. Therefore the principle of whimsical identity implies that there is an inaccessible cardinal, by the following lemma.

Lemma 9.13. If U is a Grothendieck universe, then $\sup_{x \in U} \text{card } x$ is an inaccessible cardinal. Conversely, if κ is inaccessible, then V_κ is a Grothendieck universe.

Proof. If U is a Grothendieck universe, let

$$\kappa = \sup_{x \in U} \text{card } x.$$

Since U is uncountable, there is an $x \in U \setminus V_\omega$, so x must not be hereditarily finite, and since U is transitive, this implies that $\kappa \geq \omega$. Since U is closed under power set, this implies that $\kappa > \omega$, and that if $\gamma < \kappa$, $2^\gamma < 2^\kappa$.

We claim that κ is regular and hence inaccessible. Let Γ be a set of cardinals $< \kappa$ such that $\text{card } \Gamma < \kappa$. By definition of κ , we can replace any $\gamma \in \Gamma$ with a γ' which is the cardinality of a set in U , to get a set Γ' of cardinals $< \kappa$ such that $\text{card } \Gamma' < \kappa$, every $\gamma' \in \Gamma'$ is the cardinality of a set in U , and $\sup \Gamma \leq \sup \Gamma'$. We must show $\sup \Gamma' < \kappa$.

In fact, if $f : \Gamma' \rightarrow U$ sends each cardinal in Γ' to its respective set, then

$$\sup \Gamma' = \text{card} \bigcup_{\gamma' \in \Gamma'} f(\gamma') \in U.$$

Therefore $\sup \Gamma' < \kappa$.

For the converse, let κ be inaccessible; then V_κ is obviously transitive and uncountable, and closed under pairing and power set. It is closed under small union since κ is regular. \square

Corollary 9.14. The following are equivalent:

1. Every set is contained in a Grothendieck universe.
2. There is a proper class of inaccessible cardinals.

Proof. Immediate from the above lemma. \square

Axiom 9.15 (universe). Every set is contained in a Grothendieck universe.

Note that this axiom is implied by the principle of whimsical identity.

9.3 Clubs

Here we record facts about clubs that we will need in the sequel.

Recall that ordinals carry the order topology, so a limit (in the topological sense) is the same as a limit (in the set-theoretic sense), and so a subset of an ordinal is said to be closed if it is closed under taking limits.

Definition 9.16. A *club* in an ordinal α is a set $\subseteq \alpha$ which is closed and unbounded, i.e. closed and cofinal.

Lemma 9.17. Let κ be an uncountable regular cardinal. Let c_0, c_1 be clubs in κ . Then $c_0 \cap c_1$ is a club.

Proof. Let $\alpha < \kappa$; to show that $c_0 \cap c_1$ is cofinal, we will find an element of $c_0 \cap c_1$ which is larger than α . Let β_n be defined inductively to be larger than $\alpha, \beta_0, \dots, \beta_{n-1}$, and the least element of c_j with these properties where $n = j \bmod 2$. Then $\beta = \lim_{n \rightarrow \omega} \beta_n$ is a limit of $c_0 \cap c_1$, which is closed and hence $\exists \beta$, and $\alpha < \beta < \kappa$. \square

Definition 9.18. A *stationary set* in an ordinal α is a set $S \subseteq \alpha$ such that for every club C in α , $S \cap C$ is nonempty.

Lemma 9.19. The intersection of a stationary set and a club in an uncountable regular cardinal is stationary.

Proof. Let S be stationary and C a club; then if D is a club, $S \cap C \cap D = S \cap (C \cap D)$, and $C \cap D$ is a club. \square

So one can think of a club set as a very large subset of α , and its elements are “generic” elements of α . Stationary sets, then, are those that are “open” in α .

Lemma 9.20. Let $S = S_0 \cup S_1$ be a partition of a stationary set S . Then one of the S_0, S_1 is stationary.

Proof. Suppose neither S_0 nor S_1 is stationary, so there are clubs c_0, c_1 they do not meet. If $\beta \in S_0$ then $\beta \notin c_0 \cap c_1$; similarly for S_1 . So S is not stationary. \square

The above notions are mainly of interest for uncountable regular cardinals.

Example 9.21. Suppose that $\text{cof } \kappa = \aleph_0$. Then the above notions are trivial.

To see this, let $f : \aleph_0 \rightarrow \kappa$ be cofinal. (If $\kappa = \aleph_0$ we can take $f = \text{id}$, and if $\kappa = \aleph_\omega$ we can take $f(n) = \aleph_n$.) Then the image C of f is a discrete subset of κ , so C is a club. In particular, by translating C by 1 we obtain two disjoint clubs in κ , and in fact we can partition a subset A of κ such that A^c is bounded into \aleph_0 many disjoint clubs.

Let $S \subseteq \kappa$; then S is stationary iff $\text{card } S^c < \kappa$, which happens iff S^c is bounded. Clearly if S^c is bounded, then S meets every club. For the converse, S must meet every club in the set A , but this is only possible if S^c is bounded in A .

Of course, we will mainly be interested in clubs in large cardinals, which are by definition uncountable and regular.

Definition 9.22. A *Mahlo cardinal* is an inaccessible cardinal κ such that

$$\{\lambda < \kappa : \lambda \text{ is an inaccessible cardinal}\}$$

is a stationary set in κ .

So Mahlo cardinals dwarf inaccessibles in the same sense that inaccessibles dwarf worldlies and worldlies dwarf ZFC: if there is just one Mahlo cardinal κ in V , then V_κ is a model of ZFC plus the universe axiom, so V thinks that the universe axiom is consistent, as is the consistency of the consistency of the consistency ... of the consistency of the universe axiom.

If we believe whimsical identity, then we may conclude Grothendieck's universe axiom; in particular, that **Ord** is Mahlo and hence a “dense set” of ordinals are inaccessible cardinals! In particular, by reflection, one should believe that there is a Mahlo cardinal, and then conclude by whimsical identity again that there are lots of Mahlo cardinals. Actually, assuming that there is just one measurable cardinal, then there will be lots of Mahlo cardinals, as we will prove later.

Definition 9.23. Let ν be a regular uncountable cardinal. The *club filter* of ν , C_ν , is the filter on ν generated by all clubs in ν .

Lemma 9.24. For any regular uncountable cardinal ν , C_ν is a ν -complete filter.

Proof. Let $\{c_\alpha\}_{\alpha < \lambda}$ be a λ -sequence of clubs where $\lambda < \nu$. We must show that $\bigcap_\alpha c_\alpha$ is a club. We prove this by induction on λ . It is clear if λ is a successor, since then we can apply the fact that a finite intersection of clubs is a club, so we assume that λ is a limit ordinal.

We replace c_α with $\bigcap_{\gamma < \alpha} c_\gamma$ and hence assume that we have a decreasing chain of clubs. We are allowed to do this by induction. Then we let β_α be the least element of c_α which is greater than β_γ if $\gamma < \alpha$, and let $\beta = \lim_{\alpha \rightarrow \lambda} \beta_\alpha$, so β is a limit of all the c_α which can be made arbitrarily large by choosing β_0 arbitrarily large. \square

The idea of the above proof is just the usual trick of passing to subsequences, but done λ many times.

Lemma 9.25. The club filter C_κ of an uncountable regular cardinal κ is a normal filter.

Proof. Let $(c_\alpha : \alpha < \kappa)$ be a κ -sequence of clubs in κ . We must show their diagonal intersection $c = \Delta_\alpha c_\alpha$ is also a club. We replace them with the sequence $\bigcap_{\beta < \alpha} c_\beta$, which is a κ -sequence of clubs such that

$$\Delta_\alpha \bigcap_{\beta < \alpha} c_\beta = c,$$

but has the advantage that it is a decreasing sequence. A diagonal intersection is an intersection, so c is closed.

It remains to show that c is cofinal. For any $\zeta < \kappa$ we can find a κ -sequence $(\gamma_\alpha : \alpha < \kappa)$ where $\gamma_\alpha \in c_\alpha$ and $\gamma_\alpha > \zeta$. Since the c_α form a decreasing sequence of closed sets, the limit of the γ_α lies in the diagonal intersection. \square

Lemma 9.26 (Fodor). Let κ be an uncountable regular cardinal, $S \subseteq \kappa$ stationary, and $f : S \rightarrow \kappa$ such that $f(\alpha) < \alpha$ if $\alpha > 0$. Then there is a $\gamma < \kappa$ and stationary $S_0 \subseteq \kappa$ such that $f(\alpha) = \gamma$ for any $\alpha \in S_0$.

Proof. We can remove 0 from S without affecting the statement of the theorem, so assume $0 \notin S$, and assume the lemma fails. Then for every $\alpha < \kappa$ there is a club c_α such that $c_\alpha \cap f^{-1}(\alpha)$ is empty. Let $c = \Delta_\alpha c_\alpha$ be the diagonal intersection of all such clubs. Since the club filter C_κ is a normal filter, c is a club and so there is an $\alpha \in S \cap c$. Therefore $\alpha \in c_\beta$ for every $\beta < \alpha$, and for such β , $\alpha \notin f^{-1}(\beta)$. Therefore $f(\alpha) \geq \alpha$, a contradiction. \square

Theorem 9.27 (Solovay). Let κ be an uncountable regular cardinal, $S \subseteq \kappa$ stationary. Then there is a partition of S into κ many stationary sets.

Proof. Let T be the set of all limit ordinals $\alpha \in S$ such that $\text{cof } \alpha = \omega$ or $S \cap \alpha$ is not stationary in α . Then T is stationary. In fact, if $c \subseteq \kappa$ is a club, then its accumulation set c' is also a club, so $c' \cap S$ is nonempty. Let $\alpha = \min c' \cap S$; then either $\alpha \in T$ or $\text{cof } \alpha$ is uncountable. Suppose that $\text{cof } \alpha$ is uncountable; we will show that $\alpha \in T$ regardless. Since $\alpha \in c'$, α is a limit of c , so $c \cap \alpha$ is cofinal and hence a club in α . Therefore $(c \cap \alpha)' = c' \cap \alpha$ is also a club in α . If $\alpha \notin T$, then $S \cap \alpha$ is stationary in α , so it meets $c' \cap \alpha$, so $c' \cap S$ has an element under α , a contradiction. Therefore $\alpha \in T$, so T is stationary.

Let $\alpha \in T$. We claim there is a cardinal λ and a function $f_\alpha : \lambda \rightarrow \alpha$ which is cofinal, such that $f_\alpha" \lambda \cap T$ is empty. If $\text{cof } \alpha = \omega$, then let $f_\alpha : \omega \rightarrow \alpha$ be cofinal, and if $f_\alpha(n)$ is ever a limit ordinal, replace it with $f_\alpha(n) + 1$. Since T only consists of limit ordinals, $f_\alpha" \omega \cap T$ is empty. If instead $S \cap \alpha$ is not stationary in α , then neither is $T \cap \alpha$, so there is a club $c \subseteq \alpha$ such that $c \cap T$ is empty. Then let $f_\alpha : \text{card } \alpha \rightarrow \alpha$ be cofinal and map onto c . This verifies the claim.

Now let

$$T_\eta^\xi = \{\alpha \in T : \xi \in \text{dom } f_\alpha, f_\alpha(\xi) \geq \eta\}.$$

Lemma 9.28. There is a $\xi < \kappa$ such that for every $\eta < \kappa$, T_η^ξ is stationary in κ .

Proof. Suppose not; then for every ξ there is an $\eta(\xi)$ and a club c_ξ such that $T_{\eta(\xi)}^\xi \cap c_\xi$ is empty. Let $c = \Delta_\xi c_\xi$. Then c is a club.

Let $E = \{\alpha < \kappa : \eta" \alpha \subseteq \alpha\}$. Then E is a club. Indeed, α is a limit of E , then for any β approximating α , $\eta" \beta \subseteq \beta$; taking the union over all such β , we see that $\eta" \alpha \subseteq \alpha$, so E is closed. To see that E is cofinal, suppose not; so there is an α such that for every $\beta > \alpha$,

$\eta''\beta \not\subseteq \beta$. Let $\gamma_0 = \alpha + 1$, and for every n , choose $\gamma_{n+1} > \gamma_n$ so that $\eta''\gamma_n \subseteq \gamma_{n+1}$. Then let $\gamma = \lim_n \gamma_n$. Since $\text{cof } \kappa > \omega$ we have $\alpha < \gamma < \kappa$, but then $\gamma \in E$, a contradiction. Therefore E is a club, and therefore so is $D = c \cap E$.

Since D is a club and T is stationary, there are $\alpha < \gamma$ such that $\alpha, \gamma \in D \cap T$. In particular, if $\xi < \alpha$, $\alpha \in c_\xi$. Thus $\alpha \notin T_{\eta(\xi)}^\xi$. Now let $\xi \in \gamma \cap \text{dom } f_\alpha$; since $\alpha \notin T_{\eta(\xi)}^\xi$, $f_\alpha(\xi) < \eta(\xi)$. Since $\gamma \in E$, and $\xi < \gamma$, $\eta(\xi) < \gamma$, so $\sup f_\alpha''\gamma \leq \gamma < \alpha$. Since f_α is cofinal in α , $\gamma \in \text{dom } f_\alpha$.

Now f_α is continuous and $\gamma \in T$ is a limit ordinal, $f_\alpha(\gamma) = \sup f_\alpha''\gamma \leq \gamma$. But as f_α is increasing, $f_\alpha(\gamma) \geq \gamma$. So $f_\alpha(\gamma) = \gamma$. Therefore $f_\alpha''(\text{dom } \alpha) \cap T$ is nonempty, a contradiction. \square

Let $f : T \rightarrow \kappa$ satisfy $f(\alpha) = f_\alpha(\xi)$ if $\xi \in \text{dom } f_\alpha$, and otherwise $f(\alpha) = 0$. By definition, $f(\alpha) < \alpha$, and for every $\eta < \kappa$,

$$T_\eta^\xi = \{\alpha \in T : f(\alpha) \geq \eta\},$$

but T_η^ξ is stationary, hence nonempty, so f is cofinal. Therefore, by Fodor's lemma, for every $\eta < \kappa$ there is a $\gamma \geq \eta$ such that

$$\{\alpha \in T_\eta^\xi : f(\alpha) = \gamma\}$$

is stationary.

We claim that there is a strictly increasing sequence $(\gamma_\eta : \eta < \kappa)$ such that

$$S_\eta = \{\alpha \in T_{\gamma_\eta+1}^\xi : f(\alpha) = \gamma_\eta\}$$

is stationary. This will complete the proof of Solovay's theorem, as the S_η are disjoint, and if $x \in S \setminus \bigcup_\eta S_\eta$ then we can simply add x to S_0 . To see the claim, let γ_λ be the least γ that has not yet been used for λ a limit ordinal or 0, and for any γ_η , let $\gamma_{\eta+1}$ be the least possible choice of γ such that $\{\alpha \in T_{\gamma_\eta+1}^\xi : f(\alpha) = \gamma\}$ is stationary. \square

9.4 Ramsey theory

Most large cardinal axioms can be viewed as asserting the truth of a certain combinatorial sentence, which also happens to imply a deeper statement about the nature of the universe as a whole. This is well-illustrated by the notion of a weakly compact cardinal, which is one that ostensibly generalizes Ramsey theory, but actually guarantees the existence of inaccessibles.

Recall that if X is a set and δ is a cardinal, then we may view a partition (i.e. coloring) of X into δ many sets (resp. many colors) as a function $X \rightarrow \delta$. Recall also that if λ is a cardinal,

$$[X]^\lambda = \{x \subseteq X : \text{card } x = \lambda\}.$$

This is mainly of interest when $\lambda < \omega$.

Definition 9.29. Let $f : [X]^n \rightarrow \delta$ be a partition. We say that $H \subseteq X$ is a *homogeneous set* for f if $f|H^n$ is constant.

Definition 9.30. Let κ, λ, δ be cardinals, $n < \omega$. We write $\kappa \rightarrow (\lambda)_\delta^n$ if for every partition $[\kappa]^n \rightarrow \delta$ there is a homogeneous set $H \subseteq \kappa$ such that $\text{card } H = \lambda$.

Example 9.31. By the Ramsey friendship theorem, any partition $[6]^2 \rightarrow 2$ has a homogeneous subset of size 3. In other words, $6 \rightarrow (3)_2^2$. More generally, we can interpret any partition law of the form $\kappa \rightarrow (\lambda)_2^2$ can be interpreted as “For any party with κ individuals, either λ many are friends or λ many are enemies.”

Theorem 9.32 (Ramsey theorem). For every $n, m < \omega$, $\omega \rightarrow (\omega)_m^n$.

Proof. We prove this by induction on n . If $n = 1$, then we are considering a partition $\omega \rightarrow m$, and the result is just the infinite pigeonhole principle.

Now consider a partition $[\omega]^{n+1} \rightarrow m$. Define $f_a : [\omega \setminus \{a\}]^n \rightarrow m$, $f_a(x) = f(x \cup \{a\})$. By induction, every f_a has an infinite homogeneous subset. Let $H_0 = \omega$ and let a_0 be arbitrary. Define H_{i+1} to be an infinite f_{a_i} -homogeneous subset of H_i and $a_{i+1} \in H_{i+1}$ such that $a_{i+1} \neq a_j$ for any $j \leq i$. This exists because H_{i+1} is infinite.

Then $(a_i : i \geq k)$ is homogeneous for f_{a_k} . Let $g(i) = f_{a_i}(a_i)$. Then $g \in m^\omega$, so has an infinite homogeneous set by the infinite pigeonhole principle. \square

Corollary 9.33 (finite Ramsey theorem). For every $n, m, T < \omega$, there is a $k < \omega$ such that $k \rightarrow (T)_m^n$.

Proof. If not, then there are n, m, T such that for every k , $k \not\rightarrow (T)_m^n$. Consider the language $(c_0, c_1, \dots, c_{m-1}, d_i : i < \omega, f)$ where f is an n -ary function symbol. Consider the theory T generated by the axioms

1. $\forall x(f(x, \dots, x) = c_0)$.
2. $\forall x_1 \dots \forall x_n((f(x_1, \dots, x_n) = c_0) \vee (f(x_1, \dots, x_n) = c_1) \vee \dots \vee (f(x_1, \dots, x_n) = c_m))$.
3. The schema that for all $i \neq j$, $d_i \neq d_j$.
4. The schema that for every permutation $\sigma : n \rightarrow n$,

$$\forall x_1 \dots \forall x_n(f(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})).$$

5. The schema that for every set $X \subset \omega$ such that $T \leq \text{card } X < \aleph_0$,

$$\bigvee_{\substack{i \neq j \\ i' \neq j' \\ i, j, i', j' \in X}} f(d_i, d_j) \neq f(d_{i'}, d_{j'}).$$

Consider the finite subset $T_k \subset T$ which restricts the schema to range over X such that $T \leq \text{card } X \leq k$. We have a model $M_k \models T_k$, where the universe is $\omega \cup \{c_0, \dots, c_{m-1}\}$, f is a coloring function, and the interpretation of the d_i is that $i \mapsto d_i$. So any finite subset of T is consistent, so by compactness, there is a model $M \models T$. This contradicts the fact that $\omega \rightarrow (\omega)_m^n$. \square

Definition 9.34. A *weakly compact cardinal* is an uncountable cardinal κ such that $\kappa \rightarrow (\kappa)_2^2$.

Without the uncountability hypothesis we would have ω weakly compact. So by whimsical identity, weakly compact cardinals exist. However, we will show they are quite big soon enough.

The reason why we restrict $n < \omega$, rather than allow n to be any infinite cardinal, is the following lemma.

Theorem 9.35 (Erdős-Rado). For any infinite cardinals κ, λ , $\kappa \not\rightarrow (\lambda)_2^\lambda$.

Proof. Let $<$ be a well-ordering of $[\kappa]^\lambda$. When we talk about minimality it will be with reference to this well-ordering. Consider the partition $f : [\kappa]^\lambda \rightarrow 2$ given by $f(A) = 1$ if A is minimal among all infinite subsets of A , or 0 otherwise. Suppose there is a homogeneous $H \subseteq \lambda$ of size λ , so $H \in [\kappa]^\lambda$.

If H isn't minimal among subsets of H , let $A \subset H$ be minimal among subsets of H . Then $f(A) = 1$, and since H is homogeneous it follows that $f(H) = 1$, a contradiction. Therefore H is minimal among subsets of H . Let A_0 be an infinite, cofinite subset of H .

For every i , let A_{i+1} be a cofinite subset of H which contains A_i and infinitely many elements of $H \setminus A_i$. Thus we have a chain $A_i \subset A_{i+1}$, and since $f(A_i) = 1$ by homogeneity of H , A_i is minimal among its subsets, so $A_i > A_{i+1}$. Therefore $([\kappa]^\lambda, <)$ is not well-founded, a contradiction. \square

So it really is important that n be finite! Otherwise we can talk about sets whose cardinality is the same as their own subsets. Note that the above theorem used the axiom of choice in a very strong way.

Definition 9.36. A *partition cardinal* is a cardinal κ such that $\kappa \rightarrow (\kappa)_2^\kappa$.

So partition cardinals are a form of large cardinal that contradicts the axiom of choice.

We now show that power sets are not weakly compact. This is basically just the proof of the Bolzano-Weierstrass theorem again.

Theorem 9.37. For any infinite cardinal κ , $2^\kappa \not\rightarrow (\kappa^+)_2^2$.

Proof. Let $(x_\alpha : \alpha < 2^\kappa)$ enumerate the binary sequences $\kappa \rightarrow 2$, and let $<$ denote the lexicographic ordering on $(x_\alpha : \alpha < 2^\kappa)$. Consider the partition $[2^\kappa]^2 \rightarrow 2$ where $\alpha < \beta$ are friends iff $x_\alpha < x_\beta$. Suppose that H is a set of all friends or all strangers, and $\text{card } H = \kappa^+$. Then H indexes a monotone sequence $(x_\alpha : \alpha \in H)$; assume it is increasing by throwing away a constant subsequence if necessary (there can be no decreasing subsequence since lexicographic orderings are well-founded). Thus all members of H are friends.

Inductively define $(\alpha_\xi, \beta_\xi : \xi < \kappa)$ by saying that β_0 is the first ordinal β such that $x_\beta(\alpha) = 1$ for some α , and α_0 is that ordinal. Then let

$$\beta_\xi = \min\{\beta : (\beta > \sup_{\xi' < \xi} \beta_{\xi'}) \wedge (\exists \alpha \in H((\alpha > 1 + \sup_{\xi' < \xi} \alpha_{\xi'}) \wedge (x_\beta(\alpha) = 1)))\},$$

and α_ξ is the minimal witness that appears in the definition of β_ξ . We take suprema at limit stages, which is possible since κ^+ is a successor cardinal, therefore regular, so any map $\xi \rightarrow \kappa^+$ cannot be cofinal.

In the above construction, the α_ξ are all in H , and they form an increasing sequence in H . But every successor stage fixes one more bit of x_{β_ξ} , and there are only κ many bits. So the construction halts at $\xi = \kappa$, yet $\text{card } H = \kappa^+$ is regular, so the α_ξ cannot form a cofinal sequence in H .

But if $\beta \in H$ and x_β is not minimal (in which case there is nothing to do), there is a $\alpha < \kappa$ such that $x_\beta(\alpha) = 1$. This implies that there is a $\xi < \kappa$ such that $x_\beta < x_{\alpha_\xi}$, and since H indexes an increasing sequence this means $\beta < \alpha_\xi$. So $(\alpha_\xi : \xi < \kappa)$ is cofinal in H , a contradiction. \square

Corollary 9.38. $(2^\kappa) \not\rightarrow (2^\kappa)_2^2$.

9.5 Weakly compact cardinals

Recall that κ is said to be a weakly compact cardinal if it is uncountable and satisfies the partition law $\kappa \rightarrow (\kappa)_2^2$. In other words, at every party with κ guests, either there are κ mutual friends or κ mutual strangers.

We first show that weakly compact cardinals are large cardinals.

Lemma 9.39. Weakly compact cardinals are inaccessible.

Proof. Suppose κ is weakly compact. Since $2^\lambda \not\rightarrow (\lambda^+)_2^2$, if $\lambda < \kappa$, $2^\lambda \not\rightarrow (\kappa)_2^2$. This implies $2^\lambda < \kappa$.

So it suffices to show that κ is regular. If not, let $(A_\xi : \xi < \lambda)$ be a partition of κ , where $\text{card } A_\xi = \delta_\xi$, $\lambda, \delta_\xi < \kappa$. For $\alpha, \beta < \kappa$, say that α, β are friends if there is a $\xi < \lambda$ such that $\alpha, \beta \in A_\xi$. Then there is an H such that $\text{card } H = \kappa$ and all elements of H are friends, or all elements of H are strangers. If they are strangers, then choose a function $f : \kappa \rightarrow \lambda$ which sends α to the ξ such that $\alpha \in A_\xi$; then f is injective, so $\lambda < \kappa \leq \lambda$, a contradiction. So they are all friends, and there is a ξ such that $H \subseteq A_\xi$. But then $\delta_\xi < \kappa \leq \delta_\xi$, a contradiction. \square

Theorem 9.40. The following are equivalent for an infinite cardinal κ :

1. κ is weakly compact or $\kappa = \omega$.
2. $\kappa \rightarrow (\kappa)_2^2$.
3. For every linear order of length κ , either there is an ascending sequence of length κ or a descending sequence of length κ .

Proof. The equivalence between the first two criteria is clear.

We now show $2 \Leftrightarrow 3$. Suppose $\kappa \rightarrow (\kappa)_2^2$ and $(x_\alpha : \alpha < \kappa)$ enumerates a linear order. Say that α and β are friends if $x_\alpha < x_\beta$. Then there is a set $H \subseteq \kappa$ of cardinality κ consisting of friends or strangers. If H consists of friends we're done. Otherwise, since $\alpha \mapsto x_\alpha$ is surjective, it cannot be constant, so throw away a constant subsequence of $(x_\alpha : \alpha \in H)$ to get a decreasing subsequence.

Conversely, suppose that for every linear order of cardinality κ , either there is an ascending subsequence of length κ or a descending subsequence of length κ . Given a partition of $[\kappa]^2$ into friends and strangers, let $\alpha \preceq \beta$ iff $\alpha \leq \beta$ are friends. Otherwise let $\beta \preceq \alpha$. So we obtain a linear order of length κ , so it has a monotone sequence of length κ . If it is increasing we have friends; otherwise, we have strangers. Therefore $2 \Leftrightarrow 3$. \square

We obtain a very concrete corollary from the above analysis.

Corollary 9.41 (Bolzano-Weierstrass theorem). $[0, 1]$ is compact.

Proof. It suffices to show that every sequence of real numbers has a monotone subsequence; then if that sequence is bounded, its monotone subsequence will be Cauchy, and we will be done. But this follows by applying the above theorem to the image of ω under the sequence. \square

Another characterization of weakly compact cardinals is in terms of trees.

Definition 9.42. Let κ be a cardinal. An *Aronszajn tree* of cardinality κ is a tree T with κ nodes, such that all branches of T have length less than κ and all levels of T have width less than κ .

So Koenig's lemma says that there are no countable Aronszajn trees. We now use this statement to generalize Ramsey's theorem.

Theorem 9.43. Let κ be an inaccessible cardinal or $\kappa = \omega$. If there are no Aronszajn trees of cardinality κ , then κ satisfies, for every $m < \omega$ and $\lambda < \kappa$, the partition law $\kappa \rightarrow (\kappa)_{\lambda}^m$.

Note that this is just the infinite Ramsey theorem in disguise. We will not need the fact that κ is uncountable, so this gives another proof of the infinite Ramsey theorem for $\kappa = \omega$.

To prove the above theorem, we first construct a tree (κ, \prec) for any partition $f : [\kappa]^{n+1} \rightarrow \eta$ where η is infinite. Put $\beta \prec \alpha$ if $\beta < \alpha$, and for every sequence $\gamma_0 \prec \dots \prec \gamma_{n-1} \prec \beta$, we have $f(\gamma_0, \dots, \gamma_{n-1}, \beta) = f(\gamma_0, \dots, \gamma_{n-1}, \alpha)$.

Lemma 9.44. For any partition, (κ, \prec) is a tree.

Proof. The relation \preceq is antisymmetric and well-founded since \leq is.

We now check that it is transitive by induction on α . Suppose that \prec is transitive in the \prec -predecessors of α , and $\delta \prec \gamma \prec \alpha$. Suppose $\gamma_0 \prec \dots \prec \gamma_{n-1} \prec \delta$. Then by induction, $\gamma_{n-1} \prec \gamma$. So

$$f(\gamma_0, \dots, \gamma_{n-1}, \delta) = f(\gamma_0, \dots, \gamma_{n-1}, \gamma) = f(\gamma_0, \dots, \gamma_{n-1}, \alpha).$$

Therefore $\delta \prec \alpha$.

Finally we must check that if $\delta \prec \beta$ and $\gamma \prec \beta$ then δ, γ are \prec -comparable. Suppose $\delta < \gamma$ and if $\zeta \prec \beta$ and $\zeta < \delta$, then ζ is comparable any predecessor of β . Suppose $\gamma_0 \prec \dots \prec \gamma_{n-1} \prec \delta$. Then by induction, $\gamma_{n-1} \prec \gamma$. So

$$f(\gamma_0, \dots, \gamma_{n-1}, \delta) = f(\gamma_0, \dots, \gamma_{n-1}, \beta) = f(\gamma_0, \dots, \gamma_{n-1}, \gamma).$$

Therefore δ and γ are comparable. \square

Lemma 9.45. Fix a partition of $[\kappa]^{n+1}$. Let λ be a limit ordinal and suppose that $\alpha, \beta \in \kappa$ are both of \prec -level λ . Then the sets of predecessors of α, β are not equal.

Proof. Suppose they are. Then if $\alpha_0 \prec \dots \prec \alpha_{n-1} \prec \alpha$, we also have $\alpha_{n-1} \prec \beta$. Since λ is a limit ordinal, pick an $\alpha_n \prec \alpha$ such that $\alpha_{n-1} \prec \alpha_n \prec \beta$. Then

$$f(\alpha_0, \dots, \alpha_{n-1}, \alpha_n) = f(\alpha_0, \dots, \alpha_{n-1}, \alpha_n) = f(\alpha_0, \dots, \alpha_{n-1}, \beta).$$

Therefore α, β are comparable and yet at the same level and not equal. \square

Lemma 9.46. Fix a partition $[\kappa]^{n+1} \rightarrow \eta$. For every $\delta < \kappa$, the δ -level of (κ, \prec) has cardinality $\leq \eta^{\text{card } \delta}$.

Proof. By induction on δ . If $\delta < \omega$ and level δ has cardinality $\leq \eta$, then level $\delta + 1$ has cardinality $\leq \eta^2 = \eta$. Take suprema at limit stages. If δ is infinite and its level has cardinality $\leq \eta^{\text{card } \delta}$, then $\text{card}[\delta]^{n-1} = \text{card } \delta$, then level $\delta + 1$ has cardinality $\leq (\eta^{\text{card } \delta})^2 = \eta^{\text{card } \delta}$ as well. \square

Proof of Theorem 9.43. By induction on m . The case $m = 1$ says that one cannot partition κ into λ subsets of size $< \kappa$, which is true since κ is regular. Now fix a partition $f : [\kappa]^{m-1} \rightarrow \lambda$ and consider the induced tree (κ, \prec) . For every level $\delta < \kappa$, its cardinality is at most $\lambda^{\text{card } \delta} < \kappa$, since κ is inaccessible or $\kappa = \omega$. Since (κ, \prec) is not an Aronszajn tree, it has a branch X of length κ . Since $\kappa \rightarrow (\kappa)_\lambda^{m-1}$ and $\text{card } X = \kappa$, we consider $f|_{[X]^{m-1}}$ and find a homogeneous subset of X of cardinality κ . \square

Lemma 9.47. Let (T, \prec) be a tree. Then there is an extension of \prec to an ordering $<$ such that $(T, <)$ is a chain, and for every $x \in T$, the successors of x form a segment of $(T, <)$.

Proof. Suppose that $(T_0, <)$ is such that $T_0 \subset T$ and $(T_0, <)$ extends (T_0, \prec) in the way prescribed by the lemma, and T_0 is closed under \prec -predecessors as a subset of T . Choose a \prec -minimal element $x \in T \setminus T_0$ and consider a \subset -maximal linear subset X of T whose minimal element is x , which exists by Zorn's lemma. Define $<$ on X by $z < y$ iff $z \prec y$. Then $(T_0 \cup X, <)$ extends $(T_0 \cup X, \prec)$, in the way prescribed by the lemma, and $T_0 \cup X$ is closed under \prec -predecessors as a subset of T . Now run Zorn's lemma on the set of all $(T_0, <)$ with this property. \square

Theorem 9.48. The following are equivalent for an infinite cardinal κ :

1. κ is a weakly compact cardinal, or $\kappa = \omega$.
2. κ is a regular cardinal, there are no Aronszajn trees of cardinality κ , and for every $\lambda < \kappa$, $2^\lambda < \kappa$.
3. For every $\lambda < \kappa$ and $m < \omega$, $\kappa \rightarrow (\kappa)_\lambda^m$.

Proof. Clearly $3 \implies 1$ and we just showed $2 \implies 3$. Now to show $1 \implies 2$.

Let (T, \prec) be a tree with κ nodes such that every level has cardinality $< \kappa$. We must prove that a Koenig lemma holds for T : T has a branch of length κ . We can extend \prec to an ordering $<$ such that $(T, <)$ is a chain, and for every $x \in T$, the successors of x form a segment of $(T, <)$. Since κ is weakly compact or $\kappa = \omega$, there is a κ -sequence C in T . We may assume it is increasing and well-ordered.

Let $T^x = \{z \in T : x \prec z\}$ and let $C^y = \{z \in C : y < z\}$. Let $D = \{x \in T : \exists y \in C (C^y \subseteq T^x)\}$. Suppose that $x, x' \in D$. Then there are $y, y' \in C$ such that $C^y \subseteq T^x$ and $C^{y'} \subseteq T^{x'}$. Assume $y \preceq y'$; then $C^y \subseteq T^x \cap T^{x'}$. Since C^y is nonempty, because C^y is a tail in the κ -sequence C , that means that x, x' have a common successor in T , and hence x, x' are comparable. Therefore D is a chain.

We now claim that $\text{card } D \geq \kappa$; then if we extend D to a branch, which is always possible by Zorn's lemma, D will be a witness that T is not Aronszajn. In fact, it suffices to show that for every level α , D has an element of level α . Since each level has cardinality $< \kappa$, and C , hence C^y by regularity of κ , has cardinality κ , C^y has elements of arbitrarily high level.

To see that D has an element of level α , let E be the set of all segments $T^x \cap C$ such that x is of level α . Every element of C of sufficiently high level is in one such segment, so $\bigcup E$ is cofinal in C and hence of cardinality κ . But $\text{card } E < \kappa$, since there are $< \kappa$ nodes at level α , so since κ is regular there must be an x at level α such that $\text{card}(T^x \cap C) = \kappa$. But then $T^x \cap C$ is cofinal in C , so $x \in D$, as required. \square

9.6 The weak compactness theorem

Another useful characterization of weakly compact cardinals, which justifies their name, is in terms of the compactness theorem.

Definition 9.49. Let κ, λ be cardinals. The language $\mathcal{L}_{\kappa, \lambda}$ consists of

1. κ variables,
2. κ relation, function, and constant symbols,
3. Finitary logical connectives and quantifiers,
4. Infinitary connectives $\bigvee_{\xi < \alpha} \varphi_\xi$ and $\bigwedge_{\xi < \alpha} \varphi_\xi$ for every $\alpha < \kappa$, and 0
5. Infinitary quantifiers $\exists_{\xi < \alpha} v_\xi$ and $\forall_{\xi < \alpha} v_\xi$ for every $\alpha < \lambda$.

Here we interpret $\bigvee_{\xi < \alpha} \varphi_\xi = \varphi_0 \vee \varphi_1 \vee \dots \vee \varphi_\omega \vee \varphi_{\omega+1} \vee \dots$ and $\exists_{\xi < \alpha} v_\xi = \exists v_0 \exists v_1 \dots \exists v_\omega \vee v_{\omega+1} \vee \dots$.

Definition 9.50. We say that $\mathcal{L}_{\kappa, \lambda}$ satisfies the weak compactness theorem if, whenever Σ is a set of sentences of $\mathcal{L}_{\kappa, \lambda}$ such that $\text{card } \Sigma \leq \kappa$ and every $\Sigma_0 \subset \Sigma$ such that $\text{card } \Sigma_0 < \kappa$ has a model, then Σ has a model.

Lemma 9.51. If κ is a weakly compact cardinal or $\kappa = \omega$ then $\mathcal{L}_{\kappa, \kappa}$ satisfies the weak compactness theorem.

This is basically trivial by what we have done so far. In fact, the usual proof of the compactness theorem goes through; we use the fact that κ has no Aronszajn trees to compensate for the lack of Koenig's lemma.

Lemma 9.52. If κ is an inaccessible cardinal and $\mathcal{L}_{\kappa, \omega}$ satisfies the weak compactness theorem, then κ is weakly compact.

Proof. Let $(T, <)$ be a tree of cardinality κ and suppose every level of T is of width $< \kappa$. We claim that T satisfies a Koenig lemma. Consider the $\mathcal{L}_{\kappa, \omega}$ language with a unary predicate B and constant symbols c_x for every $x \in T$. Let Σ be the set of sentences consisting of $\neg B(c_x) \vee \neg B(c_y)$ if x, y are incomparable, and $\bigvee_{x \in U_\alpha} B(c_x)$ where U_α is the α th level of T . So Σ says that $\{x \in T : B(c_x)\}$ is a branch of length κ through T . If Σ has a model $(M, <, X)$ and we interpret the constant symbols as intended, then Σ also has a model of the form $(T, <, Y)$ where Y is a branch of length κ through T . So it suffices to show that Σ has a model.

Let $\Sigma_0 \subset \Sigma$, $\text{card } \Sigma_0 < \kappa$. Since T is inaccessible we can take a sufficiently long initial segment of T and a branch through that initial segment to obtain a model of Σ_0 . Therefore by the weak compactness theorem, Σ has a model. \square

Thus we give our final characterization of weak compactness.

Definition 9.53. We say that κ is a *whimsically inaccessible cardinal* if κ is inaccessible or $\kappa = \omega$.

Theorem 9.54. The following are equivalent for an infinite cardinal κ :

1. κ is weakly compact or $\kappa = \omega$.
2. For every $\lambda < \kappa$ and $n < \omega$, $\kappa \rightarrow (\kappa)_\lambda^n$.
3. Every linear order of cardinality κ has a monotone κ -sequence.
4. κ is whimsically inaccessible, and there are no Aronszajn trees of size κ .
5. κ is whimsically inaccessible, and $\mathcal{L}_{\kappa, \kappa}$ satisfies the weak compactness theorem.
6. κ is whimsically inaccessible, and $\mathcal{L}_{\kappa, \omega}$ satisfies the weak compactness theorem.

Chapter 10

Measurable cardinals

Recall that given a set x , an ultrafilter U on x is the set of subsets of x of μ -measure 1, where μ is some 2-valued, finitely additive measure whose algebra of definition is 2^x . Filters are similar but we do not assume that μ is defined for every subset of x ; thus, if F is a filter, then $x \in F$, F is closed under intersection, F is upwards closed, and $\emptyset \notin F$.

Definition 10.1. Let κ be a cardinal and F a filter. We say that F is a κ -complete filter if F is closed under λ -intersection for any cardinal $\lambda < \kappa$. If F is \aleph_1 -complete, then we say that F is a countably complete filter.

Thus every filter is \aleph_0 -complete, and a filter is countably complete if it is closed under countable intersections.

Example 10.2. Let κ be an infinite cardinal. Then cofinite filter on κ is not countably complete. In fact, $\{\{n\}^c : n < \omega\}$ is a countable set whose intersection is empty, yet whose elements are cofinite subsets of $\kappa \supseteq \omega$.

In “ordinary mathematics,” the only ultrafilters we have to work with are the trivial ultrafilters, and the extensions of cofinite filters that we get by using Zorn’s lemma (or the boolean prime ideal theorem). Therefore if κ is an uncountable cardinal, we are going to have to look very hard to find a nontrivial κ -complete ultrafilter.

Definition 10.3. A measurable cardinal is an uncountable cardinal κ such that there is a nontrivial κ -complete ultrafilter on 2^κ .

10.1 Properties of measurable cardinals

Lemma 10.4. If κ is a measurable cardinal then κ is regular.

Proof. Let Λ be a set of cardinals $< \kappa$ such that $\text{card } \Lambda < \kappa$, and let U be a witness to the measurability of κ . Since U is nontrivial, finite sets are U -small, and since U is κ -complete this implies that for every cardinal $\lambda < \kappa$, λ is U -small. By κ -completeness again, $\sum_{\lambda \in \Lambda} \lambda$ is U -small, yet κ is U -large. \square

This gives an amusing measure-theoretic consequence.

Theorem 10.5. For every measurable cardinal κ there is a surjective probability measure on 2^κ .

Proof. Let E be the set of all limit ordinals $< \kappa$, and 0. Then $F = \{\{\lambda + n : n < \omega\} : \lambda \in E\}$ is a partition of κ into countable sets, and since κ is a regular cardinal, it follows that $\text{card } F = \kappa$. So let α_γ be the γ th ordinal in E ; then γ ranges over all ordinals $< \kappa$. Let $A_n = \{\alpha_\lambda + n : \lambda < \kappa\}$. Then $\text{card } A_n = \kappa$, so $\{A_n : n < \omega\}$ is a partition of κ into countably many sets of cardinality κ .

Let U be a witness to the measurability of κ and let μ_n be the two-valued measure induced by U on A_n . Now let

$$\eta(x) = \sum_{n=1}^{\infty} 2^{-n} \mu_n(A_n \cap x).$$

Clearly η is a probability measure, and since any real in $[0, 1]$ can be written as a series of dyadics (by writing it in base 2) it is surjective. \square

Theorem 10.6. Every measurable cardinal is inaccessible, and hence worldly.

Proof. Let κ be a measurable cardinal; then κ is regular. Let $\gamma < \kappa$ and assume $\kappa \leq 2^\gamma$. Let U witness that κ is measurable; since we have an injective function $\kappa \rightarrow 2^\gamma$ we have an ultrafilter V on 2^γ . We will find a function $f : \gamma \rightarrow 2$ such that $\{f\} \in V$, a contradiction.

If $\beta \leq \gamma$ let $A(f, \beta) = \{g \in 2^\gamma : \forall \alpha < \beta (f(\alpha) = g(\alpha))\}$. This partitions into

$$A(f, \beta) = \{g \in A(f, \beta) : g(\beta) = 0\} \cup \{g \in A(f, \beta) : g(\beta) = 1\} = A(f, \beta, 0) \cup A(f, \beta, 1).$$

Moreover $A(f, 0) = 2^\gamma$, so $A(f, 0) \in U$. Thus $A(f, \beta, i) \in U$ for some i . Let $f(0) = i$.

Assume $A(f, \beta) \in U$ for all $\beta \leq \alpha$ and α is a successor. Then let $f(\alpha) = i$ and continue the induction. Something similar happens at limits, but we have to use

$$A(f, \alpha) = \bigcap_{\beta < \alpha} A(f, \beta).$$

When we finish the induction, $U \ni A(f, \gamma) = \{f\}$, which is impossible. \square

10.2 Combinatorics of ultrafilters

In this section we establish a sufficient condition for measurable cardinals to exist, which explains why we were not able to find any ultrafilters other than the obvious ones. We then consider special types of ultrafilters on measurable cardinals.

Definition 10.7. Let U be a nontrivial ultrafilter. The *completeness* of U , $\text{comp } U$, is the smallest cardinal κ such that U is not κ -complete.

So U is not closed under intersections of size $\text{comp } U$.

Lemma 10.8. Let U be a nontrivial ultrafilter on κ . Then $\text{comp } U \leq \kappa$.

Proof. Since U is nontrivial, it contains all cofinite subsets of κ , in particular those of the form $\{\alpha\}^c$, $\alpha < \kappa$. But

$$\bigcap_{\alpha < \kappa} \{\alpha\}^c = \emptyset,$$

so U is not closed under intersections of size κ . \square

Lemma 10.9. Let U be a nontrivial ultrafilter on A , $f : A \rightarrow B$, and

$$V = \{X \subseteq B : f^{-1}(X) \in U\}$$

be the pushforward ultrafilter on B . Then V is trivial iff $f^{-1}(\{x\}) \in U$ for some $x \in B$; if V is nontrivial then $\text{comp } U \leq \text{comp } V$.

Proof. The first claim is obvious. For the second, if V is not closed under intersections of size κ , say $\{X_\alpha\}_{\alpha < \kappa}$ has $X_\alpha \in V$ but $\bigcap_{\alpha < \kappa} X_\alpha \notin V$, then $\bigcap_{\alpha < \kappa} f^{-1}(X_\alpha) \notin U$, so neither is U closed under intersections of size κ . \square

Theorem 10.10. There is a countably complete nontrivial ultrafilter iff there is a measurable cardinal.

Proof. If there is a measurable cardinal then by definition there is a countably complete ultrafilter. So we just have to prove the converse.

Let U be a nontrivial, countably complete ultrafilter on A and let $\kappa = \text{comp } U$. Since U is countably complete, κ is uncountable. We claim that κ is measurable.

We know that $\kappa \leq \text{card } A$. Let $\{X_\alpha\}_{\alpha < \kappa}$ be such that $X_\alpha \in U$ and $\bigcap_\alpha X_\alpha \notin U$. Then $B = A \setminus \bigcap_\alpha X_\alpha \in U$. Let

$$\begin{aligned} f : A &\rightarrow \kappa \\ a &\mapsto \begin{cases} \text{least } \gamma \text{ such that } a \notin X_\gamma, \text{ if } a \in B \\ 0, \text{ else.} \end{cases} \end{aligned}$$

Then $f(a) = 0$ provided that $a \notin B$, and the pushforward of U by f ,

$$V = \{X \subseteq \kappa : f^{-1}(X) \in U\},$$

is an ultrafilter such that $\kappa \leq \text{comp } V \leq \kappa$, hence $\text{comp } V = \kappa$, provided that V is nontrivial.

So it suffices to show that $f^{-1}(\beta) \notin U$ for every $\beta < \kappa$. In fact, $f^{-1}(0) = A \setminus B \notin U$. Otherwise, $f^{-1}(\beta)$ is disjoint from X_β . But $X_\beta \in U$ and no filter can contain two disjoint sets (or else it would contain \emptyset). \square

Definition 10.11. A *normal ultrafilter* is an ultrafilter U on an infinite cardinal κ such that for every function $f : \kappa \rightarrow \kappa$, if

$$\{\alpha < \kappa : f(\alpha) < \alpha\} \in U,$$

then there is a $\beta < \kappa$ such that $f^{-1}(\beta) \in U$.

Thus, an ultrafilter is normal provided that for every function which sends most elements downwards, it is actually constant on a large set. The β should be “small” relative to κ in the sense that the interval $(\beta, \kappa) \in U$.

Definition 10.12. Let κ be a regular cardinal and let $\{X_\alpha\}_{\alpha < \kappa}$ be a κ -sequence of subsets of κ . The *diagonal intersection* is

$$\Delta_{\alpha < \kappa} X_\alpha = \bigcup_{\alpha} [0, \alpha] \cap X_\alpha.$$

Thus an ordinal β is in the diagonal intersection if it is in X_α , for every $\alpha < \beta$ (rather than $\alpha < \kappa$!) Thus diagonal intersection is less strict than intersection; this also motivates the terminology “diagonal intersection.”

Example 10.13. Let $X_\alpha = (\alpha, \kappa)$. Then $\bigcap_\alpha = \emptyset$ but $\Delta_{\alpha < \kappa} X_\alpha = \kappa$.

Lemma 10.14. Let κ be a regular cardinal and U an ultrafilter on κ . Then U is normal iff U is closed under diagonal intersection.

Proof. Suppose that U is abnormal. So there is an $X \in U$ and a $f : \kappa \rightarrow \kappa$ such that $f(\alpha) < \alpha$ for every $\alpha \in X$ and $X_\alpha = f^{-1}(\alpha) \notin U$ for all $\alpha < \kappa$. Let $Y = \Delta_{\alpha < \kappa} X_\alpha^c$. If $\gamma \in X \cap Y$ then $\gamma \notin X_{f(\gamma)}$ since $f(\gamma) < \gamma$, yet $\gamma \in X_{f(\gamma)}$. So $X \cap Y$ is nonempty, and since $X \in U$ it follows that $Y^c \in U$, so U is not closed under diagonal intersection.

Conversely, suppose that U is not closed under diagonal intersection, so that there are $X_\alpha \in U$ such that $X = \Delta_{\alpha < \kappa} X_\alpha \notin U$. Then $X^c \in U$, and for every $\alpha \notin X$ there is a $f(\alpha) < \alpha$ such that $\alpha \notin X_{f(\alpha)}$. Then $f^{-1}(\beta) \subseteq X^c \setminus X_\beta \notin U$ for any β so U is abnormal. \square

Definition 10.15. An ultrafilter U on an infinite cardinal κ is a *uniform ultrafilter* if for every $X \in U$, $\text{card } X = \kappa$. It is a *weakly uniform ultrafilter* if for every $\gamma < \kappa$, $(\gamma, \kappa) \in U$.

Example 10.16. Every trivial ultrafilter is normal, and neither uniform nor weakly uniform. Suppose that $U = \{X \subseteq \kappa : \alpha \in X\}$, and assume $f(\alpha) = \beta$. Then $f^{-1}(\beta) \in U$, so U is normal.

Definition 10.17. The *generalized cofinite filter* on an infinite set X is the filter of all $A \in 2^X$ such that $\text{card } A^c < \text{card } X$.

Lemma 10.18. An ultrafilter on an infinite set X is uniform iff it contains the generalized cofinite filter of X .

Proof. Let F be the generalized cofinite filter and U an ultrafilter. If U does not contain F , let $A \in F \setminus U$; then $A^c \in U$ but $\text{card } A^c < \text{card } X$, so U is not uniform. Conversely, if U is not uniform, let $A \in U$ be such that $\text{card } A < \text{card } X$. Then $A^c \in F \setminus U$. \square

Lemma 10.19. An ultrafilter on a regular cardinal is uniform iff it is weakly uniform.

Proof. Let U be an ultrafilter on κ . Suppose that U is uniform; if there is a tail (γ, κ) which is not in U , then its complement $\gamma \in U$, but $\text{card } \gamma < \kappa$, a contradiction; so U is weakly uniform. Conversely, if U is weakly uniform and $X \in U$, suppose that $\text{card } X < \kappa$. Then X is bounded, so let $\gamma = \sup X$; it follows that $[0, \gamma] \in U$, contradicting that U is weakly uniform. \square

Lemma 10.20. Let U be a nontrivial ultrafilter on a regular cardinal κ . Then U is κ -complete iff U is uniform.

Proof. Suppose that U is κ -complete; to show uniformity, it suffices to show that U contains the generalized cofinite filter F . In fact, if $A \in F$, then $\text{card } A^c < \kappa$, so A^c can be written as a union of fewer than κ singletons. Since U is nontrivial this implies that $A^c \notin U$.

Suppose that U is uniform and weakly uniform and let $(X_\alpha)_{\alpha < \kappa} \in U$. If $\bigcap_\alpha U_\alpha \notin U$, then its complement $\bigcup_\alpha U_\alpha^c$ is in U , so let $\gamma = \sup \bigcup_\alpha U_\alpha^c$. If $\gamma < \kappa$ then $[0, \gamma] \in U$, a contradiction. Therefore $\bigcup_\alpha U_\alpha^c$ is cofinal, implying that $\bigcap_\alpha U_\alpha$ is bounded and hence $\text{card } \bigcup_\alpha U_\alpha < \kappa$, contradicting that U is uniform. \square

Theorem 10.21 (Scott). If κ is a measurable cardinal, then there is a normal uniform ultrafilter on κ .

Proof. Let U be a κ -complete nontrivial ultrafilter on κ .

We will construct a function f such that:

1. For every β , $f^{-1}(\beta) \notin U$.
2. If $\{g < f\} \in U$ then there is a β such that $g^{-1}(\beta) \in U$.

If this is impossible, let $f_0 = \text{id}$. Since U is nontrivial, f_0 satisfies condition 1 and hence fails condition 2. Let f_{n+1} witness the failure of condition 2 for f_n , so that f_{n+1} satisfies condition 1 and we can continue the induction. Let $X_n = \{f_{n+1} < f_n\}$; by countable completeness of U , $X = \bigcap_n X_n \in U$. But if $\alpha \in X$, then $f_0(\alpha) > f_1(\alpha) > \dots$, a contradiction.

Now let V be the pushforward of U by f . By condition 1, V is a nontrivial κ -complete ultrafilter. So by the lemma V is uniform.

To see that V is normal, let $g : \kappa \rightarrow \kappa$ be a function such that

$$S = \{\alpha < \kappa : g(\alpha) < \alpha\} \in V.$$

Now $f^{-1}(S) = \{\alpha < \kappa : g(f(\alpha)) < f(\alpha)\} \in U$. By condition 2, this implies that there is a β such that $g \circ f^{-1}(\beta) \in U$, so $g(\beta) \in V$, as desired. \square

Example 10.22. Any nontrivial ultrafilter on \aleph_0 is uniform since it is \aleph_0 -complete by definition. Moreover, every nontrivial ultrafilter U on \aleph_0 is abnormal. To see this, let $A \in U$ be coinfinite. For any n , let $f(n) = n$ if $n \notin A$, or $f(n)$ be the greatest element of A^c under n if $n \in A$. Since A is infinite and coinfinite, f is constant exactly on finite sets, but pushes down on A , a contradiction.

Since \aleph_0 is the only example of a nonmeasurable cardinal κ with a nontrivial κ -complete ultrafilter (hence, uniform ultrafilter), all the above theory is a bit trivial if κ is not a measurable cardinal.

We are interested in normal, uniform ultrafilters on measurable cardinals anyways, since we need them in order to take iterated ultrapowers of V in the sequel.

10.3 Ultrapowers by measurable cardinals

We now show that the existence of a measurable cardinal is actually much stronger than an inaccessible cardinal: it actually gives elementary embeddings of V into class models.

Given an ultrafilter U on some set A and functions $f, g : A \rightarrow B$, let $f \sim_U g$ mean that $\{f = g\} \in U$. Given a binary relation E on U , let $f E_U g$ mean that $\{f E g\} \in U$. Then the ultrapower is defined by

$$\Pi_U(B, E) = (B^A/U, E_U).$$

By Los' theorem, the inclusion map $B \rightarrow B^A/U$ is an elementary embedding for the relation E . We want to replace $(B^A/U, E)$ with an honest-to-god model of set theory, i.e. show that it is isomorphic to (M, \in) for some transitive set M .

Lemma 10.23 (Mostowski collapse). Let E be a well-founded binary relation on A such that (A, E) satisfies the axiom of extensionality. Then there is a transitive set M such that (M, \in) is isomorphic to (A, E) .

Proof. Since E is a well-founded binary relation, there is a minimal element 0 of A . By extensionality, 0 is uniquely defined. Now define $\pi : E \rightarrow V$ by sending $\pi(0) = \emptyset$, and

$$\pi(x) = \{\pi(y) : y E a\}.$$

Then $\pi(y) \in \pi(x)$ iff $y E x$ and the map π is injective. Let $M = \pi(E)$; by construction M is transitive and $(M, \in) \cong (A, E)$ by π . \square

Definition 10.24. Let (A, E) meet the hypotheses of Mostowski's collapse lemma and let M be the transitive set given by the conclusion of that lemma. We call M the *transitive collapse* of (A, E) .

So, provided that our model of set theory was a well-founded set to begin with, there is no loss in assuming that it is a transitive model. Later we will use a theorem known as Scott's trick to drop the assumption that A was a set, so we can even take the transitive collapse of a proper class.

Example 10.25. There exist models of ZFC which are not well-founded, and in particular do not have a transitive collapse. To see this, consider the language (\in, c_1, c_2, \dots) where c_i are new constant symbols. Let T extend ZFC by declaring that $c_{i+1} \in c_i$ for every i . Then by compactness, T has a model if ZFC does, and clearly T is not well-founded. The trouble here is that even though M satisfies the axiom of foundation, M thinks that the sequence $\vec{c} = \{c_1, c_2, \dots\}$ does not exist, even though it does in V . However, if we were to take the transitive collapse of M , then \vec{c} would appear in M , a contradiction.

Of course, we are mainly interested in models that are well-founded anyways, but the inability to take transitive collapse will be a major hurdle to overcome later when we take iterated ultrapowers.

Lemma 10.26. Let E be a well-founded binary relation on M which satisfies the axiom of extensionality, and let U be a countably complete ultrafilter. Then $\Pi_U(M, E)$ is well-founded.

Proof. Suppose not, so there are functions $\{f_n : n \in \omega\}$ such that $[f_{n+1}]E_U[f_n]$, where $[\cdot]$ sends a function to its equivalence class under U . Therefore

$$X_n = \{x \in A : f_{n+1}(x)Ef_n(x)\} \in U$$

but U is countably complete, so $\bigcap_n X_n \in U$. Therefore there is a $x \in \bigcap_n X_n$. So $f_{n+1}(x)Ef_n(x)$, yet E is well-founded, a contradiction. \square

Therefore we may take the Mostowski collapse of $\Pi_U(M, E)$.

Definition 10.27. Let (M, E) be a well-founded binary relation which satisfies the axiom of extensionality, and let U be a countably complete ultrafilter. The *transitive ultrapower* of (M, E) by U , $\text{Ult}(M, E, U)$, is the Mostowski collapse of the ultrapower of $\Pi_U(M, E)$.

The ultrapower map $M \rightarrow \text{Ult}(M, E, U)$ is then an elementary embedding.

We now carry out the above construction to remove the assumption that M is a set. We define the equivalence class of a function $f : A \rightarrow M$ under some ultrafilter U on A by

$$[f]_U = \{g \in M^A : g \sim_U f \text{ and } g \text{ has minimal rank}\}.$$

The restriction to sets of minimal rank is known as *Scott's trick*. Since the rank of elements of $[f]_U$ is bounded, $[f]_U$ is a set, and we may let $\Pi_U(M, E)$ be the class of all such sets. Los' theorem goes through, but now it is a schema of theorems rather than a first-order theorem.

However, the Mostowski collapse lemma does not go through unchanged.

Example 10.28. Take (Ord, E) where $\alpha E \beta$ if α, β have the same parity and $\alpha < \beta$, or if α is even and β is odd. This has ordertype $\text{Ord} + \text{Ord}$. So if M was the transitive collapse of (Ord, E) , we would have a nontrivial order-embedding $\text{Ord} \rightarrow \text{Ord}$. But, by transfinite induction, the only order-embedding $\text{Ord} \rightarrow \text{Ord}$ is constant, so this is a contradiction. Therefore (Ord, E) is a counterexample to the Mostowski collapse lemma.

Definition 10.29. By a *class model* we mean a pair of classes (M, E) where $E \subseteq M^2$ is a well-founded binary relation which satisfies the axiom of extensionality. A *setlike class model* is a class model (M, E) such that for every $x \in M$,

$$\{y \in M : yEx\}$$

is a set.

Lemma 10.30 (Mostowski collapse lemma with Scott's trick). Assume that (M, E) is a setlike class model. Then there is a transitive class N such that (N, \in) is isomorphic to (M, E) .

The proof is the same as before. The assumption that the class model is setlike guarantees that each stage of the induction is a set.

The above construction allows us to take ultrapowers of the universe V .

Definition 10.31. If $j : M \rightarrow N$ is an elementary embedding, the *critical point* of j , $\text{crt } j$, is the ordinal α such that $j_U|V_\alpha$ is the identity and $j_U(\alpha) > \alpha$, if one exists.

Clearly if an elementary embedding has a critical point, then it is not the identity. So to show that we can nontrivial ultrapowers of V it suffices to find a critical point.

Theorem 10.32. Let U be a countably complete ultrafilter on a set A . Let $\kappa = \text{comp } U$ and let $j_U : V \rightarrow \text{Ult}(V, U)$ be the ultrapower map. Then:

1. $\text{crt } j_U = \kappa$.
2. $V_{\kappa+1}^{\text{Ult}(V, U)} = V_{\kappa+1}$.
3. $\text{Ult}(V, U)^\kappa \subseteq \text{Ult}(V, U)$.
4. If $A = \kappa$, then $\text{Ult}(V, U)^{\kappa+1} \not\subseteq \text{Ult}(V, U)$.
5. If $A = \kappa$, then $U \notin \text{Ult}(V, U)$.
6. If $A = \kappa$, then $V_{\kappa+2} \not\subseteq \text{Ult}(V, U)$.

Lemma 10.33. Under the hypotheses of Theorem 10.32, $j_U|_\kappa$ is the identity.

Proof. By induction. Assume that $\alpha < \kappa$ and if $\gamma < \alpha$ then $j_U(\gamma) = \gamma$. Let c_α be the constant function that returns α ; we must show $[c_\alpha]_U = \alpha$. If $\gamma < \alpha$ then $\gamma = [c_\gamma]_U \in [c_\alpha]_U$ by assumption. So $\alpha \subseteq [c_\alpha]_U$.

Conversely, assume $[f]_U \in [c_\alpha]_U$. By Los, this implies $\{x \in A : f(x) < \alpha\} \in U$. We claim there is a $\gamma < \alpha$ such that $\{x \in A : f(x) = \gamma\} \in U$. If not, $\bigcap_\gamma \{x \in A : f(x) \neq \gamma\} \in U$ since U is κ -complete, so U -almost every $x \in A$ is sent to $\geq \alpha$, a contradiction. So $[f]_U = \gamma < \alpha$. \square

Proof of Theorem 10.32. We show that $j_U|_{V_\kappa}$ is the identity by induction on α . This is clear for limits, so assume that $j_U|_{V_\alpha}$ is the identity. Let $x \in V_{\alpha+1}$; then $x \subseteq V_\alpha$, and any $y \in x$ satisfies $y \in V_\alpha$, hence $j_U(y) = y$ by induction. So $x \subseteq j_U(x)$. Conversely, $j_U(x) \subseteq V_{j_U(\alpha)}^{\text{Ult}(V, U)} \subseteq V_\alpha^{\text{Ult}(V, U)}$ by the lemma and induction. By induction, this implies $j_U(x) \subseteq x$. Therefore $j_U|_{V_{\alpha+1}} = \text{id}$.

Now we show $j_U(\kappa) > \kappa$. Thus we must find an $[f] < [c_\kappa]$ such that if $\alpha < \kappa$, $[c_\alpha] < [f]$. Since $\kappa = \text{comp } U$, there are $\{X_\alpha : \alpha < \kappa\}$ such that $X_\alpha \in U$ and $\bigcap_\alpha X_\alpha \notin U$. Let

$$f(x) = \begin{cases} \text{the least } \gamma \text{ such that } a \notin X_\gamma, \text{ if one exists} \\ 0, \text{ else.} \end{cases}$$

Then $f^{-1}(0) \cap \kappa \in U$, and $[f] < [c_\kappa]$. If $\alpha < \kappa$, $X_\alpha \in U$, so $\{\gamma < \kappa : f(\gamma) > \alpha\} \in U$. Therefore $[f] > [c_\alpha]$. So $j_U(\kappa) \neq \kappa$.

For the second claim, we have already shown $V_\kappa^{\text{Ult}(V, U)} = V_\kappa$, and since $\text{Ult}(V, U) \subseteq V$, $V_{\kappa+1}^{\text{Ult}(V, U)} \subseteq V_{\kappa+1}$. If $x \in V_{\kappa+1}$, then $x \subseteq V_\kappa$. Since j_U is an elementary embedding, $j_U(x) \subseteq V_{j_U(\kappa)}^{\text{Ult}(V, U)}$ and if $y \in x$, $j_U(y) \in j_U(x)$. So $x = j_U(x) \cap V_\kappa$ and hence $V_{\kappa+1} \subseteq V_{\kappa+1}^{\text{Ult}(V, U)}$.

For the third claim, let $\{[f_\alpha]\}_{\alpha < \kappa}$ be a κ -sequence. Let

$$g(a) = \{f_\alpha(a)\}_{\alpha < \kappa}.$$

Then $[g](\alpha) = [f_\alpha]$.

Now assume $A = \kappa$. For the fourth claim, we must show that $j_U|\kappa^+ \notin \text{Ult}(V, U)$. We first show that

$$j_U(\kappa^+) = \sup_{\alpha < \kappa^+} j_U(\alpha).$$

If not, then $\sup_{\alpha} j_U(\alpha) < j_U(\kappa^+)$. Let $[f]$ denote the left-hand side. Then for U -almost every α , $f(\alpha) < \kappa^+$. So there is a $\beta < \kappa^+$ so that the image of f is not contained in β . So $[f] < [c_\beta]$, which is impossible.

For the fourth and fifth claims, it suffices to show that if $U \in \text{Ult}(V, U)$ then $j_U|\kappa^+ \in \text{Ult}(V, U)$. Indeed, $j_U|\kappa^+$ is computable from U and the κ -sequences in κ^+ (which we already showed are in $\text{Ult}(V, U)$). \square

Theorem 10.34 (Scott-Keisler). Let κ be an ordinal. The following are equivalent:

1. κ is a measurable cardinal.
2. There is a transitive class M and an elementary embedding $j : V \rightarrow M$ such that $\kappa = \text{crt } j$.
3. There is a transitive class N and an elementary embedding $j : V_{\kappa+1} \rightarrow N$ such that $\kappa = \text{crt } j$.

Proof. $1 \implies 2 \implies 3$ is obvious from the above (here $N = M \cap V_{j(\kappa)+1}$). For the final direction, let

$$U = \{x \subseteq \kappa : \kappa \in j_U(x)\}.$$

We must show that U is a nontrivial, κ -complete ultrafilter. To see that U is nontrivial, if $\alpha < \kappa$, then $\kappa \notin \{\alpha\} = j_U(\{\alpha\})$. Clearly U is closed upwards and nonempty. To see that U is κ -complete, note that if $\gamma < \kappa$, then $j_U(\gamma) = \gamma$ so

$$j_U \bigcap_{\alpha < \gamma} X_\alpha = \bigcap_{\alpha < \gamma} j_U(X_\alpha) \ni \kappa.$$

\square

The fact that

$$U = \{x \subseteq \kappa : \kappa \in j_U(x)\}$$

is important in its own right.

10.4 Measurable cardinals are very large

We now show that measurable cardinals humiliate both Mahlo cardinals and L .

Theorem 10.35. Let κ be a measurable cardinal. Then κ is the κ th Mahlo cardinal in V .

Proof. Let C be a club in κ and let $j : V \rightarrow M$ be an elementary embedding such that $\text{crt } j = \kappa$. Then $j(C)$ is a club in $j(\kappa)$, so $j(C)$ is cofinal in $j(\kappa)$. But $j(\kappa) > \kappa$ and κ is a limit of C in κ^+ , hence in $j(\kappa)$, so $\kappa \in j(C)$.

Since κ is inaccessible,

$$M \models j(C) \text{ contains an inaccessible cardinal,}$$

so C actually contains an inaccessible cardinal. Therefore $S = \{\lambda < \kappa : \lambda \text{ is inaccessible}\}$ is stationary in κ , so κ is Mahlo.

Therefore

$$M \models \kappa \text{ is Mahlo,}$$

since $V_{\kappa+1} \subseteq M$. But $j(\kappa)$ is also a Mahlo cardinal, so $\kappa < j(\kappa)$. Now if $\alpha < \kappa$, $j(\alpha) = \alpha$ and so

$$M \models \text{There is a Mahlo cardinal } j(\lambda) \text{ such that } \alpha < j(\lambda) < j(\kappa).$$

Therefore in V , there is a Mahlo cardinal λ such that $\alpha < \lambda < \kappa$.

Since α was arbitrary, it follows that

$$\kappa = \sup\{\lambda < \kappa : \lambda \text{ is a Mahlo cardinal}\}$$

but κ is a regular cardinal, so the set of Mahlo cardinals under κ must have cardinality κ . \square

In particular, κ is the κ th inaccessible cardinal in V .

We now show that not only do measurable cardinals dwarf inaccessible cardinals, they dwarf Gödel's constructible universe.

Theorem 10.36. If there is a measurable cardinal then $V \neq L$.

Proof. Assume $V = L$ and let $j : V \rightarrow M$ be an elementary embedding, so that $\text{crt } j = \kappa$ is a measurable cardinal. Since L is the smallest inner model of V , $V = L = M$. Moreover $j(\kappa) > \kappa$, but

$$L = M \models j(\kappa) \text{ is the least measurable cardinal,}$$

a contradiction. \square

Maddy, Woodin, Steel, and other philosophers argue that the axiom “There is a proper class of measurable cardinals” is true because it is falsifiable, and yet nobody can falsify it. (Of course, since this axiom gives a proper class of models of ZFC, it is not provable). Thus we may conclude that $V \neq L$. This is disappointing, because L decides a lot of independent sentences, such as the continuum hypothesis. A major program in set theory is to find a L -like model, known as the Ultimate L , that decides a lot of sentences and is compatible with suitable large cardinals.

Chapter 11

One measurable cardinal

We find the minimal model $L[U]$ which contains L and a measurable cardinal. This model isn't the Ultimate L , because it doesn't even have two measurable cardinals; in fact, two measurable cardinals humiliate $L[U]$ in the same way that one measurable cardinal humiliated L , but it's a step in the right direction. In particular, $L[U]$ will satisfy GCH, have a definable well-ordering of \mathbb{R} , and otherwise decide lots of things that we could decide in L but not from ZFC alone.

11.1 Generalizing L

Throughout this section, we both use A to mean a fixed set, and a new constant symbol in our language which refers to that set.

Definition 11.1. Let A be a set. Let $\mathcal{P}_A(X)$ be the set of all subsets of X which are definable in the language (\in, A) with parameters in X .

Now let $L_0[A] = \emptyset$, $L_{\alpha+1}[A] = \mathcal{P}_A(L_\alpha[A])$, $L_\lambda[A] = \bigcup_{\alpha < \lambda} L_\alpha[A]$, and $L[A] = \bigcup_{\alpha \in \text{Ord}} L_\alpha[A]$. For any $x \in L[A]$, let $\text{rank}^{L[A]}$ be the least α such that $x \in L_\alpha[A]$.

Thus we expand the notion of definable power set to allow us to query whether a set is in A . Notice that $L[0] = L$ and in fact \mathcal{P}_0 is the definable powerset.

Example 11.2. If x is a real number such that $x \notin L$ (for example, x is a random real), then $L[x] \supset L$. In fact, $L_n[x] = L_n$, and so $L_\omega[x] = L_\omega$. But the formula $y : y \in x$ defines x as a subset of ω in the language (\in, x) , so $L_{\omega+1}[x] \neq L_{\omega+1}$.

We should show that $L[A]$ satisfies ZFC. This isn't as trivial as it seems.

Theorem 11.3 (Woodin?). Suppose that there are infinitely many Woodin cardinals in V and a measurable cardinal above them all. Then $L[\mathbb{R}]$ is not a model of ZFC, but rather a model of ZF, the Baire category theorem, and the axiom of determinacy.

Lemma 11.4. Let $M = (M, \in)$ be a transitive class. Then:

1. $M \models$ extensionality and foundation.

2. If $\emptyset, \omega \in M$, then $M \models$ empty set and infinity.
3. $M \models$ pairing, union, and powerset iff $\forall x, y \in M (\{x, y\}, \bigcup x, 2^x \in M)$.
4. $M \models$ replacement iff $\forall F \in \text{Def } M \cap M^M (\forall a \in M (F''a \in M))$ and $\forall F \in \text{Def } M \forall a \in M (F \cap a \in M)$.
5. If there is a definable well-ordering of M , then $M \models$ choice.

Here $\text{Def } M$ is the subclass of M consisting of all sets definable from parameters in M and $F''a = \{F(b) : b \in a\}$. The proof is essentially obvious.

Lemma 11.5. $L[A] \models \text{ZF}$.

Proof. $L[A]$ is a transitive class $\ni \emptyset$ by definition. Since ω is definable from parameters in $L_\omega[A]$, $\omega \in L_{\omega+1}[A]$. Similarly, $\{x, y\}, \bigcup x, 2^x$ are all definable from the parameters x, y , so they are in $L_{\alpha+1}[A]$ if $x, y \in L_\alpha[A]$. If $a \in L_\alpha[A]$ and $F \in \text{Def } L_\alpha[A] \cap L_\alpha[A]^{L_\alpha[A]}$ then $F''a$ is definable from the parameters a, F , so $F''a \in L_{\alpha+1}[A]$. Similarly for $F \cap a$. \square

Lemma 11.6. Let κ be an inaccessible cardinal. Then $L_\kappa[A] \models \text{ZF}$.

Proof. Same as above, with some modifications, as below:

For power set, let $x \in L_\kappa[A]$ and $\alpha < \kappa$ such that $x \in L_\alpha[A]$. Let $f : 2^x \cap L_\kappa[A] \rightarrow \kappa$ satisfy $f(y) = \text{rank}^{L[A]} y$. Since $2^x < \kappa$ and κ is regular, f is not cofinal, so f carries $2^x \cap L_\kappa[A]$ into some $\lambda < \kappa$. The part of replacement concerned with $F \in L_\kappa[A]^{L_\kappa[A]}$ is similar.

For replacement, first let $C \in \text{Def } L_\kappa[A]$ and $a \in L_\kappa[A]$. Since $C \in \text{Def } L_\kappa[A]$ there is a formula φ which defines $C \cap a$ in $L_\kappa[A]$ from a and a parameter $p \in L_\kappa[A]$. By reflection, there is an α such that $L_\alpha[A]$ is an elementary substructure of $L_\kappa[A]$, and then $a, p \in L_\alpha[A]$. Therefore $C \cap a \in L_{\alpha+1}[A]$. \square

Lemma 11.7. Let A be a set. There is a Π_2 sentence “ $V = L[A]$ ” in the language (\in, A) such that for any transitive class M , $M \models V = L[A]$ iff $M = L[A]$ or there is a limit ordinal λ such that $M = L_\lambda[A]$.

Proof. The sentence in question is

$$\forall x \exists \alpha (x \in L_{\alpha+1}[A]).$$

So we need $L_{\alpha+1}[A]$ to have a Σ_1 definition. Let $\text{DEF}(x, M)$ mean that x is the set of all $y \in M$ such that y is definable in (M, \in, A) without parameters, which is a Σ_1 formula, since satisfaction is a Σ_1 relation. Here we are using the fact that V_ω interprets **PA** to apply Gödel coding.

Let $\alpha \in \text{Ord}$. We uniquely define a function f , the $(L[A], \alpha)$ -rank function, by declaring that:

1. $\text{dom } f = \alpha + 1$.
2. $f(0) = 0$.
3. $\forall \gamma, \delta \leq \alpha$, if $\delta = \gamma + 1$ then $\text{DEF}(f(\gamma), f(\delta))$, and

4. $\forall \lambda \leq \alpha$, if λ is a limit ordinal then $f(\lambda) = \bigcup_{\beta < \lambda} f(\beta)$.

Since DEF was a Σ_1 formula, so is the definition of f .

Let

$$\text{LEV}(y, \alpha) = \alpha \in \mathbf{Ord} \wedge \exists f(f \text{ is the } (L[A], \alpha)\text{-rank function} \wedge f(\alpha) = y).$$

Then $\text{LEV}(y, \alpha)$ is a Σ_1 definition of $y = L_\alpha[A]$. Finally, let

$$(x \in L_{\alpha+1}[A]) = \exists y((x \in y) \wedge \text{LEV}(y, \alpha)).$$

To check that this works, first note that if λ is a limit ordinal, then $L_\lambda[A] \models \text{DEF}(L_{\alpha+1}[A], L_\alpha[A])$. It is easy and tedious to check that among transitive sets M , $M \models V = L[A]$ iff $M = L_\lambda[A]$ for some λ . \square

11.2 Solovay's theorem and condensation

Henceforth we consider the model $L[U]$, where U is a normal uniform ultrafilter on a measurable cardinal.

Definition 11.8. Let κ be a regular cardinal and U an ultrafilter. Then $L[U]$ is a κ -model if

$$L[U] \models U \text{ is a normal uniform ultrafilter on } \kappa.$$

In particular, if $L[U]$ is a κ -model, then

$$L[U] \models \kappa \text{ is a measurable cardinal}$$

since measurable cardinals are the only cardinals with normal uniform ultrafilters.

Lemma 11.9. If κ is a measurable cardinal, then there is a κ -model.

Proof. Let U be a normal uniform ultrafilter on κ . Then $U \cap L[U] \in L[U]$ since U is a set and $L[U]$ is transitive. Now if $x \in L_\alpha[U]$ then it is equivalent to query $x \in U$ and $x \in U \cap L[U]$. Let $U' = U \cap L[U]$; then $L[U'] = L[U]$.

We show that

$$L[U] \models U' \text{ is a normal uniform ultrafilter on } \kappa;$$

then $L[U] = L[U']$ is a κ -model. Since $\kappa \in U'$, U' is nonempty. If $x \in U'$ and $x \subseteq y \subseteq \kappa$, and $y \in L[U]$, then $y \in U$ since U is an ultrafilter. Similarly U' is κ -complete and maximal, hence a uniform ultrafilter.

It suffices to show that U' is normal. Let $f \in L[U] \cap \kappa^\kappa$ be such that $\{\alpha < \kappa : f(\alpha) < \alpha\} \in U'$. Then for U -almost every α , $f(\alpha) < \alpha$, so there is a $\beta \in \mathbf{Ord} \subseteq L[U]$ such that for U -almost every α , $f(\alpha) = \beta$, since U is normal. Therefore U' is normal. \square

We now show that $L[U]$ is very L -like.

Axiom 11.10 (global choice). There is a well-ordering of V .

This really needs to be stated in class theory, rather than set theory. Clearly global choice implies choice.

Lemma 11.11. If $L[U]$ is a κ -model then $L[U]$ is a model of global choice. Moreover, the well-ordering is definable if we are allowed an oracle that tells us the well-ordering of κ .

Proof. We show that there is a well-ordering of $L_\alpha[U]$ for any α , which is definable from the well-ordering of κ . These well-orderings will all cohere in that if $\alpha < \beta$ then the well-ordering of $L_\alpha[U]$ will extend to the well-ordering of $L_\beta[U]$. So we will be able to take a union over all ordinals.

Clearly $L_0[U]$ is well-ordered and we can take unions at limit stages. Now if $L_\alpha[U]$ is well-ordered, we have a lexicographic well-ordering on $L_\alpha[U]^{<\omega}$ and

$$L_{\alpha+1}[U] = \{x \in L_\alpha[U] : \exists[\varphi] \in V_\omega \exists b \in L_\alpha[U]^{<\omega} (L_\alpha[U] \models \varphi(x, b))\}$$

so the lexicographic well-orderings on $L_\alpha[U]^{<\omega}$ and on V_ω give a well-ordering of $L_{\alpha+1}[U]$. The only trouble is that we might throw in some sets from U , but if $x \in U$ then $x \subseteq \kappa$ so x is well-ordered in a canonical way. \square

This was basically the same proof as in L . We see the problem with $L[\mathbb{R}]$: if enough subsets of \mathbb{R} are determined, which happens if there are too many large cardinals, there is no good way to make their well-ordering cohere.

Definition 11.12. Let M be a model of ZFC, and for every $\alpha \in \text{Ord}^M$ let $M_\alpha = \{x \in M : \text{rank}^M x \leq \alpha\}$. Then we say that M has *condensation* if for every limit ordinal λ in M and elementary substructure X of M_λ , there is a limit ordinal ξ such that M_ξ is the transitive collapse of X .

Lemma 11.13. Every κ -model $L[U]$ has condensation.

Proof. Since X is well-founded and satisfies extensionality, it has a transitive collapse M . But then $M \models V = L[U]$, giving a ξ such that $M = L_\xi[U]$. \square

Again, the proof that L has condensation was basically the same. The significance is that if a model M has condensation, then it is easy to prove that M satisfies GCH at least on a tail end of **Card**. Unfortunately, as κ was measurable and hence $> 2^{\aleph_0}$, we will not be able to use condensation to prove that $L[U]$ satisfies the continuum hypothesis; we will need some much more powerful machinery to do that.

Lemma 11.14. Let $L[U]$ be a κ -model. For every $\alpha \geq \kappa$, $\text{card } L_\alpha[U] = \text{card } \alpha$.

Proof. By transfinite induction starting at κ , whose base case is a transfinite induction up to κ .

The base case is that $\text{card } L_\kappa[U] = \kappa$. We have $\kappa \leq \text{card } L_\kappa[U]$ since $\kappa = \sup \text{Ord}^{L_\kappa[U]}$. Conversely, we run a second transfinite induction up to κ . Clearly the claim $\text{card } L_\alpha[U] \leq \kappa$ holds at finite and limit stages (as long as $\alpha < \kappa$) since κ is a regular cardinal. If $\text{card } L_\alpha[U] \leq \kappa$ and $\omega \leq \alpha < \kappa$, then

$$\text{card } L_{\alpha+1}[U] = \text{card } L_\alpha[U] + \aleph_0 = \text{card } L_\alpha[U].$$

Here the first equality is because everything is definable from the previous stage, we aren't adding any subsets of κ yet, and there are \aleph_0 many definitions. This implies the base case. Limit stages are clear.

For the successor case, assume $\text{card } L_\alpha[U] = \text{card } \alpha$. Then

$$\text{card } L_\alpha[U] \leq \text{card } L_{\alpha+1}[U] \leq \kappa + \aleph_0 + \text{card } L_\alpha[U] = \text{card } L_\alpha[U]$$

since we are only adding subsets of κ and stuff definable from the previous stage. Therefore

$$\text{card}(\alpha + 1) = \text{card } \alpha = \text{card } L_\alpha[U] = \text{card } L_{\alpha+1}[U].$$

□

Theorem 11.15 (Solovay). Let $L[U]$ be a κ -model. Then

$$L[U] \models \kappa \text{ is the only measurable cardinal and GCH holds above } \kappa.$$

Proof. First we prove GCH. Let $\lambda \geq \kappa$; we must show $2^\lambda \leq \lambda^+$. We know that $\text{card } L_{\lambda^+} = \lambda^+$. Therefore we show $2^\lambda \cap L[U] \subseteq L_{\lambda^+}[U]$.

Let $A \in 2^\lambda \cap L[U]$ and let $\xi = \max(\text{rank}^{L[U]} A, \kappa, U)$, so $A \in L_\xi[U]$. By reflection, there is an elementary substructure H of $L_\xi[U]$ such that $A \in H$, $\lambda \subseteq H$, and $\text{card } A \in H$. Moreover, by taking a minimal such structure we can assume $\text{card } H = \lambda$. By condensation, there is a ξ' such that $L_{\xi'}[U]$ is the transitive collapse of H , and in particular $\xi' < \lambda^+$ and $A \in L_{\xi'}[U]$. Therefore $A \in L_{\lambda^+}[U]$.

Now let $V = L[U]$, and suppose that λ is a measurable cardinal equipped with a normal uniform ultrafilter U' , such that $\lambda \neq \kappa$. Let

$$j : V \rightarrow \text{Ult}(V, U')$$

be the elementary embedding given by U' . Then $U' \notin \text{Ult}(V, U')$, yet $U' \in V$; we will prove $\text{Ult}(V, U') = V$, a contradiction.

If $\lambda > \kappa$, then $\text{Ult}(V, U') = L[j(U)]$ since j must preserve the truth of the sentence $V = L[U]$. Since $\text{crt } j > \kappa$, $j(U) = U$, so $L[j(U)] = V$, as desired.

If $\lambda < \kappa$, let $M = \text{Ult}(V, U')$.

Lemma 11.16. One has

$$j(U) = U \cap M.$$

Proof of lemma. We claim that $j(U) \subseteq U$; then $j(U) \subseteq U \cap M$, and since $U \cap M \subseteq j(U)$ since j is an elementary embedding, this gives the lemma.

Let $[\cdot]$ denote the ultraprojection $V^\lambda \rightarrow M$. Suppose $[f] \in j(U)$, so that $f \in V^\lambda$; we must show $[f] \in U$. Let $j(U) = [c_U]$. By definition, for U' -almost every $\xi < \lambda$, $f(\xi) \in c_U(\xi) = U$, so we may modify f U' -almost nowhere and assume that $f \in U^\lambda$. Since $\lambda < \kappa$,

$$x = \bigcap_{\xi < \lambda} f(\xi) \in U.$$

Let $y = j(x)$, so $y \in j(U)$.

For every $\xi' < \lambda$ we have

$$\bigcap_{\xi < \lambda} f(\xi) \subseteq f(\xi')$$

so

$$y = \left[c_{\bigcap_{\xi < \lambda} f(\xi)} \right] \subseteq [f].$$

We now claim $y \in U$. To see this, let

$$I = \{\lambda < \xi < \kappa : \xi \text{ is inaccessible}\}.$$

Since κ is Mahlo, I is cofinal in κ ; since U is uniform, it follows that $I \in U$. Therefore $x \cap I \in U$, so x is cofinal in κ , and hence y is cofinal in $j(\kappa) > \kappa$, hence cofinal in κ .

Therefore $y \in U$, and so $[f] \in U$. Therefore $U' \subseteq U$. \square

We already showed that $L[U \cap L[U]] = L[U]$ in a previous lemma. Therefore the above lemma gives

$$M = L[U \cap M] = L[U] = V,$$

our desired contradiction. \square

11.3 Iterated ultrapowers

Our goal is to show that if $L[U]$ is the κ -model of a measurable cardinal κ then $L[U]$ acts a lot like L . In particular, we are going to prove that $L[U]$ thinks GCH is true. To do this, we need a lot more machinery.

We will mainly be interested in the case when M is a model of ZFC + “There is exactly one measurable cardinal,” but we need not assume this yet; we don’t even need to assume that M has power sets. We also don’t need to assume that the ultrafilter U is actually an element of M .

Definition 11.17. Let $M = (M, \in)$ be a transitive model of ZFC minus powerset. Let U be an ultrafilter. We say that M is κ -amenable to U if for all $x \in M$ such that $\text{card}^M x = \kappa$, $x \cap U \in M$. We say that M is *fully amenable* to U if M is κ -amenable to U for all κ .

Thus if U is actually in M , it follows that M is fully amenable. Here κ is a cardinal in M .

Definition 11.18. Let M be a transitive model of ZFC minus powerset. An M -ultrafilter on κ is an ultrafilter U on κ such that

$$M \models U \text{ is a normal uniform ultrafilter on } \kappa$$

and M is κ -amenable to U .

Under these hypotheses, we can again take the ultrapower $\Pi_U(M)$. It may be that κ is countable (even finite), in which case it may not be true that $\Pi_U(M)$ is well-founded. However, if $\Pi_U(M)$ is well-founded, we again define its transitive collapse $\text{Ult}(M, U)$ using Mostowski’s lemma.

Lemma 11.19. Let M be a transitive model of ZFC minus powerset and U an M -ultrafilter on κ . Assume that $\Pi_U(M)$ is well-founded, and let

$$j : M \rightarrow \text{Ult}(M, U) = M'$$

be the ultrapower map. Then:

1. j is cofinal.
2. $j|V_\kappa^M$ is the identity.
3. $V_{\kappa+1}^M = V_{\kappa+1}^{M'}$.
4. $U \notin M'$.
5. $U' = j(U)$ is a M' -ultrafilter.
6. $\text{card } M = \text{card } M'$.

Proof. To see that j is cofinal, let $[f] \in M'$; then if $g(x) \supseteq f(x)$ for all x , we have $[f] \in [g]$. The proofs that j restricts to the identity, preserves $V_{\kappa+1}$, and $U \notin M'$ are the same as before. Since U is a M -ultrafilter, j an elementary embedding, U' is a M' -ultrafilter.

For the cardinality claim, note that in M , $\text{card } U < \text{card } M$ (since M thinks itself is a proper class); in particular, for any $\alpha \in \text{Ord}^M$ sufficiently large, we have $\text{card } U < \text{card } M_\alpha$ in M . So in M ,

$$\text{card } M'_{j(\alpha)} = \text{card } M_\alpha^\kappa / U \leq \text{card } M_\alpha.$$

Since $\text{card } M'_{j(\alpha)} \geq \text{card } M_\alpha$ in M , M thinks that $\text{card } M_\alpha = \text{card } M'_{j(\alpha)}$.

We claim that if $M \models \text{card } x = \text{card } y$, then in fact $\text{card } x = \text{card } y$. To see this, note that since M is a transitive class, there is actually an honest-to-god bijection $f : x \rightarrow y$ between honest-to-god sets in M , and so f remains a bijection in V . Therefore, for all α , $\text{card } M_\alpha = \text{card } M'_{j(\alpha)}$. Therefore

$$\text{card } M = \sup_\alpha \text{card } M_\alpha = \sup_\alpha \text{card } M'_{j(\alpha)} = \sup_\alpha \text{card } M'_\alpha = \text{card } M',$$

where we used the fact that j is cofinal as a map $\text{Ord}^M \rightarrow \text{Ord}^{M'}$. \square

We now iterate the above construction. Let $M_0 = M$, and let $M_{n+1} = M'_n$. We will need to be able to pass to the direct limit for limit ordinals.

Lemma 11.20. Suppose that λ is a limit ordinal, and $(M_\alpha, U_\alpha, \kappa_\alpha)_{\alpha < \lambda}$ is a λ -sequence such that U_α is a M_α -ultrafilter on κ_α . Suppose further that we are given elementary embeddings $j_{\alpha, \beta} : M_\alpha \rightarrow M_\beta$ if $\alpha < \beta < \lambda$, where $j_{\alpha, \alpha+1}$ is the ultrapower map for U_α , and the diagrams

$$\begin{array}{ccc} M_\alpha & \xrightarrow{j_{\alpha, \beta}} & M_\beta \\ & \searrow j_{\alpha, \gamma} & \downarrow j_{\beta, \gamma} \\ & & M_\gamma \end{array}$$

all commute. Suppose further that the direct limit of the elementary directed system $\{j_{\alpha,\beta} : \alpha < \beta < \lambda\}$ is well-founded, and let $(M_\lambda, U_\lambda, \kappa_\lambda)$ be its transitive collapse, and let

$$j_{\alpha,\lambda} : M_\alpha \rightarrow M_\lambda$$

be the induced elementary embeddings.

Then the ordinal κ_λ such that $\kappa_\lambda = j_{\alpha,\lambda}(\kappa_\alpha)$ for every α is well-defined, and U_λ is a M_λ -ultrafilter on κ_λ .

Proof. Since the diagrams commute, $j_{\alpha,\lambda}(\kappa_\alpha)$ does not depend on the choice of α . Since $j_{\alpha,\lambda}$ is an elementary embedding and U_α is a M_α ultrafilter on κ_α , the second claim also holds. \square

Definition 11.21. With $(M_\alpha, U_\alpha, \kappa_\alpha, j_{\alpha,\beta})_{\alpha < \beta < \tau}$ as in the previous lemma iterated up to τ , we call (M_0, U_0, κ_0) *initial data for the iterated ultrapower construction* and $(M_\alpha, U_\alpha, \kappa_\alpha, j_{\alpha,\beta})_{\alpha < \beta < \tau}$ the *iteration* of the initial data.

Notice that we cannot guarantee, just from the assumptions we make above on initial data, that for every α , M_α is still well-founded. If this is the case, then we cannot take the transitive collapse, and the above induction stops prematurely.

Definition 11.22. Let (M_0, U_0) be initial data for the iterated ultrapower construction and (M_α, U_α) its iteration. Let τ be the first ordinal such that M_τ is not well-founded, or $\tau = \text{Ord}$ if no such ordinal exists. If $\lambda < \tau$, we say that (M_0, U_0) is λ -iterable. If $\tau = \text{Ord}$, so (M_0, U_0) is λ -iterable for all λ , we say that (M_0, U_0) is iterable.

In the proof of GCH we will mainly use ω_1 -iterable initial data. In fact, we will prove that ω_1 -iterable initial data is actually just iterable.

Lemma 11.23. Let (M_0, U_0, κ_0) be initial data for the iterated ultrapower construction and $(M_\alpha, U_\alpha, \kappa_\alpha)$ its iteration. Assume that τ is the first ordinal such that M_τ is not well-founded. For every $\alpha < \beta < \tau$:

1. $\text{crt } j_{\alpha,\beta} = \kappa_\alpha$ and $j_{\alpha,\beta}(\kappa_\alpha) = \kappa_\beta$.
2. $j_{\alpha,\beta}|V_{\kappa_\alpha} \cap M_\alpha$ is the identity and

$$V_{\kappa_{\alpha+1}} \cap M_\alpha = V_{\kappa_{\alpha+1}} \cap M_\beta.$$

3. If λ is a limit ordinal then

$$\kappa_\lambda = \sup_{\xi < \lambda} \kappa_\xi.$$

4. If M_0 is a set then M_α is a set and $\text{card } M_\alpha = \text{card } M_0 \text{ card } \alpha$.

Proof. We already proved 1 and 2.

For 3, let $\xi < \kappa_\lambda$; we will find a η such that $\kappa_\eta > \xi$. Since the ultrapower map cannot decrease κ , this will suffice. In fact let ξ', η satisfy

$$\xi = j_{\eta,\lambda}(\xi').$$

Then $M_\eta \models \kappa_\eta > \xi'$ and since $\text{crt } j_{\eta,\lambda} = \kappa_\eta$ it follows that $\xi' = \xi$.

For 4, cardinality does not increase at successor stages, and at limit stages λ we take a union of λ many models. So

$$\text{card } M_\alpha \leq \text{card } M_0 \text{ card } \alpha.$$

Moreover $\text{card } M_0 \leq \text{card } M_\alpha$ and $\{\kappa_{\alpha'} : \alpha' < \alpha\} \in M_\alpha$ so $\text{card } M_\alpha \leq \text{card } M_0$. \square

Lemma 11.24. Let (M_0, U_0) be initial data for the iterated ultrapower construction and (M_α, U_α) its iteration. If M_0 is $\alpha + 1$ -iterable and $x \in \mathcal{P}(\kappa_\alpha) \cap M_\alpha$, then $x \in U_\alpha$ iff $\kappa_\alpha \in j_{\alpha,\alpha+1}(x)$.

Proof. We can collapse this down to the case $\alpha = 0$. Then the claim is that $x \in U$ iff $\kappa \in j_U(x)$, which follows from the proof of the Scott-Keisler theorem. \square

Lemma 11.25. Let (M_0, U_0) be initial data for the iterated ultrapower construction and (M_α, U_α) its iteration. Assume (M_0, U_0) is λ -iterable and λ is a limit ordinal. Then $x \in U_\lambda$ iff there is an $\alpha < \lambda$ such that $\{\kappa_\gamma : \alpha < \gamma < \lambda\} \subseteq x$.

Proof. Let $\gamma < \lambda, x'$ be such that

$$x = j_{\gamma,\lambda}(x').$$

Since $\text{crt } j_{\gamma+1,\lambda} \geq \kappa_\lambda$,

$$x' \in U_\gamma \leftrightarrow \kappa_\gamma \in j_{\gamma,\gamma+1}(x') \leftrightarrow \kappa_\gamma \in x.$$

\square

Lemma 11.26. Let (M_0, U_0) be initial data for the iterated ultrapower construction and (M_α, U_α) its iteration. If (M_0, U_0) is α -iterable and $x \in M_\alpha$ there are $\xi_1 < \dots < \xi_n < \alpha$ such that

$$\kappa_{\xi_1} < \dots < \kappa_{\xi_n}$$

and a function $f : \kappa_0^n \rightarrow M_0$, which exists in M_0 , such that

$$x = j_{0,\alpha}(f)(\kappa_{\xi_1}, \dots, \kappa_{\xi_n}).$$

Proof. By induction on α . Let $x \in M_{\alpha+1}$. Then there is a $g : \kappa_\alpha \rightarrow M_\alpha$ such that

$$x = j_{\alpha,\alpha+1}(g)(\kappa_\alpha).$$

Now find $g' : \kappa_0^n \rightarrow M_0$ in M_0 and ξ_i such that $g = j_{0,\alpha}(g')(\kappa_{\xi_1}, \dots, \kappa_{\xi_n})$ by induction. Then let

$$f(\gamma_1, \dots, \gamma_{n+1}) = (g'(\gamma_1, \dots, \gamma_n))(\gamma_{n+1}),$$

which works. The limit case follows by induction too, and we only need finitely many entries before the limit stage so the index n doesn't run over ω . \square

This is useful because it allows us to do combinatorial tricks to find the cardinality of the set of all x in the above lemma which satisfy some condition. For example, consider the following lemma, which allows us to bound above the image of $j_{0,\alpha}$ in several arguments that follow.

Lemma 11.27. Let (M_0, U_0) be initial data for the iterated ultrapower construction and (M_α, U_α) its iteration. Assume (M_0, U_0) is α -iterable. For any $\gamma \in \text{Ord}^M$,

$$j_{0,\alpha}(\gamma) < (\text{card}(\gamma^{\kappa_0} \cap M) \text{card} \alpha)^+.$$

Proof. If $\gamma \leq \kappa_0$ then $j_{0,\alpha}$ sends γ to itself, and

$$\gamma < \text{card}(\gamma^{\kappa_0})^+$$

by cardinal arithmetic.

If $\gamma > \kappa_0$, recall $j_{0,\alpha}(\kappa_0) = \kappa_\alpha$, and $j_{0,\alpha}$ blows up γ to be much larger than κ_0 . Let $\eta \in (\kappa_\alpha, j_{0,\alpha}(\gamma))$. By Lemma 11.26 we may write

$$\eta = j_{0,\alpha}(f)(\kappa_{\xi_1}, \dots, \kappa_{\xi_n}).$$

Since $\eta < j_{0,\alpha}(\gamma)$, elementarity implies that f maps into γ .

The number of possible choices of η is the number $\text{card} \gamma^{\kappa_0} \cap M$ of possible choices of f times the number of possible choices of inputs for f , $\text{card} \alpha^n = \text{card} \alpha$. \square

Lemma 11.28. Let (M_0, U_0) be initial data for the iterated ultrapower construction and (M_α, U_α) its iteration. Assume (M_0, U_0) is ν -iterable, ν is a cardinal, and

$$\text{card} \kappa_0^{\kappa_0} \cap M < \nu.$$

Then $\kappa_\nu = \nu$.

Proof. By a previous lemma we have $\nu \leq \kappa_\nu$. On the other hand,

$$\kappa_\nu = j_{0,\nu}(\kappa_0) < (\nu \text{card}(\nu^{\kappa_0} \cap M))^+$$

and by assumption this implies that $\kappa_\nu \leq \nu$. \square

Lemma 11.29. Let (M_0, U_0) be initial data for the iterated ultrapower construction and (M_α, U_α) its iteration. Suppose $\theta \in \text{Ord}^M$, $\alpha < \theta$, (M_0, U_0) is α -iterable, $\text{cof}^M \theta > \alpha$, and

$M \models \theta$ is a strong limit cardinal, and for every $\theta' < \theta$ there are functions $\kappa_0 \rightarrow \theta'$.

Then

$$j_{0,\alpha}(\theta) = \theta.$$

Proof. We have $\theta \leq j_{0,\alpha}(\theta)$; to show the opposite, we show that if $\eta < j_{0,\alpha}(\theta)$, $\eta < \theta$. By Lemma 11.26, we can write

$$\eta = j_{0,\alpha}(f)(\kappa_{\xi_1}, \dots, \kappa_{\xi_n}).$$

As in previous arguments, we can assume that f maps into θ . Since $\text{cof}^{M_0}(\theta) > \alpha$, there is a $\theta' < \theta$ such that f maps into θ' . Since θ is a strong limit cardinal,

$$\kappa_0^{\text{card} \theta'} = 2^{\text{card} \theta'} < \theta.$$

Therefore

$$\eta < j_{0,\alpha}(\theta') < (\text{card}((\theta')^{\kappa_0} \cap M) \text{card} \alpha)^+ \leq \theta.$$

\square

11.4 Suitable initial data for iteration

Thus we have built up a lot of machinery about how the iterated ultrapower construction moves ordinals. Now we consider when initial data is actually **Ord**-iterable.

Definition 11.30. A *weakly countably complete ultrafilter* U is an ultrafilter such that for any ω -sequence X_n in U , $\bigcap_n X_n$ is nonempty.

Definition 11.31. Let (M, U) and (N, W) be initial data. We say that (N, W) is *suitable initial data* if N is countable, W is an N -ultrafilter, and there is an elementary embedding $N \rightarrow M$ sending W to U .

Theorem 11.32. Let (M, U) and (N, W) be initial data. If U is weakly countably complete in V and (N, W) is suitable initial data, then (N, W) is $\omega_1 + 1$ -iterable.

Proof. Let W be an N -ultrafilter on ν , and let $(N_\alpha, W_\alpha, \nu_\alpha, j_{\alpha, \beta})$ be its iteration. It suffices to show that if $\alpha < \omega_1$, then there is an elementary embedding $N'_{\alpha+1} \rightarrow M$, where $N'_{\alpha+1}$ is the uncollapsed ultrapower of N_α ; since M is well-founded, so will be $N'_{\alpha+1}$ (since it's impossible to embed a non-well-founded relational structure in a well-founded relational structure). Then we can take the transitive collapse $N_{\alpha+1}$ and keep the induction going. Limit stages are no problem because the direct limit of a chain of well-founded relational structures is well-founded.

As it turns out, we can make the diagram

$$\begin{array}{ccc} & M & \\ \sigma_\alpha \uparrow & \swarrow & \\ N_\alpha & \xrightarrow{\sigma'_{\alpha+1}, j_{\alpha, \alpha+1}} & N'_{\alpha+1} \end{array}$$

commute by taking

$$\sigma'_{\alpha+1}([f]) = \sigma_\alpha(f)(\eta_\alpha)$$

for any

$$\eta_\alpha \in \bigcap_{x \in W_\alpha \cap N_\alpha} \sigma_\alpha(x).$$

Such an η_α exists since there are only countably many sets $\sigma_\alpha(x)$, since N_α is countable, and U is a weakly countably complete ultrafilter in V , and W_α maps to U in the elementary embedding σ_α . \square

Corollary 11.33. Let (M, U, κ) be initial data such that

$$M \models V_{\kappa+1} \text{ exists,}$$

and $U \in M$, and let (M_α, U_α) be its iteration. The following are equivalent:

1. (M, U) is iterable.
2. (M, U) is ω_1 -iterable.

3. (M, U) is $\omega_1 + 1$ -iterable.
4. There is an α such that (M_α, U_α) is weakly countably complete in V .
5. For all initial data (N, W) , if there is an elementary embedding $N \rightarrow M$ sending W to U , then (N, W) is iterable.

Proof. 1 \implies 2 and 5 \implies 1 are clear. 2 \implies 3 holds because limit stages don't break iterability.

For 3 \implies 4, we are given that $(M_{\omega_1}, U_{\omega_1})$ is well-founded. Let X_n be an ω -sequence in U_{ω_1} . Then for every $n < \omega$ there is an α_n such that

$$\{\kappa_\gamma : \alpha_n < \gamma < \omega_1\} \subseteq X_n.$$

Let $\beta = \sup_{n < \omega} \alpha_n$; then $\beta < \omega_1$ and $\bigcap_{n < \omega} X_n$ contains $\kappa_{\beta+1}$. So take $\alpha = \omega_1$.

For 4 \implies 1 and 1 \implies 5, use the previous theorem to show that (M_α, U_α) is iterable and hence we have inclusions $(N, W) \rightarrow (M, U) \rightarrow (M_\alpha, U_\alpha)$. \square

Corollary 11.34. Let (N, \in) be a transitive model of ZFC and Let (M, U, κ) be initial data such that

$$M \models V_{\kappa+1} \text{ exists,}$$

and $U \in M$. Assume $\omega_1 \subseteq N$, and

$$N \models U \text{ is an } M\text{-ultrafilter on } \kappa.$$

Then (M, U) is iterable iff

$$N \models (M, U) \text{ is iterable.}$$

Proof. It follows by the above corollary, since N gets ω_1 correct. \square

The point of the above results is that ω_1 -iterability is the same as **Ord**-iterability, and this is because we are mainly interested in countably complete ultrafilters (since these already give measurable cardinals and well-foundedness).

11.5 Mice

Definition 11.35 (Jensen). Let κ be a regular cardinal. A *premouse* at κ is a pair (M, U) such that $M = (M, \in)$ is a transitive class such that:

1. $M \models \text{ZFC minus powerset,}$
2. $M \models V_{\kappa+1} \text{ exists,}$
3. $M \models V = L[U], \text{ and}$
4. $M \models U \text{ is a normal uniform ultrafilter on } \kappa.$

A *mouse* is an iterable premouse.

Since $M \models V = L[U]$ and is a transitive class, $M = L[U]$ or $M = L_\lambda[U]$ for some limit ordinal $\lambda > \kappa$. Therefore we abuse notation $L_{\mathbf{Ord}}[U] = L[U]$ (so \mathbf{Ord} is an “uncountable limit ordinal”).

Lemma 11.36. Let κ be a regular cardinal, λ a limit ordinal or $\lambda = \mathbf{Ord}$, and $L[U]$ a κ -model. Then $(L_\lambda[U], U)$ is a premouse at κ . If λ is uncountable, then $L_\lambda[U]$ is actually a mouse.

Proof. Most of the definition of a premouse follows straight from the definitions. Since $L[U]$ is actually a model of ZFC, in particular it thinks that the power set of V_κ exists.

If λ is a limit ordinal, then $U \cap L_\lambda[U]$ is countably complete, so weakly countably complete. Therefore $L_\lambda[U]$ is iterable, and hence a mouse. \square

Let ν be an uncountable regular cardinal, so its club filter C_ν is a nontrivial ν -complete filter. It is not necessarily true that C_ν is actually an ultrafilter, so this does not imply that ν is measurable, but sometimes a model will get this wrong and think that club filters are witnesses to measurability. This is the case when we iterate a mouse at κ more than κ^κ times.

Lemma 11.37. Suppose (M, U) is a mouse at κ and $\nu > (\kappa^\kappa)^M$ is a regular cardinal. Then there is a λ such that

$$(M_\nu, U_\nu) = (L_\lambda[C_\nu], C_\nu \cap L_\lambda[C_\nu]).$$

Proof. We have $\kappa_\nu = \nu$ and $M_\nu = L_\lambda[U_\nu]$ by previous results. This is where we use $\nu > (\kappa^\kappa)^M$.

Since U is a normal uniform ultrafilter, so is U_ν , so for every $x \in U_\nu$, $\text{card } x = \kappa_\nu = \nu$. By a previous lemma, there is an $\alpha_x < \nu$ such that

$$f(x) = \{\kappa_\gamma : \alpha_x < \gamma < \nu\} \subseteq x.$$

We claim that $f(x)$ is a club in ν . Since ν is a limit ordinal,

$$\nu = \sup_{\gamma < \nu} \kappa_\gamma = \sup f(x),$$

so $f(x)$ is cofinal. If β is a limit point of $f(x)$ in ν , let

$$\mu = \sup \{\gamma : \kappa_\gamma < \beta\}.$$

Since β is a limit ordinal, $\kappa_\mu = \beta$. This verifies the claim.

Therefore x contains a club; but x was arbitrary, so $U_\nu \subseteq C_\nu \cap L_\lambda[C_\nu]$. Since U_ν is maximal, the claim follows. \square

Corollary 11.38. Suppose there is a κ -model, and η is a sufficiently large regular cardinal. Then $(L_\eta[C_\eta], C_\eta \cap L_\eta[C_\eta])$ is an η -model.

Proof. Let $L[U]$ be a κ -model. Then $L[U]$ is a mouse, so is $(\eta + 1)$ -iterable. By the previous lemma, if $\eta > (\kappa^\kappa)^{L[U]}$, then

$$(L[U]_\eta, U_\eta) = (L[C_\eta], C_\eta \cap L[C_\eta]),$$

Since $L[U]_\eta$ is a mouse and U_η is a normal uniform ultrafilter on $\kappa_\eta = \eta$,

$$L[U]_\eta \models \eta \text{ is an inaccessible cardinal}$$

so we may replace $L[C_\eta]$ with $L_\eta[C_\eta]$ in the above equality. \square

Here we are not assuming that κ is measurable.

The reason that mice are of interest is that they are models of set theory with large cardinals that can be sorted into a nice linear ordering by letting them iterate. This is not true for large cardinals that are even bigger than measurable cardinals, because the mice for their inner models actually sort into trees or other more complicated combinatorial structures.

Definition 11.39. Suppose (M, U) and (M', U') are premice at κ and κ' respectively. We say that $(M, U) \leq (M', U')$ if:

1. $\kappa = \kappa'$.
2. $U = U' \cap M$.
3. There is an $\alpha \in \mathbf{Ord}^M$ such that $M = L_\alpha[U']$.

Notice that the first condition $\kappa = \kappa'$ means that the mice are comparable. The second and third conditions basically say that (M, U) is a submouse of (M', U') .

Theorem 11.40 (comparison). Let (M, U, κ) and (M', U', κ') are mice. Let

$$\nu > \text{card}(\kappa^\kappa \cap M) + \text{card}((\kappa')^{\kappa'} \cap M')$$

be a regular cardinal. Then (M_ν, U_ν) and (M'_ν, U'_ν) are comparable mice.

Proof. We have

$$(M_\nu, U_\nu) = (L_\alpha[C_\nu], C_\nu \cap L_\alpha[C_\nu])$$

and similarly for (M'_ν, U'_ν) . In particular both are mice at ν . If $\alpha \leq \alpha'$ then $(M_\nu, U_\nu) \leq (M'_\nu, U'_\nu)$. \square

11.6 Applications of mice

In this section we use mice to prove two very powerful theorems about $L[U]$.

Theorem 11.41. Let κ be a measurable cardinal and $L[U]$ a κ -model. Then $L[U]$ satisfies the generalized continuum hypothesis.

Theorem 11.42. Let κ be a measurable cardinal and $L[U]$ a κ -model. Then $L[U]$ thinks \mathbb{R} has a Σ_3 -definable well-ordering.

So $L[U]$ looks a lot like L ! In fact, L also satisfied GCH and thought that \mathbb{R} had a Δ_2 -definable well-ordering. For the rest of our adventures with large cardinals, we will mainly be looking for inner models of V which generalize $L[U]$ to satisfy stronger large cardinal axioms than the relatively weak axiom “There is one measurable cardinal.”

If M is a model of global choice, let $<_M$ denote the given well-ordering on M . In particular, this will happen when M is a κ -model.

Lemma 11.43. Let κ be a measurable cardinal, $V = L[U]$, and let $(M^1, U_1), (M^2, U_2)$ be mice at κ . Let λ be a cardinal such that $\text{crt } U_1 > \lambda$ and $\text{crt } U_2 > \lambda$. Then either $((2^\lambda)^{M^1}, <_{M^1})$ is an initial segment of $((2^\lambda)^{M^2}, <_{M^2})$ or vice versa.

Proof. There is a regular cardinal θ such that M_θ^1 and M_θ^2 are comparable mice. Without loss of generality, assume that $M_\theta^1 \leq M_\theta^2$. By assumption on λ , if $x \subseteq \lambda$ and $x \in M^1 \cap M^2$ then x is a fixed point of the elementary embeddings $j_i : M^i \rightarrow M_\theta^i$. Therefore $<_{M^i}$ is an initial segment of $<_{M_\theta^i}$, and since $M_\theta^1 \leq M_\theta^2$, $<_{M_\theta^1}$ is an initial segment of $<_{M_\theta^2}$. So the $<_{M^i}$ are initial segments of $<_{M_\theta^2}$ and hence one must be an initial segment of each other. \square

Proof of GCH. Assume $V = L[U]$. By Solovay’s theorem, $\lambda^+ = 2^\lambda$ when $\lambda \geq \kappa$.

Let $\lambda < \kappa$ and $x \in 2^\lambda$, so by Solovay’s theorem, λ is not a measurable cardinal. Let γ be a limit ordinal such that $x, U \in L_\gamma[U]$, $L_\gamma[U]$ thinks $V_{\kappa+1}$ exists, and $L_\gamma[U]$ models ZFC minus powerset. By reflection, there is an elementary substructure $(H, U \cap H)$ of $L_\gamma[U]$ such that $x, U \in H$ and $\text{card } H = \lambda$. Then $(H, U \cap H) \models V = L[U]$, so there is a limit ordinal δ such that the transitive collapse of $(H, U \cap H)$ is $L_\delta[U']$ for some ultrafilter U' .

Since κ is measurable, δ is uncountable, so $M = L_\delta[U']$ is a mouse, and $\text{card } M = \lambda$. Besides, $x \in M$, and since λ is not a measurable cardinal, $\text{crt } U' > \lambda$. Since V is also a mouse, the lemma implies that $<_M$ is an initial segment of $<_V$. So

$$\text{card}\{y \in 2^\lambda : y <_V x\} = \text{card}\{y \in 2^\lambda : y <_M x\} \leq \text{card } M = \lambda.$$

Therefore any initial segment of $(2^\lambda, <_V)$ has cardinality λ . But this implies that $\text{card } 2^\lambda \leq \lambda^+$. Cantor’s theorem gives the other direction. \square

Proof of Σ_3 -well-ordering. Let $V = L[U]$. We take $\lambda = \aleph_0$ in the above lemma. Then $x <_V y$ iff there is (\exists) a countable premouse M such that for every (\forall) countable ordinal α , M is α -iterable, and $x <_M y$. Being α -iterable is a Π_2 -condition: for all $\beta < \alpha$ and every \in -chain C in M_β , there is a $n < \omega$ such that $\text{card } C = n$. Moreover a formula of the form $\exists \forall \Pi_2$ is Σ_3 . \square

Chapter 12

Extender models

Measurable cardinals are nice, but as we saw above, they have some serious deficiencies. Putting aside for now the issue that the canonical model $L[U]$ could only hold a single measurable cardinal, let's see what else is "wrong" with them.

Definition 12.1. A regular cardinal κ is an α -strong cardinal if there is an elementary embedding $j : V \rightarrow M$ such that $\text{crt } j = \kappa$, $j(\kappa) > \alpha$ and $V_\alpha \subseteq M$. A *strong cardinal* is one which is α -strong for every $\alpha \in \text{Ord}$.

Therefore κ is measurable iff κ is $\kappa + 1$ -strong. But not every measurable cardinal is $\kappa + 2$ -strong, since if U was the ultrafilter

$$U = \{x \subseteq \kappa : \kappa \in j(x)\}$$

where j witnessed measurability of κ , then $U \notin V_{\kappa+2}^M$.

Definition 12.2. A regular cardinal κ is a λ -supercompact cardinal if there is an elementary embedding $j : V \rightarrow M$ such that $j(\kappa) > \lambda$ and $M^\lambda \subseteq M$. A *supercompact cardinal* is one which is λ -supercompact for every $\lambda \in \text{Card}$.

Again we see that a measurable cardinal is κ -supercompact but in general is no better, since if j is a witness to the measurability of κ , then $j|_{\kappa^+} \notin M$, so that κ may not be κ^+ -supercompact.

This is frustrating: we had a clean definition of measurability in terms of combinatorics (or "measure theory") and one would hope that we could define supercompactness similarly. But clearly ultrapowers are not enough to define supercompactness, so we need some other formulation.

12.1 The diagram chase

Motivated by the above, we introduce another way to construct elementary embeddings from large cardinals, which will hopefully allow us to come up with a better definition of supercompactness.

Definition 12.3. Let X be a class. The *Skolem hull* $\text{Hull } X$ of X is the intersection of all classes Y such that Y is an elementary substructure of V and $X \subseteq Y$.

Let $j : V \rightarrow M$ be a nontrivial elementary embedding, $\text{crt } j = \kappa$, and

$$U = \{x \subseteq \kappa : \kappa \in j(x)\}$$

the normal uniform ultrafilter given by j . Let i be the ultrapower map of U . Then we have an elementary diagram

$$\begin{array}{ccc} V & \xrightarrow{j} & M \\ & \searrow i & \nearrow \psi \\ & \text{Ult}(V, U) & \end{array}$$

where the map ψ satisfies

$$\psi([f]) = j(f)(\kappa).$$

Moreover,

$$\psi''(\text{Ult}(V, U)) = \text{Hull}^M(j''(V) \cup \{\kappa\})$$

where as usual,

$$f''(x) = \{f(y) : y \in x\}.$$

So, given an elementary embedding j , we factored j into maps i, ψ where i is an ultrapower embedding and ψ embeds the ultrapower given by i as an appropriate Skolem hull. This is basically just the first isomorphism theorem: any ultrapower must contain $j''V$ and $\{\kappa\}$, but we could also embed this ultrapower in a much bigger structure.

Suppose that we want to run the above process but with the critical point κ replaced by a possibly much larger cardinal λ , which would witness the δ -supercompactness of κ for any $\delta < \lambda$. Indeed, if κ is δ -supercompact, we can find a cardinal $\lambda > \delta$ and an elementary embedding $j : V \rightarrow M$ such that $\text{crt } j = \kappa$ and $j(\kappa) = \lambda$. Let $H = \text{Hull}^M(j''V \cup \{\lambda\})$ and let N be the transitive collapse of H . Then we have an elementary embedding $\psi : N \rightarrow M$ such that $\psi''N = H$, and since for every $\alpha \leq \lambda$, $\alpha \in H$, it must be that $\text{crt } \psi > \lambda$. Then we can find an elementary embedding $i : V \rightarrow N$ which makes the elementary diagram

$$\begin{array}{ccc} V & \xrightarrow{j} & M \\ & \searrow i & \nearrow \psi \\ & N & \end{array}$$

commute: just take $i = j \circ \psi^{-1}$; since $\psi''N \subseteq j''V$, this is well-defined. If we can show that i is an ultrapower embedding, we have accomplished our goal. In fact, we will show that i is the direct limit of a λ -sequence of ultrapower embeddings.

Definition 12.4. Let $j : V \rightarrow M$ be an elementary embedding, $\text{crt } j = \kappa$, and $\lambda > \kappa$ regular cardinals. We say that an elementary embedding $i : V \rightarrow N$ is the (κ, λ) -derived extender for j if:

1. N is the transitive collapse of $H = \text{Hull}^M(j''V \cup \{\lambda\})$.

2. The elementary embedding $\psi : N \rightarrow M$ such that ψ^{-1} is the canonical isomorphism $H \rightarrow N$ makes the elementary diagram

$$\begin{array}{ccc} V & \xrightarrow{j} & M \\ & \searrow i & \nearrow \psi \\ & N & \end{array}$$

commute.

Fix λ and suppose we have an elementary embedding $j : V \rightarrow M$. We want to modify j to an ultrapower embedding $i : V \rightarrow N$ so that $\text{crt } i \geq \lambda$, such that i is the (κ, λ) -derived extender of j .

To this end, let N be the transitive collapse of a Skolem hull H in M which contains all of λ as well as $j''V$. For every $\alpha < \lambda$, let

$$\kappa_\alpha = \min_{j(\beta) \geq \alpha} \beta.$$

Then consider the κ_α -complete ultrafilter

$$E_\alpha = \{x \subseteq \kappa_\alpha : \alpha \in j(x)\}.$$

Indeed, E_α is pushed forward by j to a trivial ultrafilter. So $\text{Ult}(V, E_\alpha)$ is well-founded and we obtain an elementary diagram:

$$\begin{array}{ccc} V & \xrightarrow{j} & M \\ \downarrow & \nearrow j_\alpha & \uparrow \psi \\ \text{Ult}(V, E_\alpha) & \xrightarrow{k_\alpha} & N \end{array}$$

Here ψ^{-1} is the transitive collapse $H \rightarrow N$. Since $\alpha < \lambda$, $j_\alpha''(\text{Ult}(V, E_\alpha)) \subseteq j''(V) \subseteq H = \psi''N$. Therefore k_α exists and the diagram commutes.

We want to show that N is the direct limit of some infinite elementary diagram containing the $\text{Ult}(V, E_\alpha)$.

It will be convenient to introduce some notation.

Definition 12.5. Let x be a set and β a cardinal. We define

$$\binom{x}{\beta} = \{y \in 2^x : \text{card } y = \beta\}.$$

If $n < \omega$, $a \subset \lambda$, and $\text{card } a = n$, we put

$$E_a = \{x \in \binom{\kappa_a}{n} : a \in x\}.$$

Here

$$\kappa_a = \sum_{\alpha \in a} \kappa_\alpha.$$

Then we have a natural elementary embedding $\text{Ult}(V, E_a) \rightarrow \text{Ult}(V, E_b)$ whenever $a \subseteq b$: if $[f]_{E_a}$, then

$$\text{dom } f \subseteq \binom{\lambda}{\text{card } a},$$

i.e. f takes in $\text{card } a$ arguments. We can therefore define $f' : \binom{\lambda}{\text{card } b} \rightarrow V$ by

$$f'(x_1, \dots, x_{\text{card } a}, \dots, x_{\text{card } b}) = f(x_1, \dots, x_{\text{card } a}).$$

Then $[f']_{E_b} \in \text{Ult}(V, E_b)$, so the map $[f]_{E_a} \mapsto [f']_{E_b}$ is the elementary embedding we wanted. Of course, we have $E_\alpha = E_{\{\alpha\}}$, so the directed system of all E_a , $a \in \binom{\lambda}{<\omega}$, includes the E_α we constructed already. Taking the limit N' of the elementary diagram consisting of all $\text{Ult}(V, E_a)$, and recalling that we had elementary embeddings $\text{Ult}(V, E_a) \rightarrow N$, we obtain an elementary diagram

$$\begin{array}{ccc} V & \xrightarrow{j} & M \\ \downarrow & & \uparrow \\ \text{Ult}(V, E_a) & & N \\ \downarrow & & \uparrow \\ \text{Ult}(V, E_b) & \longrightarrow & N' \end{array}$$

whenever $a \subseteq b$. But N was the transitive collapse of $\text{Hull}^M(j''V \cup \{\lambda\})$, so necessarily must be minimal possible. This implies $N' = N$. Therefore we introduce the notation $N = \text{Ult}(V, E)$ where

$$\text{Ult}(V, E) = \left\{ [f, a] : \text{dom } f = \binom{\kappa_a}{\text{card } a}, a \subset \lambda, \text{ card } a < \aleph_0. \right\}.$$

Here $[f, a] = [g, b]$ if there is a $c \supseteq a \cup b$ such that $[f, a]_{E_c} = [g, b]_{E_c}$. In particular, there is a canonical embedding $V \rightarrow \text{Ult}(V, E)$ given by $x \mapsto [c_x, \emptyset]$.

12.2 The abstract definition

We now give an abstract definition of extenders which do not require us to have been given an elementary embedding a priori.

Whenever we have a directed system $X = (X_i : i \in I)$ and an embedding $X_i \rightarrow X_j$, $x \in X_i$, we will write x_i^j to mean the lift of x to X_j .

Definition 12.6. Let κ, λ be cardinals. A (κ, λ) -extender is a λ -sequence

$$E = (E_a : a \in \binom{\lambda}{<\omega}),$$

such that:

1. For every a , let

$$\kappa_a = \min \left\{ \kappa \in \mathbf{Card} : \binom{\kappa}{\text{card } a} \in E_a \right\}.$$

Then E_a is a κ -complete ultrafilter on $\binom{\kappa_a}{\text{card } a}$.

2. There are inclusion maps $E_a \rightarrow E_b$ induced whenever $a \subseteq b$, so that E is a directed system.
3. If there are $a \in \binom{\lambda}{\text{card } a}$, $\text{dom } f = \binom{\kappa_a}{\text{card } a}$, and $i < \text{card } a$, such that for E_a -almost every $x \in \binom{\kappa_a}{\text{card } a}$, $f(x) < x_i$, then there is a $\beta < a_i$ such that if $b = a \cup \{\beta\}$ then for E_b -almost every $x \in \binom{\kappa_b}{\text{card } b}$, $f_a^b(x) = x_a^b$.
4. The direct limit of the elementary diagram $(\text{Ult}(V, E_a) : a \in \binom{\lambda}{\text{card } a})$ is well-founded.

The transitive collapse of the direct limit of $(\text{Ult}(V, E_a) : a \in \binom{\lambda}{\text{card } a})$ is denoted $\text{Ult}(V, E)$, and we let j_E denote the canonical elementary embedding $V \rightarrow \text{Ult}(V, E)$.

From this definition, we end up with a directed system of ultrapowers $\text{Ult}(V, E_a)$. Note that condition (4), and hence the definition of $\text{Ult}(V, E)$ doesn't a priori make sense; we have to check that the conditions before it actually give us an elementary diagram:

Lemma 12.7. Let E be a (κ, λ) -extender. Then the $\text{Ult}(V, E_a)$, as a ranges over $\binom{\lambda}{\text{card } a}$, form an elementary diagram.

Proof. Fix $a \subseteq b$, θ a formula, and $f \in V^{\kappa_a}$; then $\text{Ult}(V, E_a) \models \theta([f])$ iff E_a -almost every $x \in \kappa_a$ satisfies $\theta(f(x))$; but this happens iff E_b -almost every $y \in \kappa_b$ satisfies $\theta(f_a^b(y))$. This happens iff $\text{Ult}(V, E_b) \models \theta([f_a^b])$. \square

Therefore, by the Mostowski collapse lemma, $\text{Ult}(V, E)$ is well-defined.

Obviously conditions (1), (2), and (4) hold for a derived extender. Meanwhile, condition (3) holds for a, f, i exactly if

$$a \in j \left\{ x \in \binom{\kappa_a}{\text{card } a} : f(x) < x_i \right\} = \left\{ x \in \binom{j(\kappa_a)}{\text{card } a} : j(f)(x) < x_i \right\},$$

so let $\beta = j(f)(a)$. Therefore every derived extender is actually an extender. We now check the converse:

Theorem 12.8. Assume that E is a (κ, λ) -extender. Then E is the (κ, λ) -derived extender of j_E .

Proof. We must show that $x \in E_a$ iff $a \in j_E(x)$. Actually, it suffices to show $[\text{id}, a] = a$: if this is true, then $a \in j_E(x)$ iff $[\text{id}, a] \in j_E(x)$ iff

$$x = \left\{ y \in \binom{\kappa_a}{\text{card } a} : y \in c_x(\emptyset) \right\} = E_{a \cup \{\emptyset\}}$$

iff $x \in E_a$.

Lemma 12.9. For every $\alpha < \lambda$,

$$[\bigcup, \{\alpha\}] = \alpha.$$

Proof. By transfinite induction. Assume this is true for $\beta < \alpha$; we must show that every element of $x = [\bigcup, \{\alpha\}]$ is an ordinal $< \alpha$. Assume $[f, \alpha] \in x$ and $a' = a \cup \{\alpha\}$. Then

$$E_{a'} \ni \left\{ x \in \binom{\kappa_{a'}}{\text{card } a'} : f_a^{a'}(x) \in x_a^{a'} \right\}.$$

By (3) in the definition, there is a $\beta < \alpha$ such that if $b = a' \cup \{\beta\}$ then

$$E_b \ni \left\{ x \in \binom{\kappa_b}{\text{card } b} : f_a^b(x) = x_\beta^b \right\}.$$

Setting $b' = a \cup \{\beta\}$, by the definition of a directed system,

$$E_{b'} \ni \left\{ x \in \binom{\kappa_{b'}}{\text{card } b'} : f_a^{b'}(x) = x_\beta^{b'} \right\}.$$

Therefore $[f, a] = [\bigcup, \beta]$, so by induction $[f, a] = \beta$. \square

We now show that $[\text{id}, a] = a$. Suppose $[f, b] \in [\text{id}, a]$. Then

$$E_{a \cup b} \ni \left\{ x \in \binom{\kappa_{a \cup b}}{\text{card } a \cup b} : f_b^{a \cup b}(x) \in x_a^{a \cup b} \right\},$$

and as $x_a^{a \cup b}$ is a finite set, the infinite pigeonhole principle gives an $\alpha \in a$ such that

$$E_{a \cup b} \ni \left\{ x \in \binom{\kappa_{a \cup b}}{\text{card } a \cup b} : f_b^{a \cup b}(x) \in \bigcup x_{\{\alpha\}}^{a \cup b} \right\},$$

so $[f, b] = \alpha$. \square

Summarizing, given an elementary embedding $j : V \rightarrow M$ we can find an extender E such that the elementary diagram

$$\begin{array}{ccc} V & \xrightarrow{j} & M \\ & \searrow j_E & \nearrow k_E \\ & \text{Ult}(V, E) & \end{array}$$

commutes, where $k_E[a, f] = j(f)(a)$ and $j_E(f)(a) = j(f)(a)$ for any function f with $\text{dom } f = \binom{\kappa_a}{\text{card } a}$, and $a \subset \lambda$ finite. On any finite subset of λ , k_E restricts to the identity, and if $j(\gamma) \leq \lambda$ then $j_E = j$ when restricted to γ .

Lemma 12.10. Let M be a transitive, well-founded model of ZFC. Then $\text{Aut}(M)$, the group of isomorphisms $M \rightarrow M$, is trivial.

Proof. Let $j : M \rightarrow M$ be an automorphism of M . We show that $j|V_\alpha^M$ is the identity for every α , by transfinite induction.

Clearly this is true if $\alpha = 0$, and at limit stages. Assume $j|V_\alpha^M = \text{id}$, and let $x \in V_{\alpha+1}^M = \mathcal{P}^M(V_\alpha^M)$. Then $y \in x$ iff $j(y) \in j(x)$ since j is elementary, and j does not add any more elements to x since it is surjective. But $j(y) = y$ since $y \in V_\alpha^M$, so this implies $j(x) = x$. \square

Theorem 12.11. Let $j : V \rightarrow M$ be an elementary embedding and E its derived extender. Let $\gamma \leq \lambda$, λ a limit ordinal such that

$$\lambda \geq \text{card}^M V_\gamma^M.$$

Then $V_\gamma^{\text{Ult}(V,E)} = V_\gamma$ and $k_E|V_\gamma^{\text{Ult}(V,E)}$ is the identity. In particular, $\text{crt } j$ is γ -strong.

Proof. Let $\nu = \text{card}^{\text{Ult}(V,E)} V_\gamma^{\text{Ult}(V,E)}$. Then $k_E(\nu) = \text{card}^M V_\gamma^M$, so

$$\nu \leq k_E(\nu) \leq \lambda.$$

If $\nu < k_E(\nu)$ then $\nu < \lambda$, but then k_E is the identity on $\{\nu\} \subset \lambda$, a contradiction. So ν is a fixed point of E .

Now let $\{X_\xi\}_{\xi < \nu}$ be a surjective ν -sequence of sets in $V_\gamma^{\text{Ult}(V,E)}$, in $\text{Ult}(V,E)$. Then $\{k_E(X_\xi)\}_{\xi < \nu}$ is a surjective ν -sequence of sets in V_γ^M , so k_E is surjective. Since k_E is elementary, it follows that k_E is an isomorphism $V_\gamma^{\text{Ult}(V,E)} \rightarrow V_\gamma^M$.

But the automorphism group of a well-founded, transitive model is trivial, so this implies that k_E is a unique isomorphism, i.e. the identity. \square

Theorem 12.12. Let E be a (κ, λ) -extender, $j_E" \gamma \in \text{Ult}(V, E)$, and $\lambda^\gamma \in \text{Ult}(V, E)$. Then $\text{crt } j_E$ is γ -supercompact, as witnessed by $\text{Ult}(V, E)$.

We omit the proof; it is extremely ugly.

Definition 12.13. Let E be a (κ, λ) -extender. Then the *critical point* $\text{crt } E$ is κ and the *length* of E is λ .

Definition 12.14. Let E be a (κ, λ) -extender. The *support* of E is

$$\text{supp } E = \sup_{a \subset \lambda} \kappa_a,$$

the *strength* of E is the largest γ such that $V_\gamma \in \text{Ult}(V, E)$, and the *closure* of E is the smallest $\text{card } \gamma$ such that $\text{Ult}(V, E)^\gamma$ is not contained in $\text{Ult}(V, E)$.

We now show that extenders allow us to give a first-order definition of “supercompact” and “strong.” This is important because the notion of an elementary embedding of V cannot be made rigorous in ZFC; it is a second-order, or class-theoretic, concept.

Theorem 12.15. Let $\gamma > \kappa$. Then κ is γ -strong iff there is a λ and a (κ, λ) -extender E such that $j_E(\kappa) > \gamma$ and E has strength $\geq \gamma$.

Proof. If such an extender exists then it witnesses strength. Otherwise, assume κ is γ -strong, witnessed by $j : V \rightarrow M$. Let $\lambda > \gamma + \text{card}^M V_\gamma$. Then there is a (κ, λ) -extender which satisfies the constraints. \square

Theorem 12.16. Let $\delta > \kappa$. Then κ is δ -supercompact iff there is a λ and a (κ, λ) -extender E such that $j_E(\kappa) > \delta$, and E has closure $> \delta$.

Proof. One direction is trivial, so assume that κ is δ -supercompact, and $j : V \rightarrow M$ is a witness to the δ -supercompactness of κ .

Let $\lambda > \delta$ be such that $\lambda = \text{card } V_\lambda$. Since $\mu \mapsto \text{card } V_\mu$ is a normal function, its set of fixed points is cofinal, so λ exists. Taking λ to be a regular cardinal, we may assume $\text{cof } \lambda > \delta$. Let E be the $(\kappa, j(\lambda))$ -extender derived from j . Since derived extenders agree,

$$V_{j(\lambda)}^{\text{Ult}(V, E)} = V_{j(\lambda)}^M,$$

and in particular,

$$j_E'' \lambda \in V_{j(\lambda)+1}^M \subset \text{Ult}(V, E).$$

It remains to show that $j(\lambda)^\delta$ is contained in $\text{Ult}(V, E)$. Let f be a sequence in $j(\lambda)^\delta$; since M is a witness to the δ -supercompactness of κ , $f \in M$. But since $\text{cof } \lambda > \delta$,

$$\text{cof}^M j(\lambda) = j(\text{cof } \lambda) > j(\delta) \geq \delta.$$

Therefore f cannot be cofinal, so V_γ^M for some $\gamma < j(\lambda)$. Moreover, $V_\gamma^M = V_\gamma^{\text{Ult}(V, E)}$, so $f \in \text{Ult}(V, E)$. \square

Expanding out the definitions of the above properties of an extender, we see that the above theorem implies that γ -strength, or γ -supercompactness, is a Σ_2 definition. In particular, strength and supercompactness are Π_3 definitions.

Corollary 12.17. Suppose that κ is $(\kappa + 2)$ -strong. Then there is a nontrivial uniform ultrafilter U on κ such that U -almost every element of κ is measurable.

Proof. Let $j : V \rightarrow \text{Ult}(V, E)$ witness that κ is $(\kappa + 2)$ -strong, and let

$$U = \{x \subseteq \kappa : \kappa \in j(x)\}$$

be the derived ultrafilter of j . Let x be the set of all measurable cardinals $< \kappa$; since κ is measurable and $\kappa < j(\kappa)$, $\kappa \in j(x)$. Therefore $x \in U$. \square

The conga line never ends! Strong cardinals humiliate measurable cardinals just as measurable cardinals humiliate Mahlo cardinals, and if there is just one strong cardinal, then $L[U]$ looks nothing at all like V .

12.3 Reinhardt cardinals

Recall that we introduced the notions of strong and supercompact cardinals to assert the existence of elementary embeddings $j : V \rightarrow M$ such that M was “not too much bigger” than V . Why don’t we just cut to the chase and assume that in fact $V = M$?

Definition 12.18. A *Reinhardt cardinal* is the critical point of an elementary embedding $V \rightarrow V$.

Definition 12.19. A *Berkeley cardinal* κ is a regular cardinal such that for every transitive set $M \ni \kappa$, there is an elementary embedding $j : M \rightarrow M$ such that $\text{crt } j = \kappa$.

Unfortunately, the above notions are inconsistent with ZFC. Actually, Woodin introduced the notion of a Berkeley cardinal as a joke, something that was obviously inconsistent with not just ZFC, but even ZF. But decades later, it is still an open problem to show that Berkeley cardinals are inconsistent with ZF.

Theorem 12.20. There is no Reinhardt cardinal.

Proof. We will use Theorem 9.27, a form of the axiom of choice due to Solovay.

Let $j : V \rightarrow V$ be an elementary embedding and $\kappa_0 = \text{crt } j$. Let $\kappa_{n+1} = j(\kappa_n)$, and let

$$\lambda = \sup_{n < \omega} \kappa_n.$$

Therefore

$$j(\lambda) = j\left(\sup_n \kappa_n\right) = \sup_n j(\kappa_n) = \lambda.$$

So $j(\lambda^+) = \lambda^+$.

Let $S = \{\alpha < \lambda^+ : \text{cof } \alpha = \omega\}$. If c is a club in λ^+ , then $c \cap \mathbf{Card}$ is also a club since $\mathbf{Card} \cap \lambda^+$ is a club, so assume that c only consists of cardinals. Then let $(\delta_n : n < \omega)$ be an increasing ω -sequence in c . By construction $\delta = \sum_n \delta_n$ satisfies $\delta > \delta_n$ for any n , so $\text{cof } \delta = \omega$. Therefore $\delta \in S \cap c$. So S is stationary. Moreover,

$$j(S) = \{\alpha < j(\lambda^+) : \text{cof } \alpha = j(\omega)\} = S$$

since λ^+, ω are fixed points of j .

By the Solovay partitioning theorem, there is a partition $S = \bigcup_{\alpha} S_{\alpha}$ of S into stationary sets of cardinality λ^+ . Then let $(T_{\alpha} : \alpha < \omega) = j(S_{\alpha} : \alpha < \omega)$, so $j(S_{\alpha}) = T_{j(\alpha)}$. Therefore $j(S_{\lambda}) = T_{\lambda}$. Since j is elementary, the T_{α} form a partition of S into stationary sets of cardinality λ^+ .

Let F be the set of fixed points of j in λ^+ . Then F is closed under ω -limits, and contains the κ_n , so is cofinal. Now T_{κ_0} is stationary, so in particular it meets the closure \overline{F} of F , say at η , and by definition of T_{κ_0} , $\text{cof } \eta = \omega$. Therefore η is an ω -limit of F , so $\eta \in F$. But then $\eta = j(\eta) \in j(S_{\gamma}) = T_{j(\gamma)}$ for some γ . Therefore $\kappa_0 = j(\gamma)$, so $\gamma < \kappa_0$ is not a fixed point of j , but κ_0 is the least ordinal moved by j . \square

12.4 Supercompact cardinals

Recall that a cardinal κ is supercompact if for every $\delta > \kappa$ there is a λ and a (κ, λ) -extender E such that E has closure $> \delta$ and $j_E(\kappa) > \delta$. We want to show that this is equivalent to some definitions in terms of ultrafilters that will be easier to work with in practice.

Definition 12.21. Let $\kappa \leq \lambda$ be cardinals. Then let

$$\mathcal{P}_{\kappa}(\lambda) = \{x \subseteq \lambda : \text{card } x < \kappa\}.$$

Definition 12.22. Let U be an ultrafilter on $\mathcal{P}_\kappa(\lambda)$. We say that U is a *fine ultrafilter* if for every $\alpha < \lambda$ and U -almost every $x \in \mathcal{P}_\kappa(\lambda)$, $\alpha \in x$. We say that U is a *normal ultrafilter* if for every $f : \mathcal{P}_\kappa(\lambda) \rightarrow \lambda$ such that for U -almost $x \in \mathcal{P}_\kappa(\lambda)$, $f(x) \in x$, then f is constant U -almost everywhere.

So a fine ultrafilter does not privilege any element of U . Normal ultrafilters are similar to in the case of normal ultrafilters on measurable cardinals, which required that if a function pushed down on a large set, then it was actually constant on a large set.

Lemma 12.23. Let $\kappa \leq \lambda$ be cardinals. Then κ is λ -supercompact iff there is a κ -complete normal fine ultrafilter on $\mathcal{P}_\kappa(\lambda)$.

Proof. First assume κ is λ -supercompact, and let $j : V \rightarrow M$ witness this, so that $\text{crt } j = \kappa$, $j(\kappa) > \lambda$, and $M^\lambda \subseteq M$. Let

$$U = \{x \in \mathcal{P}_\kappa(\lambda) : j''\lambda \in j(x)\}$$

be the derived ultrafilter. Then U is a κ -complete ultrafilter on $\mathcal{P}_\kappa(\lambda)$, as in the proof of the Scott-Keisler theorem.

Since $M^\lambda \subseteq M$ and $j''\lambda = \{j(\alpha) : \alpha < \lambda\}$ is a λ -sequence, $j''\lambda \in M$. Moreover,

$$\text{card } j''\lambda = \lambda < j(\kappa).$$

If $\alpha < \kappa$ then $j(\alpha) \in j(\kappa) \subseteq j(\lambda)$ so $j''\kappa \subseteq j(\lambda)$. So j sends $\mathcal{P}_\kappa(\lambda)$ into $\mathcal{P}_{j(\kappa)}(j(\lambda))$.

To see that U is fine, let $\alpha < \lambda$, so that U -almost every $x \in \mathcal{P}_\kappa(\lambda)$ satisfies $\alpha \in x$ iff

$$j''\lambda \in j(\{x \in \mathcal{P}_\kappa(\lambda) : \alpha \in x\}) = \{x \in \mathcal{P}_{j(\kappa)}(j(\lambda)) : j(\alpha) \in x\},$$

i.e. $j(\alpha) \in j''\lambda$, which is clear.

To see that U is normal, let $f : \mathcal{P}_\kappa(\lambda) \rightarrow \lambda$ satisfy for U -almost every x , $f(x) \in x$. Unraveling the definitions, this means that $j(f)(j''\lambda) \in j''\lambda$. Thus there is an $\alpha < \lambda$ such that $j(f)(j''\lambda) = j(\alpha)$, and so if $x \in U$, $j''\lambda \in j(x)$ implies $f(x) = \alpha$.

For the converse, let U be a κ -complete normal fine ultrafilter on $\mathcal{P}_\kappa(\lambda)$ and $j : V \rightarrow \text{Ult}(V, U)$ the ultrapower map.

Lemma 12.24. $j''\lambda = [\text{id}]$.

Proof. We first prove $j''\lambda \subseteq [\text{id}]$. If $[c_\alpha] = j(\alpha) \in j''\lambda$, then we must show $[c_\alpha] = [\text{id}]$, which is true exactly when $\alpha \in x$ for U -almost every x , which happens because U is finite.

Conversely, suppose $f \in [\text{id}]$, i.e. for U -almost every x , $f(x) \in x$. By normality this means that for some $\alpha < \lambda$ and U -almost every x , $f(x) = \alpha$. Therefore $[f] \in j(\alpha) \in j''\lambda$. \square

To see that $j(\kappa) > \lambda$, note that the map h that sends a well-ordered set to its ordertype, and a set without a designated well-ordering to card , is definable, so preserved by j . Therefore

$$\lambda = h(j''\lambda) = h([\text{id}]).$$

But for every $x \in \mathcal{P}_\kappa(\lambda)$, $h(\text{id}(x)) = h(x) < \kappa = c_\kappa(x)$ so $h([\text{id}]) < h([c_\kappa]) = j(\kappa)$. Thus $\lambda < j(\kappa)$. In particular, $\kappa < \lambda < j(\kappa)$ so $\text{crt } j \leq \kappa$, but U was κ -complete so $\text{crt } j \geq \kappa$.

Finally we must show $\text{Ult}(V, U)^\lambda \subseteq \text{Ult}(V, U)$. Let $(f_\alpha : \alpha < \lambda)$ be given. For any $g \in \mathcal{P}_\kappa(\lambda)$ let

$$g(x) = (f_\alpha(x) : \alpha < \lambda).$$

Then if $\alpha < \lambda$ and $x \in \mathcal{P}_\kappa(\lambda)$, $c_\alpha(x) = \alpha$, so $g(x)(c_\alpha(x)) = f_\alpha(x)$. This means that $[g](j(\alpha)) = f_\alpha$. Since $[g] \in \text{Ult}(V, U)$ and $j''\lambda = [\text{id}] \in \text{Ult}(V, U)$,

$$(f_\alpha : \alpha < \lambda) = ([g](j(\alpha)) : j(\alpha) \in j''\lambda) \in \text{Ult}(V, U),$$

as desired. \square

Summing up the above, we obtain the following corollary.

Corollary 12.25. Let κ be a cardinal. Then the following are equivalent:

1. κ is supercompact.
2. For every $\lambda \geq \kappa$ there is a normal fine ultrafilter on $\mathcal{P}_\kappa(\lambda)$.
3. For every $\lambda \geq \kappa$ there is an extender E such that $\text{crt } E = \kappa$, $j_E(\kappa) > \lambda$, and E has closure $\geq \lambda$.

Theorem 12.26 (Magidor). Let κ be a cardinal. Then κ is supercompact if and only if for every $\lambda \geq \kappa$ there is a $\delta < \lambda$ and an elementary embedding $j : V_\delta \rightarrow V_\lambda$ such that $j \text{ crt } j = \kappa$.

Proof. First suppose κ is supercompact. Let $\lambda \geq \kappa$ and $k : V \rightarrow M$ be a witness that κ is $(\text{card } V_\lambda)$ -supercompact, so $M^{\text{card } V_\lambda} \subseteq M$. Therefore $k|V_\lambda \in M$, and since $\lambda < k(\kappa)$,

$$M \models \text{There are } \eta < k(\kappa) \text{ and } i : V_\eta \rightarrow V_{j(\lambda)} \text{ such that } i(\text{crt } i) = k(\kappa).$$

Here the witnesses are $\eta = \lambda$ and $i = k|V_\lambda$. Since k is an elementary embedding, there are $\delta < \kappa$ and $j : V_\delta \rightarrow V_\lambda$ such that $j \text{ crt } j = \kappa$.

For the converse, let $\lambda \geq \kappa$; we need a κ -complete normal fine ultrafilter on $\mathcal{P}_\kappa(\lambda)$. By assumption, there is an elementary embedding $j : V_\delta \rightarrow V_\lambda$ for some $\delta < \lambda$, and j extends to an elementary embedding $j : V_{\delta+10} \rightarrow V_{\lambda+10}$, where $j \text{ crt } j = \kappa$. Let $\eta = \text{crt } j$. By assumption, $\mathcal{P}(\mathcal{P}_\eta(\delta)) \in V_{\delta+10}$, and

$$U' = \{x \subseteq \mathcal{P}_\eta(\delta) : j''\delta \in j(x)\}$$

is a η -complete normal fine ultrafilter on $\mathcal{P}_\eta(\delta)$. Here the $+10$ guarantees that we have enough iterates of the power set to carry out the argument above. Then $U = j(U')$ is a $j(\eta) = \kappa$ -complete normal fine ultrafilter on $\mathcal{P}_{j(\eta)}(j(\delta)) = \mathcal{P}_\kappa(\lambda)$. \square

By a similar argument we have:

Corollary 12.27. Let κ be a cardinal. Then κ is supercompact if and only if for every $\lambda \geq \kappa$ there is an extender E such that $\text{supp } E \subseteq \kappa$, the strength of E is $\geq \lambda$, and $j_E(\text{crt } E) = \kappa$.

Theorem 12.28 (Solovay). Suppose $\lambda > \kappa$ are regular cardinals. Then there is an $X \subseteq \mathcal{P}_\kappa(\lambda)$ such that $\sup |X : X \rightarrow \lambda|$ is injective, and for every κ -complete normal fine ultrafilter U on $\mathcal{P}_\kappa(\lambda)$, $X \in U$.

Proof. Let

$$S = \{\alpha < \lambda : \text{cof } \alpha = \omega\}.$$

If c is a club in λ , then so is $c \cap \text{Card}$, so let $(\delta_n : n < \omega)$ be an increasing sequence of cardinals in $c \cap \text{Card}$, which is possible since $c \cap \text{Card}$ is cofinal. Then their limit $\delta = \sup_{n < \omega} \delta_n$ is also in $c \cap \text{Card}$ since it is closed, so $\delta \in c \cap S$. Therefore S is stationary, so by Solovay's stationary theorem, there is a partition of S into stationary sets $(S_\alpha : \alpha < \lambda)$.

If $\beta < \lambda$ and $\omega < \text{cof } \beta < \kappa$, let

$$\sigma_\beta = \{\alpha < \beta : S_\alpha \cap \beta \text{ is stationary in } \beta\}.$$

Lemma 12.29. For every β such that σ_β exists, $\sigma_\beta \in \mathcal{P}_\kappa(\lambda)$.

Proof. By assumption, we have a collection $(S_\alpha \cap \beta : \alpha \in \sigma_\beta)$ of disjoint stationary sets in β .

Suppose that $\text{card } \sigma_\beta \geq \kappa$. Let c be a club in β whose ordertype is $\text{cof } \beta$, which exists since $\text{cof } \beta > \omega$. Then the sets $S_\alpha \cap \beta \cap c$ are all disjoint, and there are κ sets. Thus they induce a partition of (a subset of) $\text{cof } \beta$ into κ many nonempty sets, but $\text{cof } \beta < \kappa$ so this is a contradiction. \square

Now let

$$X = \{\sigma_\beta \in \mathcal{P}_\kappa(\lambda) : \sup \sigma_\beta = \beta\}.$$

By construction, $\sup |X|$ is injective.

Now fix a normal fine κ -complete ultrafilter U on $\mathcal{P}_\kappa(\lambda)$. We must show $X \in U$. Let $j : V \rightarrow \text{Ult}(V, U)$ be the ultrapower embedding, and let

$$(T_\alpha : \alpha < j(\lambda)) = j((S_\alpha : \alpha < \lambda)).$$

Let $\lambda^* = \sup j'' \lambda$.

Let

$$\sigma'_\beta = \{\alpha < \beta : T_\alpha \cap \beta \text{ is stationary in } \beta\}.$$

Then

$$j(X) = \{\sigma'_\beta \in \mathcal{P}_{j(\kappa)}(j(\lambda)) : \text{Ult}(V, U) \models \sup \sigma'_\beta = \beta\}.$$

Lemma 12.30. In $\text{Ult}(V, U)$, $j'' \lambda = \sigma'_{\lambda^*}$.

Proof. First we prove that if $\alpha < \lambda$ then $j(\alpha) \in \sigma'_{\lambda^*}$. In other words, we must show that $T_{j(\alpha)} \cap \lambda^* = j(\alpha) \cap \lambda^*$ is stationary in λ^* .

Let $c \in \text{Ult}(V, U)$ be a club in λ^* , and let

$$D = \{\alpha < \lambda : j(\alpha) \in c\}.$$

We claim that D is an ω -club (as a subset of λ in V) in the sense that it is cofinal and closed under ω -limits. To see this, note that $j'' \lambda$ is cofinal in λ , and any ω -limits in λ are sent to ω -limits in λ^* , so $j'' \lambda$ is an ω -club. But $D = j'' \lambda \cap c$, so D is also an ω -club. Therefore there is a $\xi \in D \cap S_\alpha$, but then

$$j(\xi) \in c \cap S_{j(\alpha)} = T_{j(\alpha)},$$

so $j(\alpha) \in \sigma'_{\lambda^*}$.

Conversely, suppose $\alpha \in \sigma'_{\lambda^*}$, so $\text{Ult}(V, U)$ thinks $T_\alpha \cap \lambda$ is stationary in λ^* . We will show $\alpha = j(\alpha')$ for some $\alpha' < \lambda$, so $\alpha \in j''\lambda$.

By the same argument as above, $j''\lambda$ is an ω -club in $\text{Ult}(V, U)$, so there is a $\xi < \lambda$ such that

$$j(\xi) \in T_\alpha \cap j''\lambda.$$

But $(j''(S_{\alpha'} \cap \lambda^*) : \alpha' < \lambda)$ is a partition of $j''\lambda$. So we may choose an α' such that

$$j(\xi) \in j''(S_{\alpha'} \cap \lambda^*).$$

Then $\alpha = j(\alpha')$. □

Therefore $\text{Ult}(V, U) \models \sup \sigma'_{\lambda^*} = \lambda^*$, i.e. $j''\lambda \in j(X)$. Since

$$U = \{Y \subseteq \mathcal{P}_\kappa(\lambda) : j''\lambda \in j(Y)\}$$

this implies that $X \in U$. □

At first it seems surprising that every normal fine ultrafilter would contain a given set, but consider that on ω , for every cofinite set X and every nontrivial ultrafilter U , $X \in U$.

Chapter 13

The Ultimate L

Recall that if κ is a measurable cardinal witnessed by an ultrafilter U , then there is a canonical model $L[U]$ where κ is the unique measurable cardinal. Then $\kappa \in L[U]$, so $\kappa \cap L[U] = \kappa \in U$. We also proved that $U \cap L[U] \in L[U]$. So $L[U]$ inherits the witness U to the measurability of κ , namely $U \cap L[U]$. We want the same thing for supercompact cardinals, so we can get a canonical model of one supercompact cardinal. When we do this, it will turn out that, pending some conjectures of Woodin, such a model would not just be a canonical model for one supercompact cardinal, but every large cardinal whatsoever,

Definition 13.1. Let T be a theory in the language (\in) . An *inner model* of T is a transitive class M such that (M, \in) is a model of T and $\mathbf{Ord} \subset M$.

We can also talk about inner models of languages which extend (\in) . For example, V and L are inner models of ZFC, V being the maximal such model and L being the minimal such model. Meanwhile $L[U]$ is an inner model of the theory ZFC + “There is a measurable cardinal”, and is also an inner model of the theory ZFC + “ U is a normal fine ultrafilter” in the language (\in, U) . Notice that since an inner model gets the relation \in correct, if it is a model of foundation then it is actually well-founded.

Definition 13.2. Let δ be a supercompact cardinal. A *weak extender model for the supercompactness* of δ is an inner model N of ZFC such that for every $\lambda \geq \delta$ there is a δ -complete normal fine ultrafilter U on $\mathcal{P}_\delta(\lambda)$ such that:

1. $\mathcal{P}_\delta(\lambda) \cap N \in U$.
2. $U \cap N \in N$.

We don’t know that $\mathcal{P}_\delta(\lambda) \cap N = \mathcal{P}_\delta(\lambda)$ a priori, so $\mathcal{P}_\delta(\lambda) \cap N \in U$ is quite a strong assumption to make on U .

Definition 13.3 (Hamkins). Let N be an inner model and κ a cardinal. We say N has the *κ -covering property* if for every $\tau \subset N$ such that $\text{card } \tau < \kappa$, there is a set $\tau' \in N$ such that $\text{card } \tau' < \kappa$ and $\tau \subseteq \tau'$.

If an inner model N has κ -covering, then N is “fat below κ ”: N contains “most” sets under κ . For example, if τ is a definable subset of N , N is an inner model of ZF, and N has

$(\text{card } \tau)^+$ -covering, then there is a $\tau' \in N$ such that $\text{card } \tau = \text{card } \tau'$ and $\tau \subseteq \tau'$. But since N has comprehension and τ was assumed definable, this implies that actually $\tau \in N$.

We now show that weak extender models for the supercompactness of δ have the δ -covering property. Since supercompact cardinals are really big, this in particular implies that weak extender models for supercompactness are really fat up to a possibly irrelevant tail.

Lemma 13.4. Suppose N is an inner model of ZFC such that for all $\lambda \geq \delta$, there is a δ -complete normal fine ultrafilter U on $\mathcal{P}_\delta(\lambda)$ such that $\mathcal{P}_\delta(\lambda) \cap N \in U$. Then N has the δ -covering property.

Proof. We first check this when $\tau \subset \text{Ord}$. Let $\lambda > \sup \tau$; then $\tau \subseteq \lambda$. Since U is fine, if $\alpha < \lambda$,

$$A_\alpha = \{\sigma \in \mathcal{P}_\delta(\lambda) : \alpha \in \sigma\} \in U.$$

Let $A = \bigcap_{\alpha \in \tau} A_\alpha$; since $\text{card } \tau < \delta$, $A \in U$. Since $\mathcal{P}_\delta(\lambda) \cap N \in U$ and U is an ultrafilter, $A \cap \mathcal{P}_\delta(\lambda) \cap N$ is nonempty. So let $\tau' \in A \cap \mathcal{P}_\delta(\lambda) \cap N$. Then $\text{card } \tau' < \delta$ and $A_\alpha \subseteq \tau'$ for any $\alpha < \tau$, hence $\tau \subseteq \tau'$.

Now if τ is arbitrary, choose a well-ordering of τ in N and precede as above. \square

13.1 Averting the dichotomy

We now give another way, besides the δ -covering property, in which weak extender models are much closer to V than L .

Definition 13.5. Let N be an inner model of ZFC and δ a regular cardinal. We say that N is *close to V* (in the sense of singular cardinals) above δ if for every singular cardinal $\gamma > \delta$:

1. $N \models \gamma$ is a singular cardinal.
2. $(\gamma^+)^N = \gamma^+$.

If N is close to V above \aleph_0 , we simply say that N is *close to V* .

For some motivation for the definition, consider the following theorem of Jensen.

Definition 13.6. If A is a set, define its *sharp* $A^\#$ to be the definable real number which codes the set of true sentences about $L[A]$ in the language $(\in, A, \aleph_n : n < \omega)$.

So $0^\#$ cannot be definable, and thus its definition makes no sense, if $V = L$, by Tarski's theorem on the undefinability of truth. Thus the existence of $0^\#$ is a “switch” that tells us how well L approximates V . Similarly $U^\#$ is a “switch” that tells us how well $L[U]$ approximates V . These switches are very dangerous to have, because of the following theorem.

Theorem 13.7 (Jensen's covering theorem). Exactly one of the following is true:

1. $0^\#$ does not exist, and L is close to V .

2. $0^\#$ exists, and for every uncountable cardinal κ ,

$$L \models \kappa \text{ is an inaccessible cardinal.}$$

Moreover, if there is a measurable cardinal, then $0^\#$ exists.

Intuitively, if there is a measurable cardinal, then L is so pathetically small that we can define truth in it, and thinks that even \aleph_1 is inaccessible. This is a bad property of L : by whimsical identity, $0^\#$ exists, so any axiom which says that V looks like L in some way needs to be thrown out. The proof of Jensen's covering theorem is notoriously difficult, and we omit it.

Notice that if N is close to V , then the calamitous behavior that N thinks that every uncountable cardinal is a large cardinal is impossible: it would then be the case that such cardinals are regular in N , and hence are actually regular cardinals, but the class of singular cardinals is proper.

A dichotomy similar to L is not possible for weak extender properties, however.

Theorem 13.8. Let N be a weak extender model for the supercompactness of δ . Then N is close to V above δ .

To prove this theorem, we will need the following lemma.

Lemma 13.9. Let N be a weak extender model for the supercompactness of δ and $\lambda \geq \delta$ is a regular cardinal in N . Then $\text{cof } \lambda = \text{card } \lambda$ in V .

Proof. Since N is a weak extender model for the supercompactness of δ and satisfies Solovay's supercompactness theorem, there is a δ -complete normal fine ultrafilter U and a set $X \in U \cap N$ such that $\mathcal{P}_\delta(\lambda) \cap N \in U$, $U \cap N \in U$, and $\sup |X|$ is injective.

Now let $c \subseteq \lambda$ be a club of ordertype $\text{cof } \lambda$. Let $X_c = \{\sigma \in X : \sup \sigma \in c\}$. Then $j_U(c)$ is a club and $j_U''c \subseteq j_U(c)$, so $j_U''\lambda \in j_U(X)$ and $\sup \sigma \in j_U(c)$, so $j_U''\lambda \in j_U(X_c)$. Therefore $X_c \in U$. Since U is fine, $\bigcup X_c = \lambda$. On the other hand, the δ -covering property applied to each $\sigma \in \mathcal{P}_\delta(\lambda)$ implies that $\delta \leq \text{cof } \lambda$, so

$$\text{card } X_c \leq \delta \text{ card } c = \delta \text{ cof } \lambda = \text{cof } \lambda.$$

Therefore

$$\text{card } \lambda \leq \delta \text{ card } X_c = \text{cof } \lambda.$$

Since $\text{card } \lambda \geq \text{cof } \lambda$, $\text{card } \lambda = \text{cof } \lambda$. □

Proof of Theorem 13.8. Suppose γ is regular in N . Then

$$\gamma < \text{cof } \gamma = \text{card } \gamma = \gamma,$$

a contradiction.

Now suppose $(\gamma^+)^N < \gamma^+$. Then $(\gamma^+)^N$ is regular in N since it is a successor in N . Therefore

$$\text{cof}(\gamma^+)^N = \text{card}(\gamma^+)^N = \gamma.$$

But $\text{cof}(\gamma^+)^N < \gamma$ since γ is singular. □

13.2 Magidor's weak extender model formulation

Theorem 13.10. Let N be an inner model of ZFC and δ a regular cardinal. Then N is a weak extender model for the supercompactness of δ iff for every $\lambda > \delta$ and $a \in V_\lambda$ there are $\delta' < \lambda' < \delta$, $a' \in V_{\lambda'}$, and

$$j : V_{\lambda'+1} \rightarrow V_{\lambda+1}$$

such that $\text{crt } j = \delta'$ where $j(\delta') = \delta$, $j(a') = a$, $j(N \cap V_{\lambda'}) = N \cap V_\lambda$, and

$$j|(N \cap V_{\lambda'+1}) \in N.$$

In other words, weak extender models N for supercompactness can be witnessed by choosing an elementary embedding $j : V_{\lambda'+1} \rightarrow V_{\lambda+1}$, where $\text{crt } j < \lambda' < \delta < \lambda$, which restricts to a member of N and preserves N and a given set a . Our goal with this theorem is to show that weak extender models actually contain most elementary embeddings, and so contain arbitrarily large cardinals. This makes them drastically different than other inner models such as $L[U]$.

To prove the above theorem we will need the following lemma.

Lemma 13.11. Let N be an inner model of ZFC, δ a regular cardinal,

$$\lambda = \text{card}^N(N \cap V_\lambda)$$

a regular cardinal such that $\lambda > \delta$, and U a witness that N is a weak extender model for the λ -supercompactness of δ , i.e. $\mathcal{P}_\delta(\lambda) \cap N \in U$ and $U \cap N \in N$.

Suppose that $j : V \rightarrow \text{Ult}(V, U)$ is the ultrapower map. Then

$$j(N \cap V_\delta) \cap V_\lambda = N \cap V_\lambda.$$

This lemma says that the ultrapower actually preserves $N \cap V_\delta$, which is really huge.

Proof. We claim

$$N \cap V_\lambda \subseteq j(N \cap V_\delta) \cap V_\lambda.$$

Now

$$U = \{X \subseteq \mathcal{P}_\delta(\lambda) : j''\lambda \in j(X)\},$$

so since $\mathcal{P}_\delta(\lambda) \cap N \in U$,

$$j''\lambda \in j(\mathcal{P}_\delta(\lambda) \cap N) \subseteq j(N).$$

Since $\lambda = \text{card}^N(N \cap V_\lambda)$, there is a bijection $e : \lambda \rightarrow N \cap V_\lambda$ such that $e \in N$. Let $x \in N \cap V_\lambda$, and let $\alpha < \lambda$ be such that $e(\alpha) = x$. Then

$$j(x) = j(e(\alpha)) = j(e)(j(\alpha)).$$

So

$$j''(N \cap V_\lambda) = j(e)''(j''\lambda) \in j(N).$$

Let $T : M \rightarrow L$ be the transitive collapse map. But $N \cap V_\lambda$ is isomorphic to $j''(N \cap V_\lambda)$, yet is a transitive class; thus $N \cap V_\lambda = T(j''(N \cap V_\lambda))$ and in particular,

$$N \cap V_\lambda = T(j(e)''(j''\lambda)).$$

So $N \cap V_\lambda \subseteq j(N)$. But $j(\delta) > \lambda$, so $N \cap V_\lambda \subseteq j(V_\delta)$ and hence

$$N \cap V_\lambda \subseteq j(N \cap V_\delta) \cap V_\lambda.$$

For the converse, note that since $N \cap V_\lambda = T(j(e)''(j''\lambda))$, it suffices to show that

$$j(N \cap V_\delta) \cap V_\lambda \subseteq T(j(e)''(j''\lambda)).$$

Let X be the set of all $\sigma \in \mathcal{P}_\delta(\lambda)$ such that:

1. $e''\sigma$ is an elementary substructure of $N \cap V_\lambda$, and
2. $T(e''\sigma) = N \cap V_\gamma$ where γ is the ordertype of σ .

Then $j(X)$ is the set of all $\sigma \in \mathcal{P}_{j(\delta)}(j(\lambda))$ such that:

1. $j(e)''\sigma$ is an elementary substructure of $j(N) \cap V_\lambda$, and
2. $T(j(e)''\sigma) = N \cap V_\gamma$ where γ is the ordertype of σ .

If $X \in U$, then $j''\lambda \in j(X)$, and hence

$$T(j(e)''(j''\lambda)) = j(N) \cap V_\lambda = j(N \cap V_\delta) \cap V_\lambda,$$

which proves the claim. So we must show $X \in U$.

Let $X' = X \cap N$; we claim $X' \in U$. Since $U' = U \cap N$ is a κ -complete normal fine ultrafilter on $\mathcal{P}_\delta(\lambda) \cap N$, if $j' : N \rightarrow \text{Ult}(N, U')$ denotes its ultrapower map, then:

1. $j'(e)''((j')''\lambda)$ is an elementary substructure of $j'(N) \cap V_{j'(\lambda)}$.
2. $T(j'(e)''((j')''\lambda)) = N \cap V_\lambda$.
3. $\text{Ult}(N, U') \cap V_\lambda = N \cap V_\lambda$.

Therefore $(j')''\lambda \in j'(X')$. So $X' \in U'$, hence $X' \in U$, hence $X \in U$. \square

Proof of Theorem 13.10. Suppose the second. Let $\gamma > \delta$ be such that $\gamma = \text{card } V_\gamma$. By hypothesis applied to $\lambda = \gamma + \omega$ (ordinal addition here) to get $\delta' < \lambda' < \delta$ and $j : V_{\lambda'+1} \rightarrow V_{\lambda+1}$ with the desired properties. Here a can be whatever.

Since γ is definable in $V_{\lambda+1}$ (since $\lambda = \gamma + \omega$), let γ' have the same definition in $V_{\lambda'+1}$, so $j(\gamma') = \gamma$. So

$$\delta' < \gamma' < \lambda' < \delta < \lambda < \gamma$$

and $j''\gamma' \in V_{\lambda+1}$.

Let

$$U' = \{X \subseteq \mathcal{P}_{\delta'}(\lambda') : j''\gamma' \in j(X)\}$$

and $U = j(U')$. Then U' is a δ' -complete normal fine ultrafilter on $\mathcal{P}_{\delta'}(\lambda')$, so U is a δ -complete normal fine ultrafilter on $\mathcal{P}_\delta(\lambda)$. In particular δ is λ -supercompact.

We must show that U is a witness to the sentence “ N is a weak extender model for the λ -supercompactness of δ .”

Lemma 13.12. $\mathcal{P}_{\delta'}(\gamma') \cap N \in U'$.

Proof. By hypothesis $j(N \cap V_{\lambda'}) = N \cap V_{\lambda}$. So if $a' \in V_{\lambda'}$,

$$j(a' \cap N) = j(a') \cap j(N) = j(a') \cap N.$$

Taking $a' = \mathcal{P}_{\delta'(\gamma')}$ $\mathcal{P}_{\delta'}(\gamma') \in V_{\lambda'}$, so $j(\mathcal{P}_{\delta'}(\gamma') \cap N) = \mathcal{P}_{\delta}(\gamma) \cap N$ and hence $j''\gamma' \in N$ implies

$$j''\gamma' \in \mathcal{P}_{\delta}(\gamma) \cap N = j(\mathcal{P}_{\delta'}(\gamma') \cap N),$$

so the claim follows by definition of U' . \square

Taking j of everything in the previous lemma we get $\mathcal{P}_{\delta}(\gamma) \cap N \in U$, which was desired.

Moreover, by hypothesis, $j|(N \cap V_{\lambda'+1}) \in N$. In particular $j(U' \cap N) \in N$. But $j(U' \cap N) = U \cap N$. So U witnesses that N is a weak extender model for the λ -supercompactness of δ .

Now suppose that N is a weak extender model for the supercompactness of δ and $\lambda > \delta$ is a regular cardinal. Let $\gamma > \lambda$ and suppose $\gamma = \text{card } V_{\gamma}$. By the lemma applied to an appropriate regular cardinal $\lambda > \gamma$, there is an ultrafilter U and an elementary embedding $j : V \rightarrow \text{Ult}(V, U)$ such that $j(\delta) > \lambda$, $\text{Ult}(V, U)$ is closed under $\text{card } V_{\gamma+1}$ -sequences, $j(N \cap V_{\delta}) \cap V_{\lambda} = N \cap V_{\lambda}$, and $j''\lambda \in j(N)$. By assumption, $V_{\lambda+1} = V_{\lambda+1}^{\text{Ult}(V, U)}$. So the elementary embedding j sends $V_{\lambda+1} = V_{\lambda+1}^{\text{Ult}(V, U)}$ to $V_{j(\lambda)+1}^{\text{Ult}(V, U)}$, hence given $a \in V_{\lambda}$, $\text{Ult}(V, U)$ thinks there are $\delta' < j(\delta)$, $\lambda' < j(\delta)$, a' and an elementary embedding $k : V_{\lambda'+1} \rightarrow V_{j(\lambda)+1}$, $\text{crt } k = \delta'$, $k(\delta') = j(\delta)$, $k(a') = j(a)$, $k(j(N) \cap V_{\lambda'+1}) \in j(N)$. The witness here is $k = j|V_{\lambda+1}$. Pulling back along $j : V \rightarrow \text{Ult}(V, U)$, we see the claim. \square

13.3 Woodin's universality theorem

We now show that weak extender models contain very large elementary embeddings.

Theorem 13.13. Suppose N is a weak extender model for the supercompactness of δ . Let $\gamma \in \text{Card}^N$ and let

$$j : \text{Hull}((\gamma^+)^N) \rightarrow M$$

be an elementary embedding such that $\text{crt } j \geq \delta$ and $M \subseteq N$. Then $j \in N$.

Proof. Let $\gamma > \lambda$ be given, where $j \in V_{\lambda}$ and $\lambda = \text{card } V_{\lambda}$. By Theorem 13.10, there are $\delta', \gamma', \lambda', j' \in V_{\lambda'}$, and an elementary embedding

$$\pi : V_{\lambda'+1} \rightarrow V_{\lambda+1}$$

such that $\text{crt } \pi = \delta'$, $\pi(\delta') = \delta$, $\pi(j') = j$, $\pi|N \in N$, and

$$\pi(N \cap V_{\lambda'}) = N \cap V_{\lambda}.$$

In particular, j' is an elementary embedding $\text{Hull}((\gamma')^+)^N \rightarrow M'$ where $M' \subseteq N$ is some class. If $j' \in N$, then since $\pi|N \in N$, $j = \pi(j') \in N$ as well. So it suffices to show $j' \in N$.

We are given $\pi|(N \cap V_{\gamma'}) \in N$. Thus

$$\pi| \text{Hull}((\gamma')^+)^N \in \text{Hull}((\gamma^+)^N) \subseteq N.$$

Let $a' \in \text{Hull}(((\gamma')^+)^N)$ and $j'(a') = b' \in M'$. Since $a' \in V_{\lambda'+1}$,

$$\pi(j'(a')) = \pi(b') = \pi(j')(\pi(a')) = j(\pi(a')).$$

Thus $\pi \circ j' = j \circ \pi$. But a', b' are in the hull, so

$$j(\pi(a')) = j(\pi| \text{Hull}(((\gamma')^+)^N))(j(a')).$$

All this junk is actually in N , so $j' \in N$. □

Thus we have shown that N contains lots of elementary embeddings, and so is nothing like the inner models under it like L or $L[U]$. In particular, there is no generalization of Scott's theorem to weak extender models for the supercompactness of δ . For example:

Corollary 13.14. Suppose that N is a weak extender model for the supercompactness of δ and $\kappa > \delta$ is a supercompact cardinal. Then

$$N \models \kappa \text{ is supercompact.}$$

Proof. Let $\lambda > \kappa$. Then there is an elementary embedding $j : V \rightarrow M$ which witnesses that κ is λ -supercompact and for every $\alpha > \lambda$, the restriction $j|_{(N \cap V_{\alpha+2})}$ is in N . Therefore $j|N \in N$. So N thinks that κ is λ -supercompact. □

13.4 The HOD dichotomy

Let N be a weak extender model for the supercompactness of δ .

Definition 13.15. A set is *ordinal-definable* if it is definable from parameters in **Ord**. It is *hereditarily ordinal-definable* if its transitive closure is ordinal-definable. The class **HOD** consists of all hereditarily ordinal-definable sets.

It's pretty easy to see that **HOD** is an inner model of ZFC; the proof is essentially the same as the proof that L is. Moreover, $L \subseteq \text{HOD}$. A major open problem is to show that $N \subseteq \text{HOD}$; this would imply that N has all the good properties of L and none of the bad ones. In fact, a conjecture of Woodin says that we can even have $N = \text{HOD}$.

But first we need to show that **HOD** is not like L in the sense that there is no analogue of 0^\sharp .

Definition 13.16. Let κ be a regular cardinal. We say that κ is a *strongly measurable cardinal in HOD* if there is a $\lambda < \kappa$ such that $(2^\lambda)^{\text{HOD}} < \kappa$, and **HOD** thinks that there is no partition of $\{\alpha < \kappa : \text{cof } \alpha = \omega\}$ into λ many stationary sets which are stationary in V .

This seems a little silly, because **HOD** is a model of ZFC and so by Solovay's stationary theorem, **HOD** thinks there is a partition of $\{\alpha < \kappa : \text{cof } \alpha = \omega\}$ into λ many sets (even κ many sets), but the point is that those sets will not actually be stationary in V .

We should check that if a cardinal is measurable in **HOD** then it is measurable in **HOD**. This is not nearly as tautological as it sounds.

Lemma 13.17. Let κ be a strongly measurable cardinal in **HOD**. Then there is a stationary set $S \subseteq \{\alpha < \kappa : \text{cof } \alpha = \omega\}$ such that $S \in \text{HOD}$ and there is no partition of S into multiple sets which are stationary in V .

Proof. Suppose not, and let λ be so large that we cannot partition $A = \{\alpha < \kappa : \text{cof } \alpha = \omega\}$ into λ many sets which are stationary in V .

Lemma 13.18. There is a binary tree $T \in \text{HOD}$ of height $\lambda + 1$ and a sequence $(S_t : t \in T)$ such that:

1. $S_\square = A$. Here \square is the empty sequence.
2. For every $t \in T$, S_t is a stationary subset of A in V .
3. For every $t \in T$, $t + 0$ and $t + 1$ are in T . Here $+$ is concatenation.
4. For every $t \in T$, $S_t = S_{t+0} \cup S_{t+1}$, and $S_{t+0} \cap S_{t+1}$ is empty.
5. For every $t \in T$ such that the length β of t is a limit ordinal,

$$S_t = \bigcap_{\alpha < \beta} S_{t|\alpha}.$$

6. For every limit ordinal $\beta \leq \lambda$ and every vertex $t : \beta \rightarrow 2$, if $t \notin T$ but for every $\alpha < \beta$, $t|\alpha \in T$, then $\bigcap_{\alpha < \beta} S_{t|\alpha}$ is not stationary.

Intuitively, the binary tree $(S_t : t \in T)$ partitions A into smaller and smaller sets which are stationary in V , splitting each set into two each time.

Proof. Since **HOD** is a model of choice we can use transfinite recursion. At stage $\alpha + 1$, choose a splitting $S_t = S_{t+0} \cup S_{t+1}$ for each t of length α , which exists by our contradiction hypothesis.

The hard work is at limit stages. Let β be a limit stage. Then

$$A = \bigcup_t \bigcap_{\alpha < \beta} S_{t|\alpha}$$

where t ranges over all $t \in \text{HOD}$ which is a branch of length β in T . Since $2^\beta < \kappa$ and the intersection of $< \kappa$ clubs is a club, so there is a branch t such that

$$S_t = \bigcap_{\beta < \alpha} S_{t|\alpha}$$

is stationary. For every such branch, put it in t and keep the induction going. \square

Let t be a branch of T of length $\lambda + 1$. Then let $S^\alpha = S_{t|\alpha} \setminus S_{t|\alpha+1}$. Since $S_{t|\alpha}$ split into two stationary sets, $S_{t|\alpha+1} = S_{(t|\alpha)=0}$ and $S_{(t|\alpha)+1}$, S^α is stationary if $\alpha \leq \lambda$. So $(S^\alpha : \alpha \leq \lambda)$ is a partition of A into λ stationary sets in V , a contradiction. \square

Lemma 13.19. Let κ be a strongly measurable cardinal in **HOD**. Then **HOD** thinks that κ is a measurable cardinal.

Proof. Using the previous lemma, let F be the club filter

$$F = \{X \subseteq S : \text{There is a club } c \text{ in } \kappa \text{ such that } c \cap S \subseteq X\}.$$

Let $F' = F \cap \mathbf{HOD}$. Then elements of F' are in \mathbf{HOD} , yet F and \mathbf{HOD} are ordinal-definable; therefore $F' \in \mathbf{HOD}$. Since F is a κ -complete filter on 2^S in V , F' is a κ -complete filter on 2^S in \mathbf{HOD} .

We claim that \mathbf{HOD} thinks that F' is an ultrafilter, which implies that \mathbf{HOD} thinks that κ is measurable. Suppose that $S_0 \in \mathbf{HOD}$ and $S_0 \subseteq S$. Let $S_1 = S \setminus S_0$. Since S is stationary in \mathbf{HOD} , at least one of the S_j is too, say S_0 . But then S_1 cannot be stationary in \mathbf{HOD} , for if it were then $S = S_0 \cup S_1$ would partition S into sets which are stationary in V , a contradiction. So $S_0 \in F'$, so \mathbf{HOD} thinks F' is a ultrafilter. \square

This is beginning to look dangerously like the situation with L and 0^\sharp . If κ is an uncountable cardinal in V and 0^\sharp exists, then L thinks that κ is an inaccessible cardinal, even if κ was actually a singular cardinal. Something similar *apparently* happens for \mathbf{HOD} .

Definition 13.20. Let κ be a regular cardinal and $\eta > \kappa$ an ordinal. We say that κ is an η -extendible cardinal if there is a $\lambda > \kappa + \eta$ and an elementary embedding

$$j : V_{\kappa+\eta} \rightarrow V_\lambda$$

such that $\text{crt } j = \kappa$. An *extendible cardinal* is one which is η -extendible for every η .

It follows straight from the definitions that every extendible cardinal is supercompact, strong, measurable, etc. Moreover, by universality, every extendible cardinal κ remains extendible in any weak extender model for the supercompactness of any cardinal δ such that $\kappa \geq \delta$. Intuitively, every extendible cardinal is so ridiculously huge that fragments of the universe all start to blur together.

Theorem 13.21 (HOD Dichotomy Theorem; Woodin). Suppose that δ is an extendible cardinal. Then exactly one of the following is true:

1. **HOD** is a weak extender model for the supercompactness of δ .
2. For every regular cardinal $\kappa > \delta$, κ is strongly measurable in **HOD**.

This looks at least superficially like the Jensen covering lemma and the 0^\sharp dichotomy. In fact, if **HOD** is a weak extender model for the supercompactness of δ , then by a previous lemma about such models, then **HOD** is close to V above δ . However, Woodin actually thinks that the **HOD** Dichotomy Theorem is not actually a dichotomy; i.e. that **HOD** is actually a weak extender model for the supercompactness of every extendible cardinal in V .

Conjecture 13.22 (HOD Conjecture; Woodin). For every extendible cardinal δ , **HOD** is a weak extender model for the supercompactness of δ .

Even though the **HOD** Dichotomy Theorem is a dichotomy, we can use it to prove lots of things about **HOD**. For example:

Corollary 13.23. If there is an extendible cardinal, then there is a measurable cardinal in **HOD**.

Proof. Let δ be an extendible cardinal. If **HOD** is a weak extender model for the supercompactness of δ , then δ is measurable in **HOD** and we can go home. Otherwise, the **HOD** Dichotomy Theorem implies that there is a proper class of cardinals which are strongly measurable in **HOD**. \square

Corollary 13.24. Let δ be an extendible cardinal. If there is a regular cardinal $\gamma_0 > \delta$ which is not strongly measurable in **HOD**, then **HOD** is close to V above δ .

To prove the **HOD** Dichotomy Theorem, we use the following lemma, which says that we can keep reflecting regular cardinals which are not strongly measurable in **HOD** higher and higher until we have a proper class of them.

Lemma 13.25. Suppose that δ is an extendible cardinal and $\gamma_0 > \delta$ is a regular cardinal which is not strongly measurable in **HOD**. Then there is a proper class of regular cardinals above γ_0 which are not strongly measurable in **HOD**.

Proof. Let $\alpha > \lambda_0 > \gamma_0$ be such that V_{λ_0} is a Σ_2 -substructure of V . This exists because we can always restrict to Σ_2 sentences and use reflection, then blow up to a sufficiently large cardinal and take transitive collapse. Suppose that $\lambda_0 = \text{card } V_{\lambda_0}$. Since **HOD** has a Σ_2 definition, we have

$$\mathbf{HOD}^{V_{\lambda_0}} = \mathbf{HOD} \cap V_{\lambda_0}.$$

Since δ is an extendible cardinal, there is an elementary embedding

$$j : V_{\alpha+1} \rightarrow V_{j(\alpha)+1}$$

such that $\text{crt } j = \delta$ and $j(\delta) > \alpha$. For every $\lambda < \gamma_0$, if $\text{card}^{\mathbf{HOD}} 2^\lambda < \gamma_0$ then there is a partition $(S_\alpha : \alpha < \lambda)$ of $S = \{\alpha < \gamma_0 : \text{cof } \alpha = \omega\}$ into stationary sets, such that $(S_\alpha : \alpha < \lambda) \in \mathbf{HOD}$. By elementarity, this remains true in V_{λ_0} , hence in $V_{j(\lambda_0)}$. \square

Proof of the HOD Dichotomy Theorem. Let $\gamma_0 > \delta$ be a regular cardinal which is not strongly measurable in **HOD**. By the lemma, there is a proper class of regular cardinals which are not strongly measurable in **HOD**. We must show that **HOD** is a weak extender model for the supercompactness of δ .

Let $\lambda > \delta$. It suffices to prove the following lemma:

Lemma 13.26. There is a δ -complete, normal fine ultrafilter U on $\mathcal{P}_\delta(\lambda)$ such that $\mathcal{P}_\delta(\lambda) \cap \mathbf{HOD} \in U$ and $U \cap \mathbf{HOD} \in \mathbf{HOD}$.

Let $\lambda_0 > \lambda$ be such that $\text{card } V_{\lambda_0} = \lambda_0$. By a previous lemma, there is a regular cardinal $\gamma_0 > \text{card}^{\mathbf{HOD}} 2^{\lambda_0}$ which is not strongly measurable in **HOD**.

Since γ_0 is not strongly measurable in **HOD** and $\text{card}^{\mathbf{HOD}} 2^{\lambda_0} < \gamma_0$, there is a partition $(S_\alpha : \alpha < \lambda_0) \in \mathbf{HOD}$ of

$$S^{\gamma_0} = \{\alpha < \gamma_0 : \text{cof } \alpha = \omega\}$$

such that for every $\alpha < \lambda_0$, S_α is stationary in V .

Using reflection, let $\lambda_1 > \gamma_0$ be such that V_{λ_1} is a Σ_2 substructure of V and $\text{card } V_{\lambda_1} = \lambda_1$. Therefore

$$\mathbf{HOD}^{V_{\lambda_1}} = \mathbf{HOD} \cap V_{\lambda_1}.$$

In particular,

$$(S_\alpha : \alpha < \lambda_0) \in \mathbf{HOD}^{V_{\lambda_1}}.$$

Since δ is an extendible cardinal, there is an elementary embedding

$$j : V_{\lambda_1+1} \rightarrow V_{j(\lambda_1)+1}$$

such that $\text{crt } j = \delta$ and $j(\delta) > \lambda_1$. So we may define

$$(T_\alpha : \alpha < j(\lambda_0)) = j(S_\alpha : \alpha < \lambda_0)$$

which is a partition in $\mathbf{HOD} \cap V_{j(\lambda_1)}$ of

$$S^{j(\gamma_0)} = \{\alpha < j(\gamma_0) : \text{cof } \alpha = \omega\}$$

into $j(\lambda_0)$ sets which are stationary in V .

We now have the ordinals

$$\delta < \lambda < \lambda_0 < \gamma_0 < \lambda_1 < j(\gamma_0).$$

Moreover, $j(\gamma_0)$ is a regular cardinal and $\text{crt } j < \gamma_0$, so $j(\gamma_0) > \sup j''\gamma_0$. Let Z be the set of all $\alpha < j(\lambda_0)$ such that $T_{j(\alpha)} \cap \sup j''\gamma_0$ is stationary in $\sup j''\gamma_0$.

Lemma 13.27. $Z = j''\lambda_0$.

Proof. We first show $j''\lambda_0 \subseteq Z$. Let $\alpha < \lambda_0$ and $C \subseteq \sup j''\gamma_0$ a club. Let D be the cofinal set $\{\alpha < \gamma_0 : j(\alpha) \in C\}$. Then \overline{D} is a club. Moreover, since j is continuous at points of cofinality ω ,

$$\overline{D} \cap S^{\gamma_0} = D \cap S^{\gamma_0}.$$

But $S_\alpha \subseteq S^{\gamma_0}$ is stationary in V , so there is a $\beta \in S_\alpha \cap D = S_\alpha \cap \overline{D}$. Thus $j(\beta) \in T_{j(\alpha)} \cap C$.

Conversely, let $\alpha \in Z$. So $T_\alpha \cap \sup j''\gamma_0$ is stationary, hence meets the club

$$C = \overline{j''\gamma_0} \cap \sup j''\gamma_0.$$

Let $\beta \in T_\alpha \cap \sup j''\gamma_0 \cap C$. But $C \cap j(S^{\gamma_0}) = j''(S^{\gamma_0})^V$ so $\beta \in j''(S^{\gamma_0})^V$. Let $\beta' \mapsto \beta$. So $j(\beta') \in T_\alpha$. The S_η are a partition, so let α' be such that $S_{\alpha'} \ni \beta'$. Then $\beta \in T_{j(\alpha')}$. But $\beta \in T_\alpha$, so $\alpha = j(\alpha')$. \square

In particular, $j''\lambda_0 \in \mathbf{HOD}$, since the definition of Z was from parameters in \mathbf{HOD} .

Since $\lambda_0 = \text{card } V_{\lambda_0}$,

$$\lambda_0 = \text{card}^{\mathbf{HOD}}(\mathbf{HOD} \cap V_{\lambda_0}).$$

In particular, \mathbf{HOD} thinks there is a bijection $\lambda_0 \rightarrow \mathbf{HOD} \cap V_{\lambda_0}$. Thus

$$j''(\mathbf{HOD} \cap V_{\lambda_0}) = j(\pi)''(j''\lambda_0).$$

We proved that $j(\pi)$ and $j(\lambda_0)$ are in \mathbf{HOD} . Therefore $j''(\mathbf{HOD} \cap V_{\lambda_0}) \in \mathbf{HOD}$.

Now let

$$U = \{X \subseteq \mathcal{P}_\delta(\lambda) : j''\lambda \in j(X)\}.$$

Since $j''(\mathbf{HOD} \cap V_{\lambda_0}) \in \mathbf{HOD}$ and $\lambda < \lambda_0$, U is definable in \mathbf{HOD} . Therefore $U \cap \mathbf{HOD} \in \mathbf{HOD}$. It follows that $\mathcal{P}_\delta(\lambda) \cap \mathbf{HOD} \in U$. \square

13.5 V is Ultimate L

We have all the pieces in play to give the ultimate generalization of L , assuming of course that several open problems are solved.

Definition 13.28. A set $A \subseteq \mathbb{R}$ is said to be *universally Baire* if for every compact Hausdorff space X and every continuous function $f : X \rightarrow \mathbb{R}$, there is an open set $U \subseteq X$ such that the symmetric difference $U \Delta f^{-1}(A)$ is meager in X .

Thus a universally Baire set has the property of Baire in every compact Hausdorff space, even those of extremely large cardinality. This is a ridiculously strong property.

Definition 13.29. An inner model N is said to be *Ultimate L* if there is a proper class of Woodin cardinals in N and, for every Σ_2 sentence φ such that $N \models \varphi$, there is a universally Baire set A such that $\mathbf{HOD}^{L(A, \mathbb{R})} \models \varphi$.

The first hint that Ultimate L might be something special is that it decides lots of sentences:

Theorem 13.30 (Woodin). If N is Ultimate L then $N \models$ the continuum hypothesis, and $N = \mathbf{HOD}^N$.

But the real significance of Ultimate L is that it cannot be broken by forcing.

Theorem 13.31 (Woodin). Suppose that \mathbb{P} is a forcing notion for $V = \text{Ultimate } L$. Then V actually is Ultimate L , and forcing by \mathbb{P} is the identity map.

If we believe that $V = \text{Ultimate } L$ is actually a true axiom, then this means that forcing fails completely, and the spectre of independence is essentially “cured,” as long as we also believe that sufficiently large cardinals are consistent.

There are two problems with the axiom $V = \text{Ultimate } L$, however.

First, it’s not clear that such an axiom is consistent; we should be able to prove its consistency from large cardinal axioms. For L and $L[U]$ we proved consistency by exhibiting a model of $V = L$ or $V = L[U]$. The former always existed, while the latter existed pending large cardinal axioms.

Secondly, we want $V = \text{Ultimate } L$ to have large cardinals. This is why we threw out L and $L[U]$ in the first place! If we are in the bad case of the **HOD** Dichotomy, everything goes to hell, as if N is Ultimate L , then $N = \mathbf{HOD}^N$, yet $V \neq \mathbf{HOD}$, so N is nothing like V .

Conjecture 13.32 (the Ultimate L Conjecture; Woodin). Suppose that δ is an extendible cardinal. Then there is an Ultimate L which is a weak extender model for the supercompactness of δ .

This is a Σ_1 statement in the language of arithmetic, so it must be either true or false. If it is true and there is an extendible cardinal, then we have found our Ultimate L and we can go home. By the universality theorem, Ultimate L has all large cardinals; it rules out forcing, and decides all statements about \mathbb{R} and other “small” sets. In particular, the continuum hypothesis is very likely true in Platonic reality if the Ultimate L Conjecture is true.

Moreover, the Ultimate L Conjecture implies the **HOD** Conjecture, showing that the **HOD** Dichotomy Theorem is not actually a dichotomy, and among other amusing consequences, showing that ZF alone can prove that Berkeley cardinals do not exist. Thus several major open problems in set theory would be solved in one fell swoop.

If the Ultimate L Conjecture is false, then Woodin's program ends in failure.